Глава 1

Диференциално смятане на функции на няколко променливи

1.1 Разстояние и норма в $\mathbb{R}^n$

Дефиниция на крайномерно пространство. В настоящата глава ще се занимаваме със свойствата на функциите на няколко променливи. Преди да започнем обаче тяхното изучаване, трябва да обърнем внимание на множествата, върху които те са дефинирани. Когато имаме функция на едно променливо $f(x)$, тя е определена за x принадлежащо на някакво подмножество (обикновено интервал) на реалната права $\mathbb{R}$. При функция на две променливи $f(x, y)$ стойността на функцията зависи от числовите стойности на двете променливи x и y, при това взети в определен ред - ясно е, че ако разменим техните места, получаваме друга стойност на функцията. Така, можем да кажем, че $f(x, y)$ е дефинирана върху някакво множество, състоящо се от *наредени двойки реални числа.* Аналогично функцията на три променливи $f(x, y, z)$ е дефинирана върху множество от *наредени тройки реални числа*, и т.н. Това ни довежда до следната дефиниция, играеща основна роля в по-нататъшните ни разсъждения:

Определение. *Под* *n-мерно евклидово пространство* $\mathbb{R}^n$ *ще разбираме* *множеството от всички наредени n-торки* $P = (x_1, x_2, \ldots, x_n)$ *от реални числа. Числата* $x_1, x_2, \ldots, x_n$ *ще наричаме съответно първа, втора,..., n-та координата на точката* P.

Пространството $\mathbb{R}^2$ от всички наредени двойки (x, y) се нарича равнина. При $n = 3$ получаваме тримерното пространство $\mathbb{R}^3$ на всички наредени тройки (x, y, z). Ще отбележим, че в линейната алгебра пространствата $\mathbb{R}^n$ се изучават от малко по-различна гледна точка: в тях се въвеждат т.н. линейни, или векторни операции - събиране на два елемента, и произведение на даден елемент с реално число. Понякога и ние ще използваме тези операции; в такъв случай, за да подчертаем наличието на такива операции, ще наричаме елементите на $\mathbb{R}^n$ вектори, и ще използваме за тях означения от вида $\vec{x}$, $\vec{y}$ и т.н.

Горната дефиниция е частен случай на възприетата в теорията на множествата операция произведение на две или повече множества. Ако $M_1, \ldots, M_n$ са някакви множества, тяхното произведение $M_1 \times M_2 \times \ldots \times M_n$ се определя като множеството от всички наредени n-торки $(m_1, m_2, \ldots, m_n)$ с $m_1 \in M_1$,...,$m_n \in M_n$. Тогава $\mathbb{R}^n$ може да се определи като произведение $\mathbb{R} \times \ldots \times \mathbb{R}$ на n екземпляра на реалната права $\mathbb{R}$. Така може да бъдат дефинирани и някои подмножества на евклидовите пространства: ако $[a, b]$ и $[c, d]$ са интервали в $\mathbb{R}$, тяхното произведение $[a, b] \times [c, d]$ представлява правоъгълник в $\mathbb{R}^2$ със страни, успоредни на координатните оси. Произведението на три интервала $[a, b] \times [c, d] \times [e, f]$ е правоъгълен паралелопипед в $\mathbb{R}^3$, и т.н.

Забележка. Дадения подход към многомерните пространства (към който ние ще се придържаме и в бъдеще) притежава следния недостатък - в него привилегирована роля играят координатните оси на променливите $x_1, \ldots, x_n$. В приложенията на анализа към природните науки такива привилегировани координатни системи рядко могат да са посочат; така например, в обичайното тримерно пространство ние не можем да предпочетем едни направления пред други. Ще изложим по-общия подход към въпроса, възприет в линейната алгебра:

Нека H е линейно пространство, т.е. пространство, в което са дефинирани векторните операции: събиране на елементи и умножение на елемент с реално число. Казваме, че пространството H е n-мерно (има размерност n), ако всеки $n + 1$ вектора в него са линейно зависими (някаква тяхна линейна комбинация с ненулеви коефициенти е нула), но съществуват n на брой линейно независими вектори. Да изберем линейно независима система от вектори $\vec{e}_1, \ldots, \vec{e}_n$ в H. Тогава всеки вектор $\vec{x} \in H$ притежава единствено разлагане от вида $\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i$. Ние можем да съпоставим на всеки вектор $\vec{x}$ n-торката от реални числа $x_1, \ldots, x_n$, с което получаваме взаимно еднозначно изображение на H в $\mathbb{R}^n$, запазващо линейните операции. В частност, на всяка функция $f(\vec{x})$ върху H ние съпоставяме функцията $f(x_1, \ldots, x_n)$, зависеща от n числови аргумента. Важно е да се помни обаче, че това съпоставяне зависи от избора на координатните вектори $\vec{e_1}, \ldots, \vec{e_n}$; по-голямата (и по-важната) част от дефинираните по-нататък понятия не зависят от избора на координатната система, и ние ще трябва да проверяваме това във всеки конкретен случай.

Най-често едно линейно пространство H се разглежда заедно с дадено скаларно произведение в него $\langle \vec{x}, \vec{y} \rangle \in \mathbb{R}$, $\vec{x}, \vec{y} \in H$. В този случай казваме, че H е *евклидово пространство*. В този случай обикновено избираме векторите $\vec{e_1}, \ldots, \vec{e_n}$ така, че да образуват *ортонормирана система* (виж по-долу).

Разстояние и норма в $\mathbb{R}^n$. По-нататък основна роля за нас ще играе понятието *евклидово разстояние между две точки в* $\mathbb{R}^n$. При $n = 1$, т.е. върху правата $\mathbb{R}$, разстоянието между точките с координати x и y е $\varrho(x, y) = |x - y|$. Разстоянието между две точки в равнината или в пространство с по-голям брой измерения може да се намери по елементарно-геометричен път с помощта на Питагоровата теорема. Нека $P = (x_1, x_2, \ldots, x_n)$, $Q = (y_1, y_2, \ldots, y_n)$. Тогава т. нар. евклидово разстояние между точките P и Q се дава с формулата

$$\varrho_2(P, Q) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

По нататък ние ще използваме не толкова дефиницията на това разстояние, колкото неговите основни свойства:

1/ $\varrho_2(P, Q) \geq 0$ за всеки две точки $P,\ Q \in \mathbb{R}^n$; $\varrho_2(P, Q) = 0$ тогава и само тогава, когато $P = Q$.

2/ Винаги $\varrho_2(P, Q) = \varrho_2(Q, P)$.

3/ (неравенство на триъгълника) За всеки три точки $P,\ Q,\ R \in \mathbb{R}^n$ имаме

$$\varrho_2(P, R) \leq \varrho_2(P, Q) + \varrho_2(Q, R).$$

Свойствата 1/ и 2/ са очевидни. Свойството 3/ лесно се интерпретира геометрически: ако разгледаме триъгълника с върхове в точките P, Q и R, то сумата от дължините на страните PQ и QR не надминава дължината на страната PR. Този факт се доказва в елементарната геометрия; разбира се, интересно е и неговото аналитично доказателство - виж твърдение1 и задача 1.

Разстоянието $\varrho_2(P,Q)$ не е единственото разстояние в $\mathbb{R}^n$, притежаващо свойствата 1/ - 3/. За илюстрация на това, да си представим, че живеем в град, разделен на правоъгълни квартали, движейки се по улиците (виж чертежа). Лесно се вижда, че всеки от най-кратките пътища (такива има много), съединяващ точките с координати $P = (x,y)$ и $Q = (x',y')$, има дължина $\varrho_1(P,Q) = |x - x'| + |y - y'|$. Аналогично разстояние може да се дефинира и в $\mathbb{R}^n$ за произволно n с формулата

$$\varrho_1(P,Q) = \sum_{i=1}^{n} |x_i - y_i| \,.$$

Очевидно разстоянието ϱ_1 също притежава свойствата 1/ - 3/.

Друго разстояние, отново удовлетворяващо 1/ - 3/, може да се дефинира с формулата:

$$\varrho_\infty(P,Q) = \sup_{i=1,\dots,n} |x_i - y_i| \,.$$

(за обяснение на индексите виж задачи 1. - 3.).

Норма в $\mathbb{R}^n$. Да разгледаме сега $\mathbb{R}^n$ като векторно пространство, и да въведем понятието евклидова норма на вектор от $\mathbb{R}^n$: ако $\vec{x} == (x_1, x_2, \dots, x_n)$, то

$$||\vec{x}||_2 = \sqrt{\sum_{i=1}^{n} (x_i)^2}.$$

Така въведената норма притежава свойствата:

$1'$/ $||\vec{x}||_2 \geq 0$; $||\vec{x}||_2 = 0$ тогава и само тогава, когато $\vec{x} = \vec{0}$.

$2'$/ $||\lambda\, \vec{x}||_2 = |\lambda|\, ||\vec{x}||_2$ за всеки $\lambda \in \mathbb{R}$ и $\vec{x} \in \mathbb{R}^n$.

$3'$/ (неравенство на триъгълника) За всеки два вектора $\vec{x},\ \vec{y} \in \mathbb{R}^n$ имаме $||\vec{x} + \vec{y}||_2 \leq ||\vec{x}||_2 + ||\vec{y}||_2$.

Нормата и разстоянието са очевидно свързани: имаме $\varrho_2(P,Q) = \left|\left|\vec{PQ}\right|\right|_2 = \left|\left|\vec{P} - \vec{Q}\right|\right|_2$, където с $\vec{P}$ и $\vec{Q}$ означаваме радиус-векторите на точките P и Q. В такъв случай ще казваме, че нормата $||\cdot||_2$ поражда разстоянието ϱ_2. Лесно се вижда, че разстоянието ϱ_1, дефинирано в забележката по-горе, се поражда от нормата

$$||\vec{x}||_1 = \sum_{i=1}^{n} |x_i| \,,$$

а разстоянието ϱ_∞ - от нормата

$$||\vec{x}||_\infty = \sup_{i=1,\dots,n} |x_i|\,,$$

(Проверете, че тези норми също удовлетворяват горните условия!)

Твърдение 1. *Нека нормата* $||\vec{x}||$ *в* $\mathbb{R}^n$ *удовлетворява условията* $1'/$ - $3'/$. *Тогава породеното от нея разстояние* $\varrho(P,Q) = \left|\left|\vec{P}-\vec{Q}\right|\right|$ *удовлетворява условията 1/ - 3/.*

Наистина, 1/ веднага следва от $1'/$. Условието 2/ се получава от $2'/$ при $\vec{x} = \vec{P} - \vec{Q}$ и $\lambda = -1$. Накрая, използвайки $3'/$, имаме

$$\left|\left|\vec{P}-\vec{R}\right|\right| = \left|\left|\left(\vec{P}-\vec{Q}\right) + \left(\vec{Q}-\vec{R}\right)\right|\right| \le \left|\left|\vec{P}-\vec{Q}\right|\right| + \left|\left|\vec{Q}-\vec{R}\right|\right|,$$

т.е. условието 3/. ■

В по-нататъшните ни разглеждания под норма $||\vec{x}||$ и разстояние $\varrho(P,Q)$ в ще разбираме евклидовата норма и евклидовото разстояние (освен ако изрично е казано противното). Добрите геометрични свойства на тази норма се дължат на факта, че тя е породена от скаларното произведение в $\mathbb{R}^n$.

Скаларно произведение. Нека H е линейно пространство. Казваме, че в H е дефинирано *скаларно произведение*, ако на всеки два вектора $\vec{x}, \vec{y} \in H$ е съпоставено реалното число $\langle \vec{x}, \vec{y} \rangle$, удовлетворяващо съотношенията:

$1''$ (линейност): За всеки $\vec{x}_1, \vec{x}_2, \vec{x}, \vec{y} \in H$ и число $\lambda \in \mathbb{R}$ имаме

$$\langle \vec{x}_1 + \vec{x}_2, \vec{y} \rangle = \langle \vec{x}_1, \vec{y} \rangle + \langle \vec{x}_2, \vec{y} \rangle\ ,\ \langle \lambda\vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{y} \rangle\,.$$

$2''$ (симетричност): За всеки $\vec{x}, \vec{y} \in H$ имаме

$$\langle \vec{x}, \vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle\,.$$

$3''$ (положителна определеност): За всеки *ненулев* вектор $\vec{x} \in H$ имаме

$$\langle \vec{x}, \vec{x} \rangle > 0.$$

В пространството $\mathbb{R}^n$ стандартното скаларно произведение се дава с формулата $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$, където $\vec{x} = (x_1, \dots, x_n)$, $\vec{y} = (y_1, \dots, y_n)$.

То може да се изрази и геометрично: $\langle \vec{x}, \vec{y} \rangle = ||x|| \, ||y|| \cos \angle (\vec{x}, \vec{y})$, където $\angle (\vec{x}, \vec{y})$ означава ъгъла между векторите $\vec{x}$ и $\vec{y}$. Лесно се вижда, че двете определения съвпадат - виж зад. 7. Разбира се, това не е единственото скаларно произведение в $\mathbb{R}^n$.

Ще изложим накратко някои основни свойства на скаларните произведения:

Твърдение 2.*(Неравенство на Коши-Шварц-Буняковски): За всеки два вектора $\vec{x}, \vec{y} \in H$ е в сила неравенството*

$$|\langle \vec{x}, \vec{y} \rangle| \leq \sqrt{\langle \vec{x}, \vec{x} \rangle} \; \sqrt{\langle \vec{y}, \vec{y} \rangle}.$$

Равенството се достига само в случая, когато векторите $\vec{x}$ и $\vec{y}$ са колинеарни.

Доказателство. Да разгледаме квадратния тричлен

$$p(t) = \langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle = At^2 + 2Bt + C,$$

където $A = \langle \vec{y}, \vec{y} \rangle$, $B = \langle \vec{x}, \vec{y} \rangle$, $C = \langle \vec{x}, \vec{x} \rangle$. От свойство $3''$ се вижда, че $p(t) \geq 0$ за всички стойности на t. Следователно $p(t)$ не може да има два различни реални корена - иначе в интервала между тях неговите значения биха били отрицателни. Оттук се вижда, че неговата дискриминанта $\Delta = 4\left(B^2 - AC\right)$ е по-малка или равна на нула, т.е. $B^2 \leq AC$.

Остава да изследваме кога се достига равенството. В такъв случай дискриминантата D е равна на нула, което означава, че $p(t)$ има точно един реален корен - да го означим с t_0. Имаме

$$p\left(t_0\right) = \langle \vec{x} + t_0\vec{y}, \vec{x} + t_0\vec{y} \rangle = 0,$$

откъдето пак по свойство $3''$ получаваме $\vec{x} + t_0\vec{y} = 0$, или, с други думи, $\vec{x} = -t_0\vec{y}$. ■

Определение. (норма, породена от скаларното произведение):*Норма на вектора $\vec{x} \in H$ ще наричаме числото*

$$||\vec{x}|| = \sqrt{\langle \vec{x}, \vec{x} \rangle}.$$

Твърдение 3.*Нормата, породена от скаларното произведение, удовлетворява условията $1'/$ - $3'/$.*

Доказателство. Свойствата $1'$ и $2'$ са очевидни. Свойството $3'$ (неравенството на триъгълника) следва от неравенството на Коши-Шварц-Буняковски:

$$||\vec{x}+\vec{y}||^2 = \langle \vec{x}+\vec{y}, \vec{x}+\vec{y}\rangle = \langle \vec{x}, \vec{x}\rangle + 2\langle \vec{x}, \vec{y}\rangle + \langle \vec{y}, \vec{y}\rangle =$$

$$= ||x||^2 + 2\langle \vec{x}, \vec{y}\rangle + ||y||^2 \le ||x||^2 + 2\,||x||\,||y|| + ||y||^2 = (||x|| + ||y||)^2 .$$

■

Нека H е n-мерно линейно пространство със скаларно произведение. Тогава в него може да се избере ортонормирана база, т.е. вектори $\vec{e}_1, \ldots, \vec{e}_n \in H$ такива, че $\langle \vec{e}_i, \vec{e}_j\rangle = 0$ при $i \neq j$ и $||\vec{e}_i|| = 1$ за всяко i. За всяко $\vec{x} \in H$ е налице разлагането

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i \text{ , където } x_i = \langle \vec{x}, \vec{e}_i\rangle .$$

Лесно се вижда, че ако $\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i$ и $\vec{y} = \sum_{i=1}^{n} y_i \vec{e}_i$, то

$$\langle \vec{x}, \vec{y}\rangle = \sum_{i=1}^{n} x_i y_i.$$

Съпоставяйки на вектора $\vec{x}$ n-торката от числа $(x_1, \ldots, x_n)$, ние получаваме взаимно еднозначно линейно съответствие между пространството H и $\mathbb{R}^n$, като скаларното произведение в H преминава в описаното по-горе стандартно скаларно произведение в $\mathbb{R}^n$. Важно е да се отбележи, че такова отъждествяване може да се извърши по много различни начини: то зависи от избора на ортонормирания базис $\vec{e}_1, \ldots, \vec{e}_n$. При смяна на базиса координатите се преобразуват чрез умножение с подходяща ортогонална матрица - виж зад. 8.

Забележка. Скаларното произведение и разлагането по ортонормален базис се прилагат не само в разгледаните тук крайно-мерни пространства, но и в по-сложни случаи като например пространства от функции. Така например, то е основен инструмент при разлагането на функции в ред на Фурие (виж глава ???).

Упражнения.

1. Да определим при $p \geq 1$ нормата $||\vec{x}||_p$ в $\mathbb{R}^n$ с формулата $||\vec{x}||_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$. Докажете, че условията $1'/$ - $3'/$ са удовлетворени.

Упътване. За да докажете $3'/$ при $p > 1$, покажете, че

$$\left(||\vec{x}+\vec{y}||_p\right)^p = \sum_{i=1}^n |x_i + y_i|^p \leq \sum_{i=1}^n |x_i|\,|x_i+y_i|^{p-1} + \sum_{i=1}^n |y_i|\,|x_i+y_i|^{p-1}$$

и приложете за двете суми вдясно неравенството на Хьолдер (виж I. §2.10, зад. 10).

2. Докажете, че при $p \in (0,1)$ неравенството на триъгълника за $||\vec{x}||_p$ не се изпълнява.

3. Докажете, че $\lim_{p\to+\infty} ||\vec{x}||_p = ||\vec{x}||_\infty$.

4. Нека H е пространство със скаларно произведение, и $||\vec{x}||$ е нормата в H, породена от него. Докажете, че е в сила теоремата на Питагор: за всеки $\vec{x}, \vec{y} \in \mathbb{R}^n$, такива че $\langle \vec{x}, \vec{y} \rangle = 0$, имаме

$$||\vec{x}+\vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2 .$$

5. Докажете, че за всяка норма, определена от скаларно произведение, е в сила равенството на успоредника:

$$2\left(||\vec{x}||^2 + ||\vec{y}||^2\right) = ||\vec{x}+\vec{y}||^2 + ||\vec{x}-\vec{y}||^2 .$$

(сумата на квадратите на страните на успоредника е равна на сумата на квадратите на неговите диагонали).

6. Използувайки резултата на зад. 5, докажете, че при $p \neq 2$ нормата $||\vec{x}||_p$ не се поражда от никое скаларно произведение в $\mathbb{R}^n$.

7. Докажете, че аналитичната и геометричната дефиниция на скаларно произведение в $\mathbb{R}^n$ съвпадат.

Упътване. От свойствата на дефинираното по аналитичен път скаларно произведение следва, че

$$\langle \vec{x}, \vec{y} \rangle = \frac{1}{2}\left(||\vec{x}||^2 + ||\vec{y}||^2 - ||\vec{x}-\vec{y}||^2\right) .$$

Да разгледаме тригълника, образуван от векторите $\vec{x}$ и $\vec{y}$, приложени в една и съща точка; страните му са равни на $||\vec{x}||$, $||\vec{y}||$ и $||\vec{x}-\vec{y}||$. Прилагайки косинусовата теорема за този триъгълник, виждаме, че дясната част на равенството съвпада с геометричната дефиниция на скаларното произведение.

8. Нека H е n-мерно векторно пространство със скаларно произведение, и $\vec{e}_1, \ldots, \vec{e}_n$ е ортонормирана система от вектори в H, а $\vec{f}_1, \ldots, \vec{f}_n$ е друга система от n вектора, изразяваща се чрез предходната с формулите

$$\vec{f}_j = \sum_{i=1}^{n} a_{i,j}\vec{e}_i \ , \ j = 1, \ldots, n.$$

а/ Докажете, че системата $\vec{f}_1, \ldots, \vec{f}_n$ е ортонормирана точно тогава, когато матрицата $\{a_{i,j}\}$ е ортогонална. Една $n \times n$ - матрица $A = \{a_{i,j}\}_{i=1,\ldots,n, j=1,\ldots,n}$ се нарича ортогонална, ако редовете и образуват ортонормирана система в $\mathbb{R}^n$, т.е.

$$\sum_{i=1}^{n} a_{i,j_1} a_{i,j_2} = \begin{cases} 0, & j_1 \neq j_2 \\ 1, & j_1 = j_2 \end{cases} .$$

б/ Покажете, че ако $\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i$ и $\vec{x} = \sum_{j=1}^{n} \tilde{x}_j \vec{f}_j$, то връзката между предишните коефициенти $\{x_i\}$ и новите $\{\tilde{x}_j\}$ се дава с формулите

$$x_i = \sum_{j=1}^{n} a_{i,j} \tilde{x}_j.$$

1.2 Топология и сходимост в $\mathbb{R}^n$

Както знаем, анализът върху реалната права се базира на отворените интервали в $\mathbb{R}$ и понятието околност (симетрична околност) на точка. За да развием анализа в многомерно пространство, ще трябва да въведем аналогични обекти в него, като имаме пред вид, че геометрията на подмножествата на многомерното пространство е много по-сложна от тази на едномерното.

Наличието на разстояние в $\mathbb{R}^n$ ни позволява да дефинираме понятието кръгова околност на дадена точка: под кръгова околност на точката $P \in \mathbb{R}^n$ с радиус r разбираме отворен кръг (евентуално кълбо) с център P и радиус r, т.е. множеството от всички точки Q от $\mathbb{R}^n$, намиращи се на разстояние по-малко от r от точката P. Означаваме:

$$B(P,r) = \{Q : \varrho(Q,P) < r\}.$$

В едномернияя случай това е симетричен интервал с център P и радиус r.

Нека сега M е подмножество на $\mathbb{R}^n$. Ще разделим точките от $\mathbb{R}^n$ на три класа в зависимост от местоположението им спрямо M.

Дефиниция. *Точката $P \in \mathbb{R}^n$ се нарича вътрешна за M, ако съществува $r > 0$ такова, че $B(P,r) \subset M$.*

С други думи, точката е вътрешна за едно множество, ако тя се съдържа в него заедно с някоя своя кръгова околност.

Множеството на всички вътрешни точки на M се бележи с $\dot{M}$.

Дефиниция. *Точката $P \in \mathbb{R}^n$ се нарича външна за M, ако съществува $r > 0$ такова, че $B(P,r) \cap M = \emptyset$.*

С други думи, точката е вътрешна за M, ако тя се съдържа в допълнението му заедно с някоя своя кръгова околност, т.е. принадлежи на $\mathbb{R}^n \dot{\setminus} M$.

Оставащите точки ще наричаме контурни:

Дефиниция. *Точката $P \in \mathbb{R}^n$ се нарича контурна за M, ако тя не е нито вътрешна, нито външна за M.*

Множеството от контурните точки се нарича контур на M и се бележи с $b\ M$ или с $b(M)$.

Горната дефиниция не е конструктивна; по-добро описание на контура на едно множество се дава в следното твърдение:

Твърдение 1. *Една точка е контурна за M точно тогава, когато всяка нейна кръгова околност се пресича и с M, и с $\mathbb{R}^n \backslash M$.*

Доказателство. Ако точката P не е нито вътрешна, нито външна за M, то за всяко $r > 0$ кръговата околност $B(P, r)$ не може да се съдържа нито в M, нито в $\mathbb{R}^n \backslash M$. Това означава, че тя ще има общи точки и с $\mathbb{R}^n \backslash M$, и с M.

Обратно, ако всяка кръгова околност на P има общи точки с M и с $\mathbb{R}^n \backslash M$, то очевидно никоя кръгова околност на P не може да се съдържа в M или в $\mathbb{R}^n \backslash M$, т.е. P не е вътрешна или външна за M. ■

Следствие. *Контурите на множеството M и на неговото допълнение $\mathbb{R}^n \backslash M$ съвпадат.*

Забележка. Дадените по-горе дефиниции зависят от това в кое евклидово пространство разглеждаме множеството M. Така например, нека I е подинтервал на правата $\mathbb{R}$. Тогава неговата вътрешност в смисъл на горната дефиниция съвпада с множеството на вътрешните му точки в обичайния смисъл, т.е. съвпада с интервала, евентуално с изключени крайни точки. От друга страна, ако разгледаме същия интервал I като подмножество на оста x в равнината $\mathbb{R}^2$, то вътрешността му е празна, и той се състои само от контурни точки (докажете!).

Ясно е, че външните за M точки не принадлежат на M; следователно точките, принадлежащи на M, са вътрешни или контурни. Това дава $M \subset \dot{M} \cup b(M)$. От друга страна, тези две множества не са длъжни да съвпадат; наистина, точките от контура на M могат и да не му принадлежат. Оттук получаваме една класификация на подмножествата на $\mathbb{R}^n$, която играе основна роля по-нататък:

Дефиниция. *Едно подмножество на $\mathbb{R}^n$ се нарича <u>отворено</u> в $\mathbb{R}^n$, ако то не съдържа точки от контура си. Под <u>околност</u> на дадена точка ще разбираме отворено множество, което я съдържа.*

Ще разгледаме и другият краен случай:

Дефиниция. *Едно подмножество на $\mathbb{R}^n$ се нарича <u>затворено</u> в $\mathbb{R}^n$, ако то съдържа всички точки на контура си.*

Ще отбележим, че в общия случай едно множество може да съдържа някои точки от контура, а други да не съдържа. С други думи, "повечето" множества в $\mathbb{R}^n$ не са нито отворени, нито затворени.

Както видяхме по-горе, контурите на едно множество и неговото допълнение съвпадат. Следователно, ако M не съдържа нито една точка от bM, то $\mathbb{R}^n \backslash M$ ще съдържа всички точки на bM, и обратно. Оттук получаваме:

Твърдение 2. *Едно подмножество на $\mathbb{R}^n$ е отворено точно тогава, когато неговото допълнение в $\mathbb{R}^n$ е затворено.*

Забележка. Лесно се вижда, че множеството $\dot{M}$ от вътрешните точки на M е отворено – очевидно това е най-голямото отворено множество, съдържащо се в M. Множеството $\overline{M} = M \cup bM$ (което се нарича затворена обвивка на M) е затворено, защото допълнението му съвпада с външността на M. Всяко затворено множество, което съдържа M, ще съдържа и $\overline{M}$ (докажете). С други думи, $\overline{M}$ е най-малкото затворено множество, съдържащо M.

Пример. Отвореният кръг е отворено множество.

Доказателство. Нека P е произволна точка от отворения кръг $B(P_0, r)$ с център P_0 и радиус r. Това означава, че $\varrho(P_0, P) < r$. Да изберем ε такова, че $0 < \varepsilon < r - \varrho(P_0, P)$; ще докажем, че $B(P, \varepsilon) \subset B(P_0, r)$. Наистина, нека $Q \in B(P, \varepsilon)$. Тогава по неравенството на триъгълника

$$\varrho(P_0, Q) \leq \varrho(P_0, P) + \varrho(P, Q) < \varrho(P_0, P) + \varepsilon = r.$$

■

По същия начин се доказва, че затвореният кръг, т.е. множеството от точки P, за които $\varrho(P_0, P) \leq r$, е затворено множество. И в двата случая контура на кръга съвпада с неговата гранична окръжност (или сфера,...), т.е. с множеството $S(P_0, r) = \{P : \varrho(P_0, P) = r\}$.

Сходимост на редици от точки в $\mathbb{R}^n$.

Дефиниция. *Казваме, че редицата $\{P_k\}_{k=1,2,\dots}$ от точки на $\mathbb{R}^n$ е сходяща към точката P_0 (пишем $P_k \to P_0$, или $\lim_{k\to\infty} P_k = P_0$), ако*

$$\varrho(P_k, P_0) \ \to \ 0.$$

По-подробно, за всяко положително ε съществува номер ν, така че за всяко естествено $k > \nu$ да имаме $\varrho(P_k, P_0) < \varepsilon$. Казано по друг начин: Всяка кръгова околност на граничната точка съдържа всички членове на редицата освен краен брой (или: всички от известно място нататък).

Сходимостта в $\mathbb{R}^n$ може лесно да се сведе към сходимост на числови редици. Наистина, нека точката P_k да е с координати $\left(x_1^k, \ldots, x_n^k\right)$, и съответно P_0 - с $(x_1^0, \ldots, x_n^0)$. Тогава имаме

Теорема 3. *Редицата* $\{P_k\}_{k=1,2,\ldots}$ *клони към* P_0 *тогава и само тогава, когато за всяко* $i = 1, 2, \ldots, n$ *редицата* $\left\{x_i^k\right\}$ *от* i*-тите координати клони по* k *към* i*-тата координата* x_i^0 *на* P_0.

Доказателство. Нека $\varrho(P_k, P_0) \to 0$. Тогава за всяко $i = 1, 2, \ldots, n$ имаме:

$$0 \leq \left|x_i^k - x_i^0\right| = \sqrt{\left|x_i^k - x_i^0\right|^2} \leq \sqrt{\sum_{j=1}^{n} \left|x_j^k - x_j^0\right|^2} = \varrho(P_k, P_0)$$

и по лемата за полицаите $\lim_{k\to\infty} \left|x_i^k - x_i^0\right| = 0$.

Обратно, ако за всяко i имаме $\left|x_i^k - x_i^0\right| \to 0$, след повдигане в квадрат, сумиране, и коренуване, получаваме, че $\sqrt{\sum_{i=1}^{n} \left|x_i^k - x_i^0\right|^2} \to 0$.
■

Забележка. Лесно се проверява, че горната теорема остава в сила, ако вместо евклидовото разстояние $\varrho(P, Q)$ в $\mathbb{R}^n$ вземем някое от разстоянията $\varrho_1(P, Q)$ или $\varrho_\infty(P, Q)$, разгледани в предния параграф. Може да се докаже и по-силно твърдение: за произволна норма в $\mathbb{R}^n$, удовлетворяваща условията $1'$ - $3'$, сходимостта на редици от точки относно разстоянието, определено от тази норма, съвпада с определената по-горе покоординатна сходимост. (Виж задачи 6-10 от §3).

От теоремата веднага се вижда, че сходимостта в $\mathbb{R}^n$ е съвместима с линейните операции. По-точно, имаме:

Следствие. Ако $P_k \to P_0$ и $Q_k \to Q_0$ в $\mathbb{R}^n$, и λ_k е числова редица, клоняща към λ_0, то

$$P_k + Q_k \to P_0 + Q_0 \quad , \quad \lambda_k\, P_k \to \lambda_0\, P_0.$$

Теорема 3 ни дава възможност да пренесем условието на Коши за сходимост, доказано в I §1.6 за числови редици, към редици от точки в $\mathbb{R}^n$.

Дефиниция. *Казваме, че редицата $\{P_k\}_{k=1,2,\dots}$ от точки на $\mathbb{R}^n$ удовлетворява условието на Коши, ако за всяко $\varepsilon > 0$ съществува число ν такова, че при $k > \nu$, $l > \nu$ да имаме $\varrho(P_k, P_l) < \varepsilon$.*

Теорема 4. *Една редица от точки на $\mathbb{R}^n$ е сходяща тогава и само тогава, когато удовлетворява условието на Коши.*

Доказателство. В едната посока твърдението лесно се доказва непосредствено от дефиницията за сходимост. Нека $P_k \to P_0$. Да вземем произволно $\varepsilon > 0$ и да изберем ν такова, че при $k > \nu$ да имаме $\varrho(P_k, P_0) < \varepsilon/2$. Тогава при $k > \nu$, $l > \nu$ имаме $\varrho(P_k, P_l) \leq \varrho(P_k, P_0) + \varrho(P_0, P_l) < \varepsilon$.

За доказателство на обратното твърдение ще използваме, че за редици от числа то вече е доказано (виж т. 8 на I §1.6). Ако условието на Коши е изпълнено, то за всяко $i = 1, \dots, n$ при $k > \nu$, $p > \nu$ имаме $\left|x_i^k - x_i^l\right| \leq \varrho(P_k, P_l) < \varepsilon$. Така редицата $\left\{x_i^k\right\}_{k=1,2,\dots}$ от i-тите координати на точките P_k удовлетворява условието на Коши и следователно е сходяща. Да означим границата и чрез x_i^0. Тогава от теорема 3 следва, че редицата P_k клони към точката $P_0 = (x_1^0, \dots, x_n^0)$. ■

Характеризация на затворените множества. Разглеждането на сходимостта позволява да се даде еквивалентна дефиниция на понятието затворено множество:

Теорема 5. *Едно подмножество на $\mathbb{R}^n$ е затворено точно тогава, когато съдържа границите на всички сходящи редици от свои елементи.*

Доказателство. Нека F е затворено подмножество на $\mathbb{R}^n$, и редицата P_k клони към P_0, като $P_k \in F$ за всяко естествено k. Трябва да докажем, че и $P_0 \in F$. Наистина, ако допуснем, че P_0 не принадлежи на F. Тъй като F е затворено, то съдържа всички свои контурни точки. Тогава P_0 може да бъде само външна за F. Следователно можем да намерим $r > 0$ такова, че $B(P_0, r)$ не пресича F. От друга страна, по дефиницията за сходимост при достатъчно голямо k имаме $P_k \in B(P_0, r)$, което противоречи на предположението, че всички точки на редицата са в F.

Обратно, нека F съдържа границите на всички сходящи редици от свои елементи; ще докажем, че F съдържа контурните си точки. Нека P_0 е произволна точка от контура на F, и да разгледаме кръговете $B\,(P_0, 1/k)$ с център P_0 и радиус $1/k$. От твърдение 1 следва, че всеки от тези кръгове има непразно сечение с F, и следователно съществува точка $P_k \in F \cap B\,(P_0, 1/k)$. Тъй като $\varrho\,(P_k, P_0) < 1/k$, то $\varrho\,(P_k, P_0) \to 0$, и следователно $P_k \to P_0$. От друга страна, тъй като всички точки P_k са в F, то от предположението следва, че и $P_0 \in F$, което трябваше да се докаже. ■

Забележка. От доказателството се вижда, че ако добавим към дадено множество M границите на всички сходящи редици от негови елементи, ще получим точно затворената му обвивка $\overline{M}$.

Точки на сгъстяване и подредици. Както и в едномерния случай, можем да въведем понятието точка на сгъстяване на дадена редица от точки на $\mathbb{R}^n$:

Дефиниция. *Казваме, че точката P_0 е точка на сгъстяване на редицата $\{P_k\}_{k=1,2,\ldots}$, ако за всяко $\varepsilon > 0$ кръговата околност $B\,(P_0, \varepsilon)$ на P_0 съдържа безкрайно много точки от редицата.*

Това означава, че всяка кръгова околност на P_0 съдържа точки с произволно големи номера. Следователно горната дефиниция може да се изкаже и по друг начин:

Точката P_0 е точка на сгъстяване на редицата $\{P_k\}_{k=1,2,\ldots}$, ако за всяко $\varepsilon > 0$ и естествено n съществува номер $m > n$ такъв, че $P_m \in B\,(P_0, \varepsilon)$.

Ясно е, че ако редицата $\{P_k\}$ притежава граница P_0, то точката P_0 е и единствена точка на сгъстяване на тази редица.* В общия случай една редица от точки може да има много точки на сгъстяване. Така, в I.§1.6 беше показано, че множеството $\mathbb{Q}$ от рационалните числа може да бъде номерирано, и така получената редица има за точки на сгъстяване всички реални числа. Не е трудно този пример да бъде пренесен в многомерния случай: да означим с $\mathbb{Q}^2$ множеството от всички точки в $\mathbb{R}^2$, за които и двете координати са реални. Тогава отново точките от $\mathbb{Q}^2$ могат да бъдат подредени в редица, и тази редица има за точки на

*Както ще покажем по-нататък, за ограничени редици е в сила и обратното твърдение.

сгъстяване всички точки от $\mathbb{R}^2$ (докажете!)

Подредици. Ще напомним понятието подредица на дадена редица: ако ни е дадена редицата $\{P_k\}_{k=1,2,\ldots}$ от точки на $\mathbb{R}^n$ и строго монотонно растяща система $k_1 < k_2 < \ldots < k_l < \ldots$ от естествени числа, ние може да си образуваме редицата $\{P_{k_l}\}_{l=1,2,\ldots}$, която наричаме подредица на дадената.

Твърдение 6. *Ако редицата $\{P_k\}_{k=1,2,\ldots}$ е сходяща към точката P_0, то всяка нейна подредица е сходяща към същата граница.*

Доказателство. Наистина, след като извън всяка кръгова околност на P_0 от вида $B(P_0, \varepsilon)$ се намират само краен брой членове на началната редица, то извън нея може да има най-много краен брой членове на подредицата. ■

Както и в едномерния случай, понятията подредица и точка на сгъстяване са тясно свързани:

Теорема 7. *Точката P_0 е точка на сгъстяване на редицата $\{P_k\}_{k=1,2,\ldots}$ точно тогава, когато съществува подредица $\{P_{k_l}\}_{l=1,2,\ldots}$, клоняща към P_0.*

Доказателство. Ако $\lim_{l\to\infty} P_{k_l} = P_0$, то всеки отворен кръг от вида $B(P_0, \varepsilon)$ съдържа безкрайно много членове на подредицата (всички освен краен брой) и следователно безбройно много членове на дадената редица.

Обратно, нека P_0 е точка на сгъстяване на редицата $\{P_k\}$. Можем да изберем номер k_1 такъв, че $P_{k_1} \in B(P_0, 1)$. Да разгледаме кръга $B(P_0, 1/2)$; тъй като според дефиницията на точка на сгъстяване той съдържа членове на редицата с произволно големи номера, ние можем да намерим естествено число k_2 такова, че $k_2 > k_1$ и $P_{k_2} \in B(P_0, 1/2)$. Продължавайки по същия начин, при намерени вече $k_1 < k_2 < \ldots < k_{l-1}$, ние можем да изберем k_l така, че $k_l > k_{l-1}$ и $P_{k_l} \in B(P_0, 1/l)$. С това получихме безкрайна редица $k_1 < k_2 < \ldots < k_l < \ldots$ от естествени числа такива, че $\varrho(P_{k_l}, P_0) < 1/l$, откъдето $\lim_{l\to\infty} P_{k_l} = P_0$. ■

Компактни множества. Теореми на Болцано-Вайерщрас и Хайне-Борел.

В тази точка ще покажем как теоремата на Болцано-Вайерщрас, известна в едномерния случай, се пренася в многомерното пространство.

Едно подмножество M на $\mathbb{R}^n$ ще наричаме <u>ограничено</u>, ако числовото множество $\{\varrho(0, P)\}_{P \in M}$ на разстоянията на всички точки от M до началото на координатите е ограничено отгоре. Лесно се вижда, че M е ограничено точно тогава, когато всички координати x_i, $i = 1, \ldots, n$ на всички негови точки $(x_1, \ldots, x_n)$ са ограничени по модул от някаква константа. Една редица $\{P_k\}$ ще наричаме ограничена, ако множеството от нейните точки е ограничено.

Теорема 8. (Болцано-Вайерщрас) *Всяка ограничена редица от точки на* $\mathbb{R}^n$ *притежава поне една точка на сгъстяване.*

Като се вземе пред вид теорема 7, се получава следната еквивалентна формулировка: Всяка ограничена редица от точки на $\mathbb{R}^n$ съдържа сходяща подредица.

Доказателство. При $n = 1$ теоремата е доказана в първата част на курса (виж I.§1.6), като се използва последователно разделяне на интервала, съдържащ членовете на редицата, на две равни части. Това доказателство може лесно да се извърши в многомерното пространство (виж теоремата на Хайне-Борел по-долу), но по-краткия начин е да използваме едномерния случай за доказателство на многомерния.

За удобство ще докажем случая $n = 2$. Нека $P_k = (x_k, y_k)$ е ограничена редица от точки на равнината. Тогава съществува константа C такава, че $|x_k|, |y_k| \leq C$. От ограничената числова редица $\{x_k\}_{k=1,2,\ldots}$ можем да изберем сходяща подредица $\{x_{k_l}\}_{l=1,2,\ldots}$; нека $\lim_{l\to\infty} x_{k_l} = x_0$. Да разгледаме сега съответната подредица от вторите координати $\{y_{k_l}\}_{l=1,2,\ldots}$. Тъй като тя също е ограничена, можем на свой ред да намерим нейна подредица $\left\{y_{k_{l_p}}\right\}_{p=1,2,\ldots} \to y_0$. Тъй като редицата $\left\{x_{k_{l_p}}\right\}_{p=1,2,\ldots}$ е подредица на редицата $\{x_{k_l}\}_{l=1,2,\ldots}$, то според твърдение 6 тя също клони към x_0. Тогава, по теорема 3, редицата от точки $P_{k_{l_p}} = \left(x_{k_{l_p}}, y_{k_{l_p}}\right)$ е подредица на дадената и клони към точката $P_0 = (x_0, y_0)$.

За редици от точки на $\mathbb{R}^n$ при по-големи стойности на n доказателството се извършва по същия начин. Разликата е, че трябва n пъти да се избират подредици на подредиците ..., като в крайна сметка отново се получава покоординатно сходяща подредица на дадената редица. ■

Следствие 9. Ако една ограничена редица от точки на $\mathbb{R}^n$ има само една точка на сгъстяване, то тя е сходяща към тази точка.

Доказателство. Нека $\{P_k\}$ е ограничена редица от точки и P_0 е нейната единствена точка на сгъстяване. Да допуснем, че $\{P_k\}$ не клони към P_0. Това означава, че съществува $\varepsilon > 0$ такова, че извън кръга $B(P_0, \varepsilon)$ остават безкрайно много членове на редицата. Повтаряйки конструкцията на теорема 7, ние можем да конструираме подредица P_{k_l} на дадената, чиито точки са извън $B(P_0, \varepsilon)$. За тази подподредица е приложима теорема 8, и ние може да намерим на свой ред нейна сходяща подредица $\left\{P_{k_{l_p}}\right\}_{p=1,2,\dots}$. Да означим границата и с $\widetilde{P}$. Ясно е, че $\widetilde{P} \neq P$. От друга страна, $\widetilde{P}$ е точка на сгъстяване на подредицата P_{k_l}, а следователно и на дадената редица $\{P_k\}$, което противоречи на условието на теоремата. ■

Дефиниция. *Едно подмножество на $\mathbb{R}^n$ наричаме <u>компактно</u>, ако то е ограничено и затворено.*

Като използваме това понятие, можем да дадем друга формулировка на теоремата на Болцано-Вайерщрас:

Теорема 10. *Едно множество $M \subset \mathbb{R}^n$ е компактно тогава и само тогава, когато от всяка редица от негови точки може да се избере подредица, клоняща към точка от M.*

Нека M е компактно и $\{P_k\}$ е редица от негови точки. Тъй като M е ограничено, то същото е вярно и за редицата $\{P_k\}$, и ние можем да изберем нейна сходяща подредица. По теорема 5 от затвореността на M следва, че границата на подредицата също принадлежи на M.

Обратно, нека M да притежава свойството за съществуване на сходяща подредица. Ще докажем, че M е ограничено и затворено. Ако M не е ограничено, то съществува редица от точки $P_k \in M$ такава, че $\varrho\left(\vec{0}, P_k\right) \to +\infty$. Да изберем нейна подредица P_{k_l}, сходяща към точката $P_0 \in M$. Тогава, от една страна, $\varrho\left(\vec{0}, P_{k_l}\right) \to \varrho\left(\vec{0}, P_0\right)$. От друга страна, тъй като $\varrho\left(\vec{0}, P_{k_l}\right)$ е подредица на $\varrho\left(\vec{0}, P_k\right)$, то $\varrho\left(\vec{0}, P_{k_l}\right) \to +\infty$, и полученото противоречие доказва ограничеността на M.

За да докажем, че M е затворено, да вземем произволна точка P_0 от контура bM на M, и да изберем, както в доказателството на теорема 5, редица P_k от точки на M, така че $P_k \to P_0$. Съществува нейна подредица P_{k_l}, сходяща към точка $P' \in M$. От твърдение 6 следва обаче, че $P' = P_0$ и следователно $P_0 \in M$. ■

Забележка. Горното твърдение е характерно за крайномерните пространства. В случая на безкрайномерни нормирани пространства то не е вярно, и затова съществуването на сходяща подредица се приема в този случай като дефиниция на понятието компактно множество.

Следната теорема отново е характерна за компактните множества. Тя обикновено не се изучава в едномерния анализ.

Нека M е подмножество на $\mathbb{R}^n$. Една фамилия $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ от подмножества на $\mathbb{R}^n$, индексирана с (крайното или безкрайно) множество A, наричаме покритие за M, ако M се съдържа в нейното обединение $\cup_{\alpha \in A} U_\alpha$. Покритието $\mathcal{U}$ наричаме отворено, ако множеството U_α е отворено за всяко $\alpha \in A$. Имаме:

Теорема 11. (теорема на Хайне-Борел) *От всяко отворено покритие на компактно множество може да се избере крайно подпокритие.*

По-точно, ако M е компактно множество в $\mathbb{R}^n$ и $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ е негово отворено покритие, то съществуват краен брой индекси $\alpha_1, \alpha_2, \ldots, \alpha_k \in A$ такива, че $M \subset U_{\alpha_1} \cup \ldots \cup U_{\alpha_k}$.

Доказателство. Първи етап. Ще докажем теоремата в случая, когато M е "n-мерен правоъгълник" в $\mathbb{R}^n$, т.е. произведение на крайни затворени интервали: $M = [a_1, b_1] \times \ldots \times [a_n, b_n]$. За удобство ще разгледаме случая $n = 2$. Нека е даден правоъгълника $M = [a, b] \times [c, d] \in \mathbb{R}^2$. Да допуснем противното – че M не може да бъде покрито с никаква крайна фамилия от елементи на $\mathcal{U}$. Да прекараме вертикалната права през средата на $[a, b]$ и хоризонталната права през средата на $[c, d]$; тези две прави разделят правоъгълника M на четири еднакви по форма правоъгълника. Тогава поне един от тях не може да бъде покрит с краен брой от елементи на $\mathcal{U}$; наистина, ако и четирите малки правоъгълника могат да бъдат покрити с краен брой множества от $\mathcal{U}$, то същото ще бъде вярно и за тяхното обединение - правоъгълника M. Следователно можем да изберем, и да означим с M_1 един от тях, за който не съществува крайно подпокритие. Тогава $M_1 = [a_1, b_1] \times [c_1, d_1] \subset M$, откъдето $[a_1, b_1] \subset [a, b]$ и $[c_1, d_1] \subset [c, d]$; при това $b_1 - a_1 = (b - a)/2$ и $d_1 - c_1 = (d - c)/2$. Да продължим процеса по-нататък, като разделим M_1 на четири еднакви правоъгълника и изберем един от тях, означен с M_2, и т.н. По този начин получаваме намаляваща редица от правоъгълници $M \supset M_1 \supset \ldots \supset M_k \supset \ldots$, като $M_k = [a_k, b_k] \times [c_k, d_k]$ и всеки от правоъгълниците M_k не може да бъде покрит с краен брой от елементи на $\mathcal{U}$. Очевидно числовата редица a_k е монотонно растяща и ограничена; да означим нейната граница с x_0. Редицата b_k е монотонно намаляваща и ограничена; тъй като $b_k - a_k = (b - a)/2^k \to 0$, то b_k клони към същата граница x_0. Оттук се вижда, че за всяко естествено k имаме $x_0 \in [a_k, b_k]$. Същото разсъждение е валидно и за вторите координати и доказва съществуването на число y_0 такова, че $y_0 \in [c_k, d_k]$ за всяко k. Очевидно точката $P_0 = (x_0, y_0)$ принадлежи на всеки от правоъгълниците M_k, $k = 1, 2, \ldots$. Тъй като $P_0 \in M$, то P_0 принадлежи на някое отворено множество U_{α_0} от покритието $\mathcal{U}$.

Според дефиницията на отворено множество съществува $e > 0$ такова, че $B(P_0, \varepsilon) \subset U_{\alpha_0}$. Лесно се вижда, че при достатъчно големи k ще имаме $M_k \subset B(P_0, \varepsilon)$ (това е изпълнено при $\sqrt{(b_k - a_k)^2 + (d_k - c_k)^2} < \varepsilon$). Така M_k е подмножество на U_{α_0}, което противоречи на допускането, че M_k не може да бъде покрито с краен брой от множествата U_α. Полученото противоречие доказва първия етап на доказателството.

В n-мерния случай доказателството се извършва по същия начин, с единствената разлика, че на всяка стъпка поредният n-мерен правоъгълник се разделя на 2^n части, и всички направени по-горе разсъждения остават в сила.

Втори етап. Като използваме първия етап, ще преминем към доказателството на теоремата в общия случай. Нека M е компактно (т.е. ограничено и затворено) множество в $\mathbb{R}^n$, и $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ е негово отворено покритие. Тъй като M е ограничено, можем да намерим n-мерен правоъгълник $\widetilde{M} = [a_1, b_1] \times \ldots \times [a_n, b_n]$ такъв, че $M \subset \widetilde{M}$. Да означим с $\mathcal{U}'$ фамилията от отворени множества, състояща се от множествата на $\mathcal{U}$ и отвореното множество $\mathbb{R}^n \backslash M$. Очевидно $\mathcal{U}'$ е покритие за $\widetilde{M}$, и по доказаното по-горе може да се намерят краен брой отворени множества $U_{\alpha_1}, \ldots, U_{\alpha_k}$ от $\mathcal{U}$, така че $\widetilde{M} \subset (\mathbb{R}^n \backslash M) \cup U_{\alpha_1} \cup \ldots \cup U_{\alpha_k}$. Това означава, че $M \subset U_{\alpha_1} \cup \ldots \cup U_{\alpha_k}$. ■

Упражнения.

1. Докажете, че:

а/ Обединение на произволна фамилия отворени множества е също отворено.

б/ Сечение на краен брой отворени множества е отворено.

2. Докажете, че:

а/ Сечение на произволна фамилия затворени множества е също затворено.

б/ Обединение на краен брой затворени множества е затворено.

Упътване. Използвайте твърдение 2 и резултата на задача 1.

3. Докажете, че множеството от точките на сгъстяване на дадена редица е затворено.

4. Нека множеството M не е компактно. Докажете, че заключението на теорема 11 не е вярно, т.е. съществува отворено покритие на M, от което не може да се избере крайно подпокритие.

Пояснение: Предположението, че M не е компактно, означава, че е налице една от двете възможности (или и двете): а/ M не е ограничено, или б/ M не е затворено. Разгледайте двете възможности поотделно.

1.3 Непрекъснатост на функции и изображения

Както и в първата част на анализа, при нас ще играят основна роля понятията функция и изображение. Нека D е подмножество на $\mathbb{R}^n$. Казваме, че е дадено изображение от D в $\mathbb{R}^m$, ако на всяка точка $x = (x_1, \dots, x_n)$ от D е съпоставена точката $f(x)$ от $\mathbb{R}^m$. В такъв случай се казва още, че f е функция с дефиниционна област D и със стойности в $\mathbb{R}^m$. Вместо $f(x)$ понякога се пише $f(x_1, \dots, x_n)$.

В случая $m = 1$, т.е. функцията f взема стойности в множеството на реалните числа, ще казваме, че f е числова, или още скаларна, функция. В многомерния случай ще казваме, че f е векторна функция. Всяка векторна функция може да бъде описана чрез числови функции; наистина, нека f е функция, дефинирана в $D \subset \mathbb{R}^n$ и вземаща стойности в $\mathbb{R}^m$. Да означим чрез $f_1(x), \dots, f_m(x)$ координатите на точката $f(x) \in \mathbb{R}^m$. Тогава $f_1(x), \dots, f_m(x)$ са числови функции върху D, които определят напълно функцията f. Тези функции се наричат координатни функции за изображението f.

Ще се занимаем с понятията граница на функция и непрекъснатост. Изложението почти дословно повтаря едномерния случай, изложен в част I, §1.8 и 1.10. Отново имаме две еквивалентни определения за граница на функция (и съответно две определения за непрекъснатост): - на Хайне и на Коши.

Определение на Хайне за граница на функция. *Дадена е функцията $f(P)$ с дефиниционна област $D \subset \mathbb{R}^n$ със значения в $\mathbb{R}^m$. Нека $P_0 \in \overline{D}$, и Q е точка от $\mathbb{R}^m$. Казваме, че функцията $f(P)$ клони към Q при $P \to P_0$, ако за всяка редица $\{P_k\}_{k=1,2,\dots}$ от точки на D, за която $P_k \to P_0$ имаме $f(P_k) \to Q$ в $\mathbb{R}^m$.*

Определение на Коши за граница на функция. *При горните условия казваме, че $f(P)$ клони към Q при $P \to P_0$, ако за всяко $\varepsilon > 0$ съществува $\delta > 0$ такова, че за всяка точка P от D, за която $\varrho(P, P_0) < \delta$, имаме $\varrho(f(P), Q) < \varepsilon$.*

Теорема 1. *Определенията на Хайне и Коши за граница на функция са еквивалентни.*

Доказателство. Ще започнем с по-лесната част - ще допуснем, че $f(P)$ клони към Q при $P \to P_0$ в смисъл на Коши, и ще докажем, че това е вярно и в смисъл на Хайне. Наистина, нека $\{P_k\}_{k=1,2,\dots}$ е редица от точки на D, клоняща към P_0. Ще докажем, че $f(P_k) \to Q$. Да изберем произволно $\varepsilon > 0$, и да вземем съответното $\delta > 0$ по определението на Коши. Тогава по определението на сходимост на редици от точки може да се намери ν такова, че $\varrho(P_k, P_0) < \delta$ за всяко $k > \nu$. Следователно при $k > \nu$ имаме $\varrho(f(P), Q) < \varepsilon$, което и трябваше да се докаже.

Доказателството в обратната посока се извършва чрез допускане на противното. Ще допуснем, че определението на Коши не е удовлетворено, и ще докажем, че не е изпълнено и определението на Хайне. Твърдението "$f(P)$ не клони в смисъл на Коши към Q при $P \to P_0$" означава, че съществува число $\varepsilon_0 > 0$ такова, че по-нататъшната част от определението не е изпълнена: т.е. за всяко $\delta > 0$ съществува точка $P_\delta \in D$ такава, че $\varrho(P_\delta, P_0) < \delta$, но $\varrho(f(P_\delta), Q) \geq \varepsilon_0$. За да сведем нещата до редици, да дадем на δ стойности $\delta_1 = 1, \delta_2 = 1/2, \dots, \delta_k = 1/k, \dots$ и да означим $P_k = P_{\delta_k}$. Тогава тези точки удовлетворяват условията $P_k \in D$, $\varrho(P_k, P_0) < 1/k$ и $\varrho(f(P_k), Q) \geq \varepsilon_0$. С други думи, редицата $\{P_k\}_{k=1,2,\dots}$ от точки на D клони към P_0, но редицата от функционалните стойности $f\{P_k\}$ не клони към Q. ■

Понятието граница на функция лесно се свежда към случая на граници на числови функции. Нека $f_1(P), \dots, f_m(P)$ са координатните функции на изображението $f(p)$, и нека $Q = (y_1, \dots, y_m)$. Тогава:

Твърдение 2. *Имаме* $\lim_{P \to P_0} f(P) = Q$ *тогава и само тогава, когато* $\lim_{P \to P_0} f_i(P) = y_i$ *за всяко* i *между* 1 *и* m.

Доказателство. По теорема 3 от §2 редицата $f(P_k)$ клони към Q точно тогава, когато $f_i(P_k) \to y_i$ за $i = 1, \dots, m$. Оттук, използвайки дефиницията на Хайне, получаваме твърдението. ■

От горното твърдение веднага следва, че при линейните операции - сума на векторни функции, и произведение на векторна функция с числова функция - може да се извършва граничен преход. По-точно, имаме:

Твърдение 3.

а/ Ако $f(P)$ *и* $g(P)$ *са векторни функции със стойности в* $\mathbb{R}^m$, *притежаващи граници при* $P \to P_0$, *то* $\lim_{P \to P_0}(f(P) + g(P)) = \lim_{P \to P_0} f(P) + \lim_{P \to P_0} g(P)$.

б/ Ако $\lambda(P)$ е числова функция, а $f(P)$ - векторна функция, и двете функции притежават граници при $P \to P_0$, то $\lim_{P\to P_0} \lambda(P)\, f(P) = \lim_{P\to P_0} \lambda(P)\ \lim_{P\to P_0} f(P)$.

Непрекъснатост на функции.

Определение. *Нека е дадена функцията $f(P)$ с дефиниционна област $D \subset \mathbb{R}^n$ и със значения в $\mathbb{R}^m$. Ще казваме, че $f(P)$ е непрекъсната в точката $P_0 \in D$, ако*

$$\lim_{P\to P_0} f(P) = f\left(P_0\right).$$

Двете еквивалентни определения на понятието граница на функция водят до две еквивалентни определения на непрекъснатостта:

Непрекъснатост по Хайне. *Функцията $f(P)$ е непрекъсната в точката P_0, ако за всяка редица $\{P_k\}_{k=1,2,\dots}$ от точки на D, за която $P_k \to P_0$ имаме $f\left(P_k\right) \to f\left(P_0\right)$ в $\mathbb{R}^m$.*

Непрекъснатост по Коши. *Функцията $f(P)$ е непрекъсната в точката P_0, ако за всяко $\varepsilon > 0$ съществува $\delta > 0$ такова, че за всяка точка P от D, за която $\varrho\left(P, P_0\right) < \delta$, имаме $\varrho\left(f(P), f\left(P_0\right)\right) < \varepsilon$.*

Разбира се, тези две определения са еквивалентни, и ние можем да използваме това определение, което е по-удобно в дадена конкретна ситуация. Така, от определението на Хайне и от твърдение 3 веднага следва:

Твърдение 4. *Сума на две непрекъснати векторни функции (със значение в едно и също пространство $\mathbb{R}^m$), както и произведение на непрекъсната скаларна функция с непрекъсната векторна функция, са също непрекъснати.*

Непрекъснатост на сложно изображение. Нека f е изображение с дефиниционна област $D \subset \mathbb{R}^n$ и стойности в $\mathbb{R}^m$, и g е изображение с дефиниционна област $E \subset \mathbb{R}^m$ и стойности в $\mathbb{R}^p$. Да предположим, че $f(D) \subset E$. Тогава можем да определим изображението F с дефиниционна област D и стойности в $\mathbb{R}^p$ с формулата $F(P) = F(g(P))$, $P \in D$. Изображението F се нарича сложно, или съставно, изображение. Казва се още, че изображението F е суперпозиция на изображенията f

и g. Както се вижда от следващата теорема, суперпозицията на две непрекъснати изображения е също непрекъснато изображение:

Теорема 5. *Нека функцията f е непрекъсната в точката $P_0 \in D$, и g е непрекъсната в $Q_0 = f(P_0)$. Тогава съставната функция F също е непрекъсната в P_0.*

Доказателство. Нека $\{P_k\}$ е редица от точки на D, клоняща към P_0. Тогава по дефиницията на Хайне редицата $Q_k = f(P_k)$ клони към Q_0 и следователно редицата $F(P_k) = g(Q_k)$ клони към $F(P_0) = g(Q_0)$. ■

Пример. Следващият пример показва, че непрекъснатостта по две променливи е нещо повече от непрекъснатостта по всяка от тях поотделно. Нека $f(x, y)$ е определена с формулата

$$f(x, y) = \frac{x\, y}{x^2 + y^2} \text{ при } (x, y) \neq (0, 0)\,,\; f(0, 0) = 0.$$

Тогава при всяко фиксирано y функцията $f(x, y)$ е непрекъсната като функция на променливата x, и при всяко фиксирано x - като функция на y. В точката $(0, 0)$ обаче функцията не е непрекъсната. Наистина, ако $P_n = (1/n, 1/n)$, то $f(P_n) \to 1/2$. Нещо повече, ако разгледаме редица от точки в $\mathbb{R}^2$, която клони към началото на координатите, намирайки се върху правата с уравнение $y = \lambda x$, то редицата от функционалните стойности клони към $\frac{\lambda}{1+\lambda^2}$, т.е. граничната стойност зависи от посоката, по която се стремим към точката.

Възможна е и по-сложна ситуация - една функция да има една и съща граница по всички прави, минаващи през дадена точка, и въпреки това да не бъде непрекъсната в тази точка - виж задача ????.

Характеризация на непрекъснатите функции. Казваме, че една функция е непрекъсната, ако тя е непрекъсната във всички точки на дефиниционната си област. Ние ще изложим една характеризация на непрекъснатите функции.

Нека $f(P)$ е функция с дефиниционна област $D \subset \mathbb{R}^n$ и със значения в $\mathbb{R}^m$. Нека M е подмножество на $\mathbb{R}^m$. Под <u>прообраз</u> на M относно f разбираме множеството от онези точки $P \in D$ такива, че $f(p) \in M$ (разбира се, това може да е и празното множество). Праобразът на M относно f ще означаваме с $f^{-1}(M)$.

Ще използваме едно обобщение на понятието отворено множество. Нека $D \subset \mathbb{R}^n$, и V е подмножество на D. Ще казваме, че множеството V е отворено в D, ако за всяка точка $P \in V$ съществува $\varepsilon > 0$, така че $B(P, \varepsilon) \cap D \subset V$. Ако D съвпада с цялото $\mathbb{R}^n$, получаваме известното вече определение на отворено множество.

Теорема 6. *Функцията $f(P)$ с дефиниционна област $D \subset \mathbb{R}^n$ и със значения в $\mathbb{R}^m$ е непрекъсната тогава и само тогава, когато праобразът на всяко отворено подмножество на $\mathbb{R}^m$ е отворен в D.*

Доказателство. Нека $f(P)$ е непрекъсната във всички точки на D, и нека $U \subset \mathbb{R}^m$ е отворено множество. Ще докажем, че $f^{-1}(U)$ е отворено в D. Наистина, нека $P_0 \in f^{-1}(U)$. Тогава $f(P_0)$ принадлежи на U, и по определението на отворено множество съществува $\varepsilon > 0$, така че $B(f(P_0), \varepsilon) \cap D \subset U$. По дефиницията на Коши за непрекъснатост можем да намерим $\delta > 0$, така че за всяко $P \in D$, за което $\varrho(P, P_0) < \delta$, имаме $\varrho(f(P), f(P_0)) < \varepsilon$. Преведено на езика на множествата, това означава, че $B(P_0, \delta) \cap D \subset f^{-1}(U)$, което и трябваше да се докаже.

Обратно, нека е дадено, че праобразът на всяко отворено множество чрез f е отворен в D. Да фиксираме произволна точка $P_0 \in D$. Ще докажем, че f е непрекъсната по Коши в P_0. Наистина, да вземем произволно $\varepsilon > 0$, и да означим с U отвореното множество $B(f(P_0), \varepsilon)$ в $\mathbb{R}^m$. Тогава $f^{-1}(U)$ е отворено в D, което означава, че съществува $\delta > 0$, така че $B(P_0, \delta) \cap D \subset f^{-1}(U)$. Това значи, че за всяко $P \in D$, удовлетворяващо $\varrho(P, P_0) < \delta$, е изпълнено $\varrho(f(P), f(P_0)) < \varepsilon$.
■

Свойства на непрекъснатите функции върху компактно множество. Теоремите на Вайерщрас за непрекъснати функции върху краен и затворен интервал лесно се пренасят за случая на непрекъснати числови функции върху крайни подмножества на $\mathbb{R}^n$:

Теорема 7. (Първа теорема на Вайерщрас). *Всяка непрекъсната числова функция, дефинирана върху компактно множество, е ограничена.*

Теорема 8. (Втора теорема на Вайерщрас). *Всяка непрекъсната числова функция, дефинирана върху компактно множество, достига най-голямата и най-малката си стойност.*

Доказателство. Нека числовата функция $f(P)$ е дефинирана и непрекъсната за $P \in D$, където D е компактно подмножество на $\mathbb{R}^n$. Да допуснем, че f не е ограничена, например, отгоре. Тогава съществува редица от точки $P_k \in D$ такива. че $f(P_k) \geq k$. Според §2.10 ние можем да изберем нейна сходяща подредица $P_{k_l} \to P_0 \in D$. Тогава от една страна $f(P_{k_l}) > k_l$ и следователно $f(P_{k_l}) \to +\infty$. От друга страна, от непрекъснатостта на f в точката P_0 следва, че $f(P_{k_l}) \to f(P_0)$, и полученото противоречие доказва първата теорема.

Нека сега $M = \sup_{P \in D} f(P)$. Тогава за всяко естествено k числото $M - 1/k$ вече не е горна граница за значенията на функцията, и ние може да намерим точка $P_k \in D$ такава. че $M - 1/k < f(P_k) \leq M$. Отново ще изберем сходяща подредица $P_{k_l} \to P_0$. Като извършим граничен преход в неравенствата

$$M - \frac{1}{k_l} < f(P_{k_l}) \leq M,$$

получаваме $M \leq f(P_0) \leq M$, т.е. $f(P_0) = M$.

Твърденията за ограниченост отдолу и за достигане на точната долна граница се получават аналогично. ■

За векторнозначни функции горните твърдения не се пренасят дословно, но се заместват от следващата теорема. За случая на числови функции тя е еквивалентна с теореми 6 и 7 (покажете!).

Теорема 9. *Образът на компактно множество при непрекъснато изображение е компактен.*

Доказателство. Нека $f(P)$ е непрекъсната функция, дефинирана върху компактното множество $D \subset \mathbb{R}^n$ и вземаща стойности в $\mathbb{R}^m$. Ще покажем, че нейният образ $f(D) \subset \mathbb{R}^m$, състоящ се от всички точки $f(P), p \in D$, е компактно подмножество на $\mathbb{R}^m$. По теорема §2.10 е достатъчно да покажем, че от всяка редица от точки $Q_k \in f(D)$ можем да изберем подредица, сходяща към точка от $f(D)$. Нека $Q_k = f(P_k)$. Избираме подредица P_{k_l} от точки на D, клоняща към $P_0 \in D$. Тогава от непрекъснатостта в точката P_0 получаваме, че $Q_{k_l} \to Q_0 = f(P_0)$. ■

Равномерна непрекъснатост. И това понятие се пренася от едномерния случай почти без изменения:

Определение. *Изображението f с дефиниционна област $D \subset \mathbb{R}^n$ се нарича <u>равномерно непрекъснато</u> в D, ако за всяко $\varepsilon > 0$ съществува*

$\delta > 0$, така че за всеки две точки P и Q от D, за които $\varrho(P, Q) < \delta$, имаме $\varrho(f(P), f(Q)) < \varepsilon$.

Теорема 10. (Теорема на Кантор). *Всяко изображение, дефинирано върху компактно множество и непрекъснато в него, е и равномерно непрекъснато.*

Доказателство. Да допуснем, че избражението f, дефинирано и непрекъснато върху компактното множество D, не е равномерно непрекъснато. Тогава съществува число $\varepsilon_0 > 0$ такова, че понататъшната част от определението не е изпълнена: т.е. за всяко $\delta > 0$ съществуват точки $P_\delta, Q_\delta \in D$ такава, че $\varrho(P_\delta, Q_\delta) < \delta$, но $\varrho(f(P_\delta), f(Q_\delta)) \geq \varepsilon_0$. Да дадем на δ стойности $\delta_k = 1/k, \ldots$ и да означим $P_k = P_{\delta_k}, Q_k = Q_{\delta_k}$. Тогава тези точки удовлетворяват условията $\varrho(P_k, Q_k) < 1/k$ и $\varrho(f(P_k), f(Q_k)) \geq \varepsilon_0$. Нека $\{P_{k_l}\} l = 1, 2, \ldots$ е сходяща подредица на редицата $\{P_k\}$. Да означим границата и с P_0. Очевидно редицата $\{Q_{k_l}\} l = 1, 2, \ldots$ клони към същата граница. Тогава $f(P_{k_l}) \to f(P_0)$, $f(Q_{k_l}) \to f(Q_0)$, и $f(P_{k_l}) - f(Q_{k_l}) \to \vec{0}$, което противоречи на допускането, че $\varrho(f(P_{k_l}), f(Q_{k_l})) \geq \varepsilon_0$. ■

Линейно свързани множества. Теорема за междинните стойности. В първата част на анализа се доказва следното свойство на непрекъснатите функции: нека $f(x)$ е непрекъсната функция, дефинирана върху интервал (вида на интервала не е от значение). Ако A и B са значенията на функцията в две точки от дефиниционния интервал, и C е число между A и B, то съществува точка, в която $f(x)$ взема стойност C. Теоремата не е вярна, ако пропуснем условието, че дефиниционата област е интервал; например, нека дефиниционното множество е обединение на два непресичащи се интервала, и $f(x)$ е равна на константата 0 върху първия, и на 1 - върху втория. Така определената функция е непрекъсната, но междинната стойност $1/2$ не се достига в никоя точка.

За да успеем да пренесем тази теорема в многомерния случай, трябва да можем да правим разграничение между двата случая за подмножества на $\mathbb{R}^n$. Общо казано, трябва да разделим множествата, които "се състоят от едно парче", от тези, които са "от няколко парчета". Това води до понятието линейна свързаност на множество. Това е множество, всеки две точки на което може да бъдат съединени с непрекъсната линия. Предварително че трябва да въведем понятието непрекъсната крива линия в $\mathbb{R}^n$:

Определение. *Под непрекъсната крива линия в $\mathbb{R}^n$ разбираме непрекъснато изображение на крайния и затворен интервал $[a, b]$ в пространството $\mathbb{R}^n$.*

Така, всяка непрекъсната крива в $\mathbb{R}^n$ се задава с векторна функция $\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$, $t \in [a, b]$. Със същия термин ще означаваме и множеството от стойностите на $\varphi(t)$: $\Gamma_\varphi = \{\varphi(t)\}_{t\in[a,b]}$. Точките $\varphi(a)$ и $\varphi(b)$ се наричат съответно начална и крайна точка на кривата Γ_φ.

Определение. *Едно множество $D \subset \mathbb{R}^n$ се нарича линейно свързано, ако за всеки две точки P, Q от D съществува непрекъсната крива Γ_φ, съдържаща се в D, с начална точка P и крайна точка Q.*

Например, всяко изпъкнало множество е линейно свързано. Ще напомним, че едно множество се нарича изпъкнало, ако то съдържа заедно с всеки две свои точки и отсечката, която ги свързва. Такива множества са кръга, правоъгълника, триъгълника и др. Отсечката, която свързва точките P и Q, може да се параметризира по формулата $\varphi(t) = (1-t)\,P + t\,Q$, и следователно представлява непрекъсната крива, съединяваща тези точки.

Обратно, обединението на два непресичащи се кръга не е свързано (докажете!),

Забележка. Друго, близко по характер свойство на множествата в $\mathbb{R}^n$ се дава в определението на свързано множество - виж задачи 12-15.

Теорема 11. *Нека f е непрекъсната функция върху линейно свързаното множество $D \subset \mathbb{R}^n$, P и Q са две точки от D, и $A = f(P)$, $B = f(Q)$. Нека C е число между A и B. Тогава съществува точка $L \in D$ такава, че $f(L) = C$.*

Доказателство. Ще сведем задачата към едномерния случай, доказан в I §1.11, следствие 2. Нека изображението $\varphi(t) :\ [a, b] \to \mathbb{R}^n$ да определя непрекъсната крива, лежаща в D и съединяваща точките P и Q. Да определим функцията на едно променливо $F(t)$ по формулата $F(t) = f(\varphi(t))$, $t \in [a, b]$. Тогава имаме $F(a) = A$, $F(b) = B$, и следователно съществува точка $\xi \in (a, b)$ такава, че $F(\xi) = C$. Нека $L = \varphi(\xi)$; тогава $f(L) = F(\xi) = C$. ■

Упражнения.

1. Докажете, че евклидовото разстояние $\varrho(P,Q)$ в $\mathbb{R}^n$ е равномерно непрекъснато по двете променливи едновременно, т.е. се явява непрекъсната функция на точката $(P,Q) \in \mathbb{R}^{2n}$.

Упътване. Докажете неравенството

$$|\varrho\,(P_1,Q_1) - \varrho\,(P_2,Q_2)| \leq \varrho\,(P_1,P_2) + \varrho\,(Q_1,Q_2)\,.$$

2. (Разстояние между точка и множество.) Нека P е точка и A е подмножество на $\mathbb{R}^n$. Разстоянието между P и A се определя по формулата

$$\varrho(P,A) = \inf_{Q\in A} \varrho(P,Q).$$

а/ Докажете, че $|\varrho(P,A) - \varrho(Q,A)| \leq \varrho(P,Q)$ и следователно $\varrho(P,A)$ е равномерно непрекъсната функция на P.

б/ Нека A е затворено. Докажете, че в горната дефиниция минимумът се достига, т.е. съществува $Q \in A$ такова, че $\varrho(P,A) = \varrho(P,Q)$.

3. (Разстояние между две множества.) Нека A и B са подмножества на $\mathbb{R}^n$. Да положим

$$\varrho(A,B) = \inf_{P\in A\ ,\ Q\in B} \varrho(P,Q).$$

а/ Докажете, че $\varrho(A,B) = \inf_{P\in A} \varrho(P,B) = \inf_{Q\in B} \varrho(Q,A)$.

б/ Докажете, че ако A и B са непресичащи се подмножества на $\mathbb{R}^n$, като A е компактно, а B е затворено, то $\varrho(A,B) > 0$.

4. Нека A и B са непресичащи се затворени подмножества на $\mathbb{R}^n$. Постройте непрекъсната в $\mathbb{R}^n$ функция $f(P)$ такава, че $f(P) = 0$ за $P \in A$ и $f(P) = 1$ за $P \in B$.

Упътване. Определете $f(P)$ с формулата:

$$f(P) = \frac{\varrho(P,A)}{\varrho(P,A) + \varrho(P,B)}.$$

5. Нека A и B са като в горната задача. Докажете, че те могат да бъдат отделени с непресичащи се околности, т.е. съществуват две отворени подмножества U, V на $\mathbb{R}^n$ такива, че $A \subset U$, $B \subset V$, и $U \cap V = \emptyset$.

Упътване. Положете $U = f^{-1}\left((-\infty, 1/2)\right)$, $V = f^{-1}\left((1/2, +\infty)\right)$.

6. Нека $||P||'$ е норма в $\mathbb{R}^n$, т.е. числова функция, удовлетворяваща условията 1'/ - 3'/ от §1. Докажете, че $||P||'$ е непрекъсната в точката $\vec{0}$.

Упътване. Нека $\vec{e}_1, \ldots \vec{e}_n$ са координатните вектори в $\mathbb{R}^n$. Тогава за точката P с координати $(x_1, \ldots, x_n)$ е в сила векторното равенство $P = x_1\vec{e}_1 + \ldots + x_n\vec{e}_n$, и от неравенството на триъгълника получаваме

$$||P||' \leq \sum_{i=1}^{n} |x_i|\, ||\vec{e}_i||' \leq \sqrt{\sum \left(||\vec{e}_i||'\right)^2}\ \sqrt{\sum x_i^2}.$$

7. Ако $||P||'$ е както в предната задача, докажете, че $||P||'$ е равномерно непрекъсната функция в $\mathbb{R}^n$.

Упътване. Покажете, че $\left|\,||P||' - ||Q||'\,\right| \leq ||P - Q||'$, и използвайте горното неравенство за $P - Q$.

8. Нека $||P||'$ е произволна норма в $\mathbb{R}^n$, удовлетворяваща условията 1'/ - 3'/, и нека $||P||$ е обичайната евклидова норма в $\mathbb{R}^n$. Докажете, че съществуват константи c, C, $0 < c \leq C$, така че за всяко $P \in \mathbb{R}^n$ да имаме

$$c\,||P|| \leq ||P||' \leq C\,||P||\,.$$

Упътване. Нека S е сфера в $\mathbb{R}^n$ с център в началото и радиус единица (т.е. множеството от точките в $\mathbb{R}^n$ с евклидова норма единица). Да означим съответно с c и C най-малката и най-голямата стойност на функцията $||P||'$ върху S (които съществуват според теоремите на Вайерщрас). Покажете, че те удовлетворяват горните неравенства.

Забележка. Казваме, че нормите $||P||$ и $||P||'$ са <u>еквивалентни</u>, ако съществуват константи c и C. за които неравенството от задачата винаги е удовлетворено. В тази терминология, твърдението на задачата може да се формулира така: Всяка норма в $\mathbb{R}^n$ е еквивалентна на евклидовата.

9. Нека $||P||$ и $||P||'$ са норми в $\mathbb{R}^n$, и $\varrho(P, Q)$, $\varrho'(P, Q)$ са породените от тях разстояния (виж §1.) Докажете, че $||P||$ и $||P||'$ са еквивалентни точно тогава, когато разсточнията ϱ, ϱ' пораждат една и съща сходимост на редици от точки в $\mathbb{R}^n$.

Упътване. Достатъчно е да докажете задачата за редици, клонящи към нулата. В едната посока твърдението следва непосредствено от дефиницията на еквивалентност. Обратно, нека за всяко естествено k съществува точка $P_k \in \mathbb{R}^n$ такава, че $||P_k||' > k\,||P_k||$. Тогава редицата $Q_k = \dfrac{1}{\sqrt{k}\,||P_k||} P_k$ удовлетворява условията $||Q_k|| \to 0$, $||Q_k||' \to +\infty$.

10. Нека $||P||'$ е произволна норма в $\mathbb{R}^n$, удовлетворяваща условията 1'/ - 3'/. Използвайки твърденията на задачи 8 и 9, докажете, че сходимостта относно $||P||'$ на редици от точки в $\mathbb{R}^n$ съвпада с покоординатната сходимост (виж теорема 3 от §2).

Пояснение. Казваме, че редицата $\{P_k\}$ клони към P_0 относно нормата $||\,||'$, ако $||P_k - P_0||' \to 0$ (в дефиницията за сходимост в §2 за тази цел е използвана евклидовата норма.)

11. Нека $f(x, y)$ е функцията в $\mathbb{R}^2$, дефинирана с формулите:

$$f(x,y) = \frac{x^2\,y}{x^4 + y^2} \text{ при } (x,y) \neq (0,0)\,,\; f(0,0) = 0.$$

Докажете, че $f(x, y)$ е не е непрекъсната в началото на координатната система, но е непрекъсната по всяка права, минаваща през началото.

Упътване. Всяка права през началото се параметризира с формулите $x = \lambda t$, $y = \mu t$, $t \in \mathbb{R}$, при подходящи λ и μ. Лесно се вижда, че функцията $f(\lambda t, \mu t)$, получена чрез заместване по тези формули, е непрекъсната функция на t. От друга страна, редицата $P_n = (1/n, 1/n^2)$ клони към $(0, 0)$, но $f(P_n) \to 1/2$.

12. Докажете, че подмножеството V на множеството D е отворено в D (според дефиницията, предхождаща теорема 6) тогава и само тогава, когато съществува отворено множество U, така че $V = D \cap U$.

13. Едно подмножество M на $\mathbb{R}^n$ се нарича свързано, ако не съществуват две непресичащи се отворени множества U и V, така че $M \subset U \cup V$ (С други думи, M не може да се разложи на две отворени в него подмножества.) Докажете, че всеки интервал е свързано подмножество на $\mathbb{R}$.

Упътване. Нека $I = [a,b] \subset U \cup V$, като $a \in U$, $b \in V$. Разгледайте $c = \sup_{x \in I \cap U} x$, и покажете, че c е контурна точка за U и V. Следователно, c не може да принадлежи нито на U, нито на V.

14. Докажете, че всяко линейно свързано множество е свързано.

Упътване. Да допуснем, че $M \subset U \cup V$, и $\varphi(t), t \in [a,b]$, е параметризация на непрекъсната крива, свързваща точка от U с точка от V. Тогава интервалът $[a,b]$ се представя като обединение на непресичащите се отворени множества $\varphi^{-1}(U)$ и $\varphi^{-1}(V)$.

15. Да означим с M графиката на функцията $f(x)$, зададена с формулата

$$f(x) = \sin \frac{1}{x} \text{ при } x \neq 0\,,\; f(0) = 0$$

(виж графиката в I, стр. 92.)

Докажете, че M е свързано, но не е линейно свързано.

16. Докажете, че всяко *отворено* свързано подмножество на $\mathbb{R}^n$ е и линейно свързано.

17. Докажете, че ако f е непрекъсната числова функция върху свързаното множество M, то за нея е валидна теоремата за междинните стойности (теорема 11).

Упътване. Да допуснем, че числата A и B са стойности на функцията f, $A < C < B$, и числото C не е стойност на f. Тогава ще имаме $M \subset f^{-1}\left((-\infty, C)\right) \cup f^{-1}\left((C, +\infty)\right)$.

18. Нека $f(p)$ е функция, дефинирана и непрекъсната върху отвореното и ограничено множество $D \subset \mathbb{R}^n$. Докажете, че f е равномерно непрекъсната върху D тогава и само тогава, когато тя се продължава до непрекъсната функция $\overline{f}(p)$ върху затворената обвивка $\overline{D}$ на D.

Забележка. В едната посока горното твърдение е непосредствено следствие на теорема 10; наистина, ако $\overline{f}(p)$ съществува, то тя е равномерно непрекъсната върху $\overline{D}$, и следователно същото е вярно и за $f(p)$. Трудната част на задачата е да се построи непрекъснато продължение на равномерно непрекъснатата функция $f(p)$ в точките от контура bD на D.

Упътване. Нека $P_0 \in bD$ и $P_k \to P_0$, $P_k \in D$. Докажете, че редицата $\{f\left(P_k\right)\}$ е фундаментална в смисъл на Коши и следователно сходяща. Покажете, че граничната стойност $\overline{f}\,(P_0)$ не зависи от избора на редицата $\{P_k\}$, клоняща към P_0. Докажете, че така продължената функция $\overline{f}\,(P)$ е непрекъсната върху $\overline{D}$.

1.4 Диференцируемост на функции и изображения. Диференциране на съставни функции.

Нека $f(x_1, \ldots, x_n)$ е функция, дефинирана в отвореното подмножество D на $\mathbb{R}^n$. Най-често е трудно да си представим нагледно поведението на тази функция; за тази цел можем да разгледаме зависимостта на функцията f от всяко едно променливо поотделно. Нека фиксираме точка $x^0 = (x_1^0, \ldots, x_n^0)$ от D, и да въведем следните (n на брой) функции на едно променливо:

$$\varphi_i(t) = f\left(x_1^0, \ldots, x_{i-1}^0, t, x_{i+1}^0, \ldots, x_n^0\right),$$

като функцията $\varphi_i(t)$ е дефинирана в околност на точката x_i^0. Функциите $\varphi_1, \ldots, \varphi_n$ ще наричаме частични функции, съответстващи на функцията на много променливи $f(x)$. На чертежа е дадена графиката на функцията $f(x,y) = x^2 - y^2$, т.е. множеството от точки $(x,y,z) \in \mathbb{R}^3$, за които $z = f(x,y)$, и лежащите върху нея графики на някои от частичните и функции.

(Тук е даден и друг начин за онагледяване на поведението на една функция на две променливи – чрез така наречените линии на ниво, ("хоризонтали") т.е. линиите, зададени с уравнението $f(x,y) = a$ при различните стойности на a. Този начин се използува например за представяне на релефа на местността върху физическите карти.)

Частни производни на функция. Можем да разгледаме производните на написаните по-горе частични функции (стига те да съществуват). С други думи, можем да диференцираме функцията $f(x)$ по една от променливите $x_1, \ldots, x_n$, като останалите променливи са фиксирани. Така стигаме до определението:

Под частна производна на функцията $f(x_1, \ldots, x_n)$ по променливата x_i в точката $x^0 = (x_1^0, \ldots, x_n^0)$ ще разбираме (ако съществува) границата

$$\frac{\partial f}{\partial x_i}\left(x^0\right) = \varphi_i'(x_i^0) =$$

$$= \lim_{h \to 0} \frac{f\left(x_1^0, \ldots, x_{i-1}^0, x_i^0 + h, x_{i+1}^0, \ldots, x_n^0\right) - f\left(x_1^0, \ldots, x_{i-1}^0, x_i^0, x_{i+1}^0, \ldots, x_n^0\right)}{h}.$$

Частната производна на $f(x)$ по x_i се означава още и с f'_{x_i}.

За нагледност ще повторим тази дефиниция за функция на две променливи $f(x,y)$:

$$\begin{aligned} f'_x(x_0,y_0) &= \lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h}, \\ f'_y(x_0,y_0) &= \lim_{k\to 0} \frac{f(x_0,y_0+k)-f(x_0,y_0)}{k}. \end{aligned}$$

Забележка. Трябва да се отбележи, че (за разлика от случая на едно променливо) наличието на частни производни, по всички променливи и във всяка точка, още не означава, че функцията е "добра". Да си припомним функцията, разгледана в примера в §3: $f(x,y) = \frac{x\ y}{x^2+y^2}$ при $(x,y) \neq (0,0)$, $f(0,0) = 0$. Лесно се вижда, че $f(x,y)$ притежава навсякъде частни производни по x и y; от друга страна, както беше показано там, функцията не е непрекъсната в $(0,0)$.

Формула за нарастването на функция на много променливи. За да обясним смисъла на изведеното по-долу равенство, отначало ще си припомним случая на функция на едно променливо. Нека $f(x)$ е функция, дефинирана в околност на точката $x_0 \in \mathbb{R}$ и диференцируема в тази точка. Знаем, че $f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$. Ако означим

$$\alpha(h) = \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0),$$

то ще получим, че $\lim_{h\to 0} \alpha(h) = 0$, и

$$f(x_0+h) = f(x_0) + f'(x_0)h + \alpha(h)h.$$

Това равенство лесно се интерпретира геометрически: ако заместим h с $x - x_0$, първите две събираеми отдясно дават уравнението на допирателната права към графиката на f в точката x_0, а третото събираемо е "остатък", намаляващ по-бързо от величината h. Ще изведем подобна формула за функция на няколко променливи.

Теорема 1 (формула за нарастването за функция на много променливи) *Нека $f(x_1,\ldots,x_n)$ е функция, дефинирана в околност на* *точката $x^0 = (x_1^0, \ldots, x_n^0)$, и нека всички частни производни f'_{x_i}, $i = 1, \ldots, n$, съществуват и са непрекъснати в тази околност. Тогава за достатъчно малки по норма вектори $h = (h_1, \ldots, h_n)$ е в сила равенството:*

$$(*) \qquad f(x+h) \ = \ f(x) \ + \ \sum_{i=1}^{n} \ f'_{x_i}(x) \ h_i \ + \ \sum_{i=1}^{n} \alpha_i(h) \ h_i,$$

където $\alpha_i(h)$ са функции, клонящи към нула при $h \to 0$.

Преди да докажем тази формула. ще обясним нейния смисъл, като я напишем по друг начин. Първите две събираеми отдясно се наричат линейна част на нарастването, а третото – остатък. Линейната част на функцията представлява приблизителен израз за стойностите на функцията близо до x, лесен за пресмятане и опериране. Остатъкът ни дава представа за точността на това приближение. Ако означим с $r(h)$ остатъка в горната формула:

$$r(h) = \sum_{i=1}^{n} \alpha_i(h) \ h_i,$$

то очевидно имаме $\lim\limits_{h \to 0} \dfrac{r(h)}{||h||} = 0$, и получаваме:

II форма на формулата за нарастването:

$$(**) \qquad f(x+h) \ = \ f(x) \ + \ \sum_{i=1}^{n} \ f'_{x_i}(x) \ h_i \ + \ r(h),$$

където $r(h) = o\,(||h||)$. [*]

Понякога е удобно горната формула да се записва по друг начин: ако означим $\Delta f = f(x+h) - f(x)$, и вместо h_i напишем Δx_i (нарастването на променливата x_i), имаме

$$\Delta f \ = \ \sum_{i=1}^{n} \ f'_{x_i}(x) \ \Delta x_i \ + \ o\,(||h||)\,.$$

[*] Двете форми на формулата за нарастването очевидно са равносилни; наистина, ако имаме формулата $(**)$, и означим $\alpha_i(h) = \frac{h_i}{||h||^2} \ r(h)$, то от нея се получава формулата $(*)$.

Доказателство на формулата за нарастването: Ще използваме теремата за крайните нараствания (виж I, §2.3): Ако функцията $\varphi(x)$ е непрекъсната в интервала $[x_0, x_0+h]$ и диференцируема във вътрешността му, то съществува число $\theta \in (0,1)$ такова, че $\varphi\left(x_0+h\right)-\varphi\left(x_0\right)=h \varphi^{\prime}\left(x_0+\theta h\right)$.

За яснота най-напред ще дадем доказателството за функция на две променливи. Имаме:

$$f\left(x+h, y+k\right)-f(x, y) \ =$$

$$= \ (f(x+h, y+k)-f(x, y+k)) \ + \ (f(x, y+k)-f(x, y)) \ =$$
$$= \ h \ f_x^{\prime}(x+\theta_1 h, y+k) \ + \ k \ f_y^{\prime}(x, y+\theta_2 k),$$

за подходящи $\theta_1, \theta_2 \in(0,1)$. Да означим

$$\alpha(h, k)=f_x^{\prime}(x+\theta_1 h, y+k)-f_x^{\prime}(x, y+k),\ \beta(h, k)=f_y^{\prime}(x, y+\theta_2 k)-f_y^{\prime}(x, y).$$

От непрекъснатостта на частните производни на f следва, че функциите $\alpha(h, k), \beta(h, k)$ клонят към нула при $h, k \rightarrow 0$. Като направим заместване, горното равенство добива търсения вид

$$f\left(x+h, y+k\right)-f(x, y) \ = \ h \ f_x^{\prime}(x, y)+k \ f_y^{\prime}(x, y)+h \ \alpha(h, k)+k \ \beta(h, k).$$

В общия случай на функция на n променливи $f\left(x_1, \ldots, x_n\right)$ доказателството се провежда по същия начин. Нека $h=\left(h_1, \ldots, h_n\right)$ е достатъчно малко по норма. Да означим $\Delta f=f(x+h)-f(x)$, и нека

$$\Delta_i f=f\left(x_1, \ldots, x_{i-1}, x_i+h_i, x_{i+1}+h_{i+1}, \ldots, x_n+h_n\right) \ -$$

$$- \ f\left(x_1, \ldots, x_{i-1}, x_i, x_{i+1}+h_{i+1}, \ldots, x_n+h_n\right).$$

Разглеждайки $f\left(x_1, \ldots, x_n\right)$ като функция само на променливото x_i (т.е. фиксирайки всички останали променливи), и прилагайки към тази функция теоремата за крайните нараствания, имаме

$$\Delta_i f=h_i f_{x_i}^{\prime}\left(x_1, \ldots, x_{i-1}, x_i+\theta_i h_i, x_{i+1}+h_{i+1}, \ldots, x_n+h_n\right), \ \theta_i \in(0,1).$$

Ако означим

$$\alpha_i(h)=f_{x_i}^{\prime}\left(x_1, \ldots, x_i+\theta_i h_i, x_{i+1}+h_{i+1}, \ldots, x_n+h_n\right)-f_{x_i}^{\prime}\left(x_1, \ldots, x_n\right),$$

то, поради непрекъснатостта на частните производни, имаме $\alpha_i(h) \to 0$ при $h \to 0$.

От друга страна, лесно се проверява, че $\Delta f = \sum_{i=1}^n \Delta_i f$, откъдето

$$\Delta f = \sum_{i=1}^n \left(f'_{x_i}(x_1, \ldots, x_n) + \alpha_i(h) \right) h_i = \sum_{i=1}^n f'_{x_i}(x) h_i + \sum_{i=1}^n \alpha_i(h)\, h_i. \blacksquare$$

Забележка. Само съществуването на частните производни (без тяхната непрекъснатост) не е достатъчно за валидността на доказаната формула; наистина, лесно се вижда, че ако една функция удовлетворява формулата за нарастването, то тя е непрекъсната. По-горе обаче ние видяхме пример на функция, притежаваща частни производни, която не е непрекъсната.

Определение на диференцируема функция. Както често се случва в математиката, за определение на даден клас от обекти се приема това тяхно свойство, което се използва при работата с тях. В случая това е формулата за нарастването. Така, (макар че това може да изглежда изкуствено), диференцируеми функции на няколко променливи се наричат тези функции, за които тя е удовлетворена. По-точно, имаме:

Определение. *Нека $f(x_1, \ldots, x_n)$ е функция на n променливи, определена в околност на точката $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Казваме, че функцията f е диференцируема в точката x, ако съществуват константи $A_1, \ldots, A_n$, както и функции $\alpha_1(h), \ldots, \alpha_n(h)$, клонящи към нула при $h \to 0$, така че за достатъчно малки по норма вектори $h = (h_1, \ldots, h_n)$ е в сила равенството*

$$f(x+h) = f(x) + \sum_{i=1}^n A_i\, h_i + \sum_{i=1}^n \alpha_i(h)\, h_i$$

В частност, веднага се вижда, че ако функцията е диференцируема в една точка, то тя е непрекъсната в тази точка.

Горното определение се пренася и за изображения:

Определение. *Вектор-функцията на n променливи $f(x_1, \ldots, x_n) = (f_k(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n))$ със стойности в $\mathbb{R}^k$ се нарича* *диференцируема в точката x, ако всичките и компоненти f_i, $i = 1, \ldots, k$, са диференцируеми в тази точка.*

С други думи, диференцируеми функции са тези, които добре се приближават с линейни функции около дадената точка.

Забележка. По-горе беше показано, че ако една функция на едно променливо притежава производна в дадена точка, то тя удовлетворява в тази точка формулата за нарастването. Следователно, за функции на едно променливо горната дефиниция съвпада с обичайната.

Сега теорема 1 може да се формулира така:

Теорема $1'$. *Ако функцията f притежава непрекъснати частни производни в околност на точката x, то тя е диференцируема в тази точка.*

Обратното твърдение не е вярно, но може да се докаже нещо по-слабо:

Теорема 2. *Ако функцията f е диференцируема в точката x, то тя притежава частни производни в тази точка, и са в сила равенствата*

$$A_i \;=\; \frac{\partial f}{\partial x_i}(x), \quad i = 1, \ldots, n.$$

Доказателство. Да вземем вектор h, който има само една ненулева координата с номер i: $h = t\vec{e}_i = (0, \ldots 0, t, 0, \ldots, 0)$ (тук и по-долу с $\vec{e}_i$ означаваме вектор с дължина единица, насочен по оста x_i). Тогава равенството от дефиницията добива вида:

$$f\left(x + t\vec{e}_i\right) \;=\; f(x) + t\;A_i + t\;\alpha_i\left(t\vec{e}_i\right),$$

и според определението на частна производна имаме

$$\frac{\partial f}{\partial x_i}(x) \;=\; \lim_{t\to 0}\frac{f\left(x + t\vec{e}_i\right) - f\left(x\right)}{t} \;=\; \lim_{t\to 0}\left(A_i + \alpha_i\left(t\vec{e}_i\right)\right) \;=\; A_i\;. \blacksquare$$

Оттук се вижда в частност, че константите A_i в дефиницията за диференцируема функция са определени еднозначно.

Разбира се, от диференцируемостта в дадена точка не следват нито съществуването на частни производни в околност на точката, нито –

още повече – тяхната непрекъснатост. Читателят сам може да намери съответните примери дори за функции на едно променливо.

Допирателна равнина към графиката на функция на няколко променливи. Нека $f(x_1, \ldots, x_n)$ е функция на n променливи, дефинирана и диференцируема в областта $D \in \mathbb{R}^n$. Както знаем, нейната графика G_f е подмножеството на $\mathbb{R}^{n+1}$, определено с условията

$$G_f = \Big\{ (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} : (x_1, \ldots, x_n) \in D, x_{n+1} = f(x_1, \ldots, x_n) \Big\} .$$

Определение. *Нека $x^0 = (x_1^0, \ldots, x_n^0)$ е точка от D, и $(x^0, f(x^0))$ е съответната точка от G_f. Под допирателна равнина в точката $(x^0, f(x^0))$ ще разбираме равнината в $\mathbb{R}^{n+1}$ с уравнение $y = l(x_1, \ldots, x_n)$, където*

$$l(x_1, \ldots, x_n) = f\left(x^0\right) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right)\left(x_i - x_i^0\right).$$

От формулата за нарастване на диференцируема функция следва следната характеризация на допирателната равнина:

Функцията $l(x)$ е единствената линейна функция, удовлетворяваща условието

$$f(x) - l(x) = o\left(\left\|x - x^0\right\|\right).$$

С други думи, измежду всички равнини в $\mathbb{R}^{n+1}$, минаващи през точката $(x^0, f(x^0))$, допирателната равнина е разположена най-близко до графиката на $f(x)$.

Определение. *Под допирателен вектор към G_f в точката $(x^0, f(x^0))$ разбираме всеки (свободен) вектор, колинеарен с допирателната равнина в тази точка.*

Очевидно допирателните вектори в точката $(x^0, f(x^0))$ образуват n-мерно линейно подпространство на $\mathbb{R}^{n+1}$, наричано допирателно подпространство към графиката в тази точка. Лесно можем да посочим един базис в това пространство: той се състои от допирателните към графиките на частичните функции в тази точка (ще напомним, че под i-та частична функция разбираме функцията на променливата

x_i, получена от $f(x_1, \ldots, x_n)$ чрез замразяване на всички останали променливи).

Така получаваме $n+1$-мерните вектори

$$l_i = \left(0, \ldots, 0, 1, 0, \ldots, 0, \frac{\partial f}{\partial x_i}\left(x^0\right)\right),$$

като единицата в дясната част е поставена на i-то място. Така написаните n на брой вектори са линейно независими и колинеарни с допирателната равнина (проверете!).

Допирателна равнина към графиката на изображение. Ще въведем аналогични на горните понятия и за векторно-значна функция, т.е. за изображение от $\mathbb{R}^n$ в $\mathbb{R}^k$. Нека $f(x_1, \ldots, x_n) = (f_k(x_1, \ldots, x_n), \ldots, f_k(x_1, \ldots, x_n))$ е диференцируема вектор-функция, дефинирана в подмножество D на $\mathbb{R}^n$ и вземаща стойности в $\mathbb{R}^k$. Ще означаваме точките на пространството $\mathbb{R}^{n+k}$ чрез (x, y), където $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_k)$ са точки в $\mathbb{R}^n$ и $\mathbb{R}^k$ съответно. Под графика G_f на изображението f ще разбираме подмножеството на $\mathbb{R}^{n+k}$, определено с условията

$$G_f = \left\{(x, y) \in \mathbb{R}^{n+k} : x \in D, y = f(x)\right\}.$$

Определение. *Нека $x^0 \in D$. Под допирателна равнина към G_f в точката $(x^0, f(x^0))$ ще разбираме n-мерната равнина в $\mathbb{R}^{n+k}$, определена с равенствата*

$$y_j = f_j\left(x^0\right) + \sum_{i=1}^{n} \frac{\partial f_j}{\partial x_i}\left(x^0\right)\left(x_i - x_i^0\right), \quad j = 1, \ldots, k.$$

Ако отново въведем понятията допирателен вектор и допирателно подпространство по същия начин, както по-горе, ще получим базис от n на брой допирателни вектори от вида

$$l_i = \left(0, \ldots, 0, 1, 0, \ldots, 0, \frac{\partial f_1}{\partial x_i}\left(x^0\right), \ldots, \frac{\partial f_k}{\partial x_i}\left(x^0\right)\right),$$

като отново единицата в дясната част е поставена на i-то място. Породеното от тези вектори n-мерно линейно подпространство на

$\mathbb{R}^{n+k}$ се нарича допирателно подпространство към графиката G_f на изображението f.

Аритметични операции с диференцируеми функции.

Както и в случая на функции на една променлива, при аритметични операции с диференцируеми функции резултатът отново е диференцируема функция:

Теорема 3. *Ако функциите на n променливи $f(x)$ и $g(x)$ са диференцируеми в точката $x^0 \in \mathbb{R}^n$, то тяхната сума $f(x)+g(x)$, произведение $f(x).g(x)$, и, ако $g(x^0) \neq 0$, тяхното частно $\frac{f(x)}{g(x)}$, са също диференцируеми в тази точка.*

Доказателство. Нека h е достатъчно малък по норма вектор от $\mathbb{R}^n$. Да напишем формулата за нарастването в точката x^0 за функциите $f(x)$ и $g(x)$:

$$\begin{aligned} f\left(x^0+h\right) &= f\left(x^0\right)+\sum_{i=1}^{n} A_i h_i + r_f(h), \\ g\left(x^0+h\right) &= g\left(x^0\right)+\sum_{i=1}^{n} B_i h_i + r_g(h), \end{aligned}$$

където $A_1,\dots,A_n,B_1,\dots,B_n$ са частните производни на $f(x)$ и $g(x)$ в x^0, и $r_f(h), r_g(h) = o\left(||h||\right)$. Като съберем двете равенства, получаваме

$$(f+g)\left(x^0+h\right) = (f+g)\left(x^0\right)+\sum_{i=1}^{n}\left(A_i+B_i\right)h_i + r_f(h)+r_g(h).$$

Тъй като за остатъка в горната формула имаме $r(h) = r_f(h) + r_g(h) = o\left(||h||\right)$, то полученото равенство изразява диференцируемостта на $(f+g)(x)$ в x^0.

За да получим формулата за нарастването за $f(x)g(x)$, да умножим двете равенства. Получаваме:

$$f\left(x^0+h\right)g\left(x^0+h\right) = f\left(x^0\right)g\left(x^0\right)+\sum_{i=1}^{n} C_i h_i + r(h),$$

където $C_i = g\left(x^0\right)A_i + f\left(x^0\right)B_i$, и

$$r(h) = \left(f\left(x^0\right)+\sum_{i=1}^{n} A_i h_i\right) r_g(h) + \left(g\left(x^0\right)+\sum_{i=1}^{n} B_i h_i\right) r_f(h) + r_f(h) r_g(h).$$

Имаме

$$\frac{r(h)}{||h||} = \left(f\left(x^0\right) + \sum_{i=1}^{n} A_i h_i\right) \frac{r_g(h)}{||h||} + \left(g\left(x^0\right) + \sum_{i=1}^{n} B_i h_i\right) \frac{r_f(h)}{||h||} + r_f(h) \frac{r_g(h)}{||h||},$$

и следователно $\frac{r(h)}{||h||} \to 0$ при $h \to 0$.

Накрая, нека $g\left(x^0\right) \neq 0$. За да докажем, че $\frac{f(x)}{g(x)}$ е диференцируема в x^0, е достатъчно да видим, че $\frac{1}{g(x)}$ е диференцируема в тази точка. Наистина, тъй като $\frac{f(x)}{g(x)} = f(x).\frac{1}{g(x)}$, диференцируемостта на частното ще следва от доказаната по-горе диференцируемост на произведение на диференцируеми функции.

За да докажем, че $\frac{1}{g(x)}$ е диференцируема в x^0, да означим

$$\Delta\left(\frac{1}{g}\right) = \frac{1}{g\left(x^0 + h\right)} - \frac{1}{g\left(x^0\right)}.$$

Тъй като $g(x)$ е непрекъсната в x^0, то същото е вярно и за $\frac{1}{g(x)}$, т.е. $\Delta\left(\frac{1}{g}\right) \to 0$ при $h \to 0$. Имаме

$$\Delta\left(\frac{1}{g}\right) = \frac{g\left(x^0\right) - g\left(x^0 + h\right)}{g\left(x^0\right) g\left(x^0 + h\right)} = -\frac{1}{g\left(x^0\right)}\ \Delta g \left(\frac{1}{g\left(x^0\right)} + \Delta\left(\frac{1}{g}\right)\right) =$$

$$= -\frac{1}{g^2\left(x^0\right)}\ \Delta g\ -\ \frac{1}{g\left(x^0\right)}\ \Delta g\ \Delta\left(\frac{1}{g}\right)\ =\ \sum_{i=1}^{n}\left(-\frac{B_i}{g^2\left(x^0\right)}\right) h_i + r(h),$$

където

$$r(h) = -\frac{1}{g^2\left(x^0\right)}\ r_g(h)\ -\ \frac{1}{g\left(x^0\right)}\ \Delta g\ \Delta\left(\frac{1}{g}\right).$$

Тогава

$$\frac{r(h)}{||h||}\ =\ -\frac{1}{g^2\left(x^0\right)}\ \frac{r_g(h)}{||h||}\ -\ \frac{1}{g\left(x^0\right)}\ \frac{\Delta g}{||h||}\ \Delta\left(\frac{1}{g}\right).$$

От диференцируемостта на $g(x)$ следва, че $\frac{\Delta g}{||h||}$ е ограничена при достатъчно малки h (докажете!), и следователно $\frac{r(h)}{||h||} \to 0$ при $h \to 0$. ■

Диференциране на съставни функции.

При намиране на частни производни на сума, произведение и т.н. на функции на няколко променливи не е необходимо да се знае нищо повече от известните формули за производна на аритметични операции с функции на една променлива [*]. Един въпрос, при който положението е различно, е диференцирането на сложни, или съставни, функции. Ще напомним формулата за случая на едно променливо: при дадени функции $f(x)$ и $\varphi(t)$ (външна и вътрешна функция), производната на съставната функция $F(t) = f(\varphi(t))$ се дава с формулата

$$F'(t) \ = \ f'(\varphi(t)).\varphi'(t),$$

т.е. производната на съставната функция е равна на произведението на производните на външната и вътрешната функция. Ще докажем, че подобна формула съществува и за съставна функция на няколко променливи. Ще започнем с по-простия вариант на теоремата:

Теорема 4.*Нека $f(x) = f(x_1, \ldots, x_n)$ е функция, дефинирана в околност на точката $x^0 = (x_1^0, \ldots, x_n^0)$, и нека функциите на едно променливо $(\varphi_1(t), \ldots, \varphi_n(t))$ са дефинирани и диференцируеми в околност на точката $t^0 \in \mathbb{R}$, при което $x_i^0 = \varphi_i(t^0)$, $i = 1, \ldots, n$. Да предположим, че $f(x)$ е диференцируема (в смисъл на горното определение) в x^0, а $\varphi_1(t), \ldots, \varphi_n(t)$ - в t^0. Тогава съставната функция*

$$F(t) \ = \ f(\varphi_1(t), \ldots, \varphi_n(t))$$

е диференцируема в точката t^0, като нейната производна се дава с формулата

$$F'\left(t^0\right) \ = \ \frac{\partial f}{\partial x_1}\left(x^0\right)\varphi_1'\left(t^0\right) + \ldots + \frac{\partial f}{\partial x_n}\left(x^0\right)\varphi_n'\left(t^0\right) \ = \ \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\left(x^0\right)\varphi_i'\left(t^0\right).$$

С други думи, формулата за производната на съставната функция се състои от толкова събираеми, колкото са променливите, като всяко от тях наподобява по форма израза за производна на съставна функция на едно променливо.

Доказателство на теоремата. Грубо казано, доказателството се състои в заменяне на функцията с нейната линейна част. Нека t е точка

[*] Всъщност горните изчисления доказват още веднаж тези формули.

достатъчно близка до t^0. Да означим $\Delta t = t - t^0$, $\Delta x_i = \varphi_i(t) - \varphi_i\left(t^0\right)$, $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$. Имаме

$$\lim_{\Delta t \to 0} \frac{\Delta x_i}{\Delta t} = \varphi_i'\left(t^0\right).$$

Нека $\Delta F = F(t) - F\left(t^0\right) = f\left(x^0 + \Delta x\right) - f\left(x^0\right)$. Тогава от формулата за нарастването имаме

$$\Delta F = \sum_{i=1}^{n} f'_{x_i}\left(x^0\right) \Delta x_i + \sum_{i=1}^{n} \alpha_i\left(\Delta x\right) \Delta x_i,$$

като $\alpha_i\left(\Delta x\right) \to 0$ при $\Delta x \to 0$. Следователно,

$$\frac{\Delta F}{\Delta t} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right) \frac{\Delta x_i}{\Delta t} + \sum_{i=1}^{n} \alpha_i\left(\Delta x\right) \frac{\Delta x_i}{\Delta t} \to \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right) \varphi_i'\left(t^0\right). \blacksquare$$

Пример. Нека $f(u,v) = u^v$, $u > 0, v > 0$, и нека $\varphi(t), \psi(t)$ са диференцируеми функции с положителни стойности. Ако положим $F(t) = \varphi(t)^{\psi(t)} = f(\varphi(t), \psi(t))$, то от горната теорема получаваме

$$F'(t) = \psi(t).\varphi(t)^{\psi(t)-1} + \ln\varphi(t).\varphi(t)^{\psi(t)}.$$

Тази формула е известна от първата част на курса, където тя се извежда чрез предварително логаритмуване.

Ще докажем аналогичната теорема за векторни функции (изображения):

Теорема 4′. *Нека $f(x) = f\left(x_1, \ldots, x_n\right)$ е функция, дефинирана и диференцируема в околност на точката $x^0 = \left(x_1^0, \ldots, x_n^0\right)$, а векторната функция $\varphi(t) = (\varphi_1(t), \ldots, \varphi_n(t))$, зависеща от аргумента $t = (t_1, \ldots, t_k) \in \mathbb{R}^k$, е дефинирана в околност на точката $t^0 = \left(t_1^0, \ldots, t_k^0\right)$, при което $\varphi\left(t^0\right) = x^0$. Да предположим, че функцията на n променливи $f(x)$ е диференцируема в x^0, а изображението $\varphi(t)$ (т.е. неговите компоненти $\varphi_1(t), \ldots, \varphi_n(t)$) - в точката t^0. Тогава съставната функция $F(t) = f(\varphi(t))$ е диференцируема в t^0, и са в сила равенствата:*

$$\frac{\partial F}{\partial t_j}\left(t^0\right) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right) \frac{\partial \varphi_i}{\partial t_j}\left(t^0\right), \quad j = 1, \ldots, k.$$

Забележка. Тук новото в сравнение с теорема 4 се състои единствено във факта, че функцията $F(t)$ е диференцируема като функция на k променливи според определението на настоящия параграф, а не само по всяка променлива поотделно.

Доказателство. Нека, както по-горе, $\Delta t = t - t^0$, $\Delta t_j = t_j - t_j^0$, $\Delta x_i = \varphi_i(t) - \varphi_i(t^0)$, $\Delta x = (\Delta x_1, \ldots, \Delta x_n)$. Формулата за нарастванията, приложена към функциите $f, \varphi_1(t), \ldots, \varphi_n(t)$, ни дава равенствата

$$\Delta f \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right) \Delta x_i \;+\; \sum_{i=1}^{n} \alpha_i\left(\Delta x\right) \Delta x_i,$$

$$\Delta \varphi_i \;=\; \sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial t_j}\left(t^0\right) \Delta t_j \;+\; \sum_{i=1}^{n} \beta_{i,j}\left(\Delta t\right) \Delta t_j,$$

където функциите $\alpha_i(\Delta x)$, $\beta_{i,j}(\Delta t)$ клонят към нула при Δx, респ. Δt клонящи към нула. Преминавайки в първото от равенствата към променливите t, заместваме Δx_i с $\Delta \varphi_i$, и вместо Δf пишем ΔF. Ще отделим в отделно събираемо членовете, съдържащи някоя от безкрайно малките функции $\alpha_i(\Delta x)$, $\beta_{i,j}(\Delta t)$. Размествайки реда на сумиране, получаваме:

$$\Delta F = \sum_{j=1}^{k} \left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(x^0\right) \frac{\partial \varphi_i}{\partial t_j}\left(t^0\right) \right) \Delta t_j + r\left(\Delta t\right),$$

където остатъкът $r(\Delta t)$ се дава с формулите

$$r\left(\Delta t\right) = \sum_{j=1}^{k} \gamma_j\left(\Delta t\right) \Delta t_j,$$

като

$$\gamma_j\left(\Delta t\right) = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\left(x^0\right) \beta_{i,j}\left(\Delta t\right) + \alpha_i\left(\Delta \varphi\right) \frac{\partial \varphi_i}{\partial t_j}\left(t^0\right) + \alpha_i\left(\Delta \varphi\right) \beta_{i,j}\left(\Delta t\right) \right).$$

Очевидно имаме $\gamma_j(\Delta t) \to 0$ при $\Delta t \to 0$, и следователно за функцията $F(t)$ в точката t^0 е изпълнена формулата за нарастването, т.е. тя е диференцируема в тази точка. Попътно равенството за ΔF ни дава още веднаж формули за частните производни на $F(t)$ в t^0. ■

Матрична производна на изображение.

Нека $D \subset \mathbb{R}^n$ и изображението $f(x) : D \to \mathbb{R}^m$ $x = (x_1, \ldots, x_n) \in D$ има компоненти $f(x) = (f_1(x), \ldots, f_m(x))$, където $f_1(x), \ldots, f_m(x)$ са диференцируеми функции в D. Ако разгледаме всички първи частни производни $\frac{\partial f_i}{\partial x_j}(x)$ на координатните функции, получаваме $n.m$ на брой различни производни, и картината изглежда доста сложна. Ситуацията става значително по-обозрима, ако подредим тези производни в една $n \times m$ - матрица. Така стигаме до следната дефиниция:

Определение. *Под* <u>*матрична производна*</u> *на изображението* $f(x)$ *в точката* x^0 *ще разбираме матрицата*

$$Df\left(x^0\right) = \left\{\frac{\partial f_i}{\partial x_j}\left(x^0\right)\right\}_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}\left(x^0\right) & \frac{\partial f_1}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial f_1}{\partial x_n}\left(x^0\right) \\ \frac{\partial f_2}{\partial x_1}\left(x^0\right) & \frac{\partial f_2}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial f_2}{\partial x_n}\left(x^0\right) \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_m}{\partial x_1}\left(x^0\right) & \frac{\partial f_m}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial f_m}{\partial x_n}\left(x^0\right) \end{pmatrix}.$$

В частност, матричната производна на една скаларна функция е вектора, съставен от нейните производни.

Използването на операциите с матрици позволява да се даде прост вид на формулата за нарастването и формулата за производна на съставна функция за изображения. Ще напомним, че всяка $n \times m$ – матрица може да се разглежда като линеен оператор от $\mathbb{R}^n$ в $\mathbb{R}^m$. Нека, както по-горе, $f(x) : D \to \mathbb{R}^m$ $x = (x_1, \ldots, x_n) \in D$ е m – мерна вектор-функция, и $h = (h_1, \ldots, h_n)$ е вектора на нарастването на аргумента. Да напишем за всяка от компонентите $f_1(x), \ldots, f_m(x)$ на изображението $f(x)$ формулата за нарастването във формата $(**)$ за точката x^0 и нарастването h. Написвайки получените m на брой скаларни равенства като равенство между m-мерни вектори, получаваме:

Формула за нарастването за вектор-функции. При горните условия имаме

$$f\left(x^0 + h\right) \; = \; f\left(x^0\right) \; + \; Df\left(x^0\right) \circ h \; + \; r(h),$$

където $r(h)$ е m – мерна вектор-функция, удовлетворяваща условието:

$$||r(h)|| \; = \; o\left(||h||\right).$$

Матрицата $Df\left(x^0\right)$, разглеждана като оператор от $\mathbb{R}^n$ в $\mathbb{R}^n$, се нарича диференциал на изображението $f(x)$ в точката x^0.

Ще покажем, че при горния подход и теоремата за диференциране на съставно изображение добива съвсем прост вид, аналогичен на този при функции на едно променливо. Ще напомним, че на всяка двойка матрици A, B с размерност съотв. $n \times m$ и $k \times n$ се съпоставя матрицата $A \circ B$ с размерност $k \times m$: ако $A = \{a_{ij}\}$, $B = \{b_{pq}\}$, то $A \circ B = \{c_{ls}\}$ с $c_{ls} = \sum_{i=1}^{n} b_{li}.a_{is}$.

Нека сега вектор-функцията $\varphi(t) = (\varphi_1(t), \ldots, \varphi_n(t))$ е определена в околност на точката $t^0 = (t_1^0, \ldots, t_k^0) \in \mathbb{R}^k$, а вектор-функцията $f(x) = (f_1(x), \ldots, f_m(x))$ - в околност на точката $x^0 = \varphi\left(t^0\right) \in \mathbb{R}^n$. Тогава съставната функция $F(t) = f(\varphi(t))$ е определена в околност на точката t^0 и взема стойности в пространството $\mathbb{R}^k$.

Теорема 4. *Да предположим, че функциите $f(x)$ и $\varphi(t)$ са диференцируеми съответно в точките x^0 и t^0. Тогава съставната функция $F(t)$ е диференцируема в точката t^0, и е в сила равенството*

$$DF\left(t^0\right) = Df\left(x^0\right) \circ D\varphi\left(t^0\right).$$

Доказателство. В същност твърдението представлява само друга формулировка на резултата на теорема $4'$. Наистина, то се получава непосредствено от даденото по-горе определение на произведение на матрици и от доказаните по-горе равенства

$$\frac{\partial F_l}{\partial t_j}\left(t^0\right) = \sum_{i=1}^{n} \frac{\partial f_l}{\partial x_i}\left(x^0\right) \frac{\partial \varphi_i}{\partial t_j}\left(t^0\right).$$

■

Функционална детерминанта на изображение.

По-нататък често ще се налага да разглеждаме изображения, които действат от $\mathbb{R}^n$ в $\mathbb{R}^n$, или, по-точно, са дефинирани в подмножество на $\mathbb{R}^n$ и вземат стойности в $\mathbb{R}^n$. Тогава тяхната матрична производна представлява квадратна матрица, и ние може да разгледаме нейната детерминанта, която играе много важна роля в по-нататъшните разглеждания.

Определение. *Нека вектор-функцията $f(x) = (f_1(x), \ldots, f_n(x))$, вземаща стойности в $\mathbb{R}^n$, е дефинирана в околност на точката $x^0 = \varphi(t^0) \in \mathbb{R}^n$. Под функционална детерминанта, или якобиан, на $f(x)$ в точката x^0, ще разбираме числото*

$$\frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}\left(x^0\right) \;=\; \det\left(Df\left(x^0\right)\right).$$

Теорема 5.(основно свойство на функционалните детерминанти). *Функционалната детерминанта на произведението на две изображения е равна на произведението на техните функционални детерминанти.*

Доказателство. Нека $F(t) = f(\varphi(t))$. Тъй като детерминанта на произведение на две квадратни матрици е равна на произведението на техните детерминанти, то твърдението непосредствено следва от теорема 4. ■

Доказаното равенство се записва във вида

$$\frac{D(F_1, \ldots, F_n)}{D(t_1, \ldots, t_n)}\left(t^0\right) \;=\; \frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}\left(x^0\right) \,\cdot\, \frac{D(\varphi_1, \ldots, \varphi_n)}{D(t_1, \ldots, t_n)}\left(t^0\right).$$

Вижда се, че начинът на записване е така съставен, че да подсеща читателя за правилната формула: горното равенство може да се тълкува като "съкращаване"на $D(x_1, \ldots, x_n)$ с "равното му" $D(\varphi_1, \ldots, \varphi_n)$.

Производна на обратното изображение. Както знаем, изображението $f : X \to Y$ от множеството X в множеството Y е биективно (взаимно еднозначно), ако за всяка точка $y \in Y$ съществува точно един праобраз в X, т.е. такова $x \in X$, че $f(x) = y$. Полагайки $x = g(y)$, получаваме обратното изображение $g = f^{-1} : Y \to X$. Лесно се вижда, че са изпълнени равенствата $g \circ f = I_X$, $f \circ g = I_Y$, където I_X и I_Y са идентичните изображения на пространствата X и Y в себе си.

Теорема 6.*Нека $f(x)$ е биективно изображение на множеството $D \subset \mathbb{R}^n$ в множеството $E \subset \mathbb{R}^n$ и $g(y)$ е неговото обратно изображение. Нека изображението $f(x)$ е диференцируемо в точката*

$x^0 \in D$, а изображението $g(y)$ - в точката $y^0 = f\left(x^0\right)$. Тогава са в сила равенствата:

$$Dg\left(y^0\right) = Df\left(x^0\right)^{-1}\ ,\ \frac{D\left(g_1,\ldots,g_n\right)}{D\left(y_1,\ldots,y_n\right)}\left(y^0\right) = \frac{1}{\frac{D(f_1,\ldots,f_n)}{D(x_1,\ldots,x_n)}\left(x^0\right)}.$$

Доказателство. От равенството $g(f(x)) = x$ и теоремата за производна на съставно изображение следва матричното равенство $Dg\left(y^0\right) \circ Df\left(x^0\right) = I$, където I е единичната $n \times n$ матрица, което е първото от горните равенства. Второто равенство следва от цитираното по-горе свойство на детерминантата. ■

В частност, от тук се вижда, че при условията на теоремата имаме $\frac{D(f_1,\ldots,f_n)}{D(x_1,\ldots,x_n)}\left(x^0\right) \neq 0$.

Забележка. В параграф ??? ние ще видим, че е вярно и обратното твърдение: ако функционалната детерминанта на едно диференцируемо биективно изображение е различна от нула, то обратното му изображение също е диференцируемо.

Упражнения.

1. Функцията $f\left(\vec{x}\right) = f\left(x_1,\ldots,x_n\right)$ се нарича *хомогенна от степен* k, ако за всяко число t е изпълнено равенството

$$f\left(t\vec{x}\right)\ =\ t^k f\left(\vec{x}\right).$$

Докажете, че за функциите, хомогенни от степен k, е в сила равенството на Ойлер:

$$x_1\frac{\partial f}{\partial x_1}\left(\vec{x}\right) + \ldots + x_n\frac{\partial f}{\partial x_n}\left(\vec{x}\right) = k\ f\left(t\vec{x}\right).$$

1.5 Производна по направление и градиент

За функции на едно променливо първата производна в дадена точка има прост геометричен смисъл - тя е равна на ъгловия коефициент на допирателната към графиката на функцията в тази точка. С други думи, тя изразява "стръмността"на графиката в тази точка (ако функцията нараства, производната е положителна, и т.н.). За функции на n променливи първите производни са също n на брой, и тяхната интерпретация не е толкова очевидна.

За си изясним геометричният смисъл на първите производни, ще си послужим с аналогията, която споменахме в началото на предния параграф: ще разгледаме графиката на дадена функция на две променливи като релеф на част от земната повърхност. Така, да си представим, че се намираме на склона на някакъв хълм (виж чертежа). Тогава стръмността на пътеката, по която вървим, зависи от нейната посока: ние може да тръгнем право нагоре (голям положителен наклон), право надолу (голям отрицателен), или да поемем по хоризонталата - тогава пътеката е хоризонтална и наклонът е нулев. Тези съображения водят до понятието производна по направление, което ще изложим по-долу.

Производна по направление. Нека $f(x) = f(x_1, \ldots, x_n)$ е функция на n променливи, диференцируема в точката x^0, и нека $\vec{e} = (e_1, \ldots, e_n)$ е n-мерен вектор с норма единица (т.е. $||e|| = \sqrt{\sum_{i=1}^{n} e_i^2} = 1$. [*])

Определение. *Под производна на функцията $f(x)$ по направление $\vec{e}$ в точката x^0 ще разбираме границата*

$$\frac{df}{d\vec{e}}\left(x^0\right) = \lim_{t \to 0} \frac{f\left(x^0 + t\vec{e}\right) - f\left(x^0\right)}{t}.$$

[*] Компонентите $e_1, \ldots, e_n$ на единичния вектор $\vec{e}$ се наричат негови директорни косинуси. Ако означим с $\vec{e^i}$ единичния координатен вектор по оста x_i, то компонентата e_i е равна на скаларното произведение $\left\langle \vec{e}, \vec{e^i} \right\rangle$ и следователно на косинуса на ъгъла между $\vec{e}$ и оста x_i.

Ще изразим $\frac{df}{d\vec{e}}(x^0)$ чрез частните производни на $f(x)$. Правата, минаваща през точката x^0 и колинеарна на вектора $\vec{e}$, може да се параметризира с уравненията

$$x(t) = \left(x_1^0 \ + \ te_1, \ldots, x_n^0 \ + \ te_n\right).$$

(Тъй като $||e|| = 1$, параметърът t съвпада с ориентираното разстояние от текущата точка $x(t)$ до началната точка x^0.)

Тогава ограничението на $f\left(x\right)$ върху тази права се представя с функцията на едно променливо

$$\varphi(t) \ = \ f(x(t)) \ = \ f\left(x_1^0 \ + \ te_1, \ldots, x_n^0 \ + \ te_n\right).$$

Прилагайки за функцията $\varphi(t) = f\left(x_1^0 + te_1, \ldots, x_n^0 + te_n\right)$ правилото за диференциране на съставна функция, получаваме равенството

$$(*) \qquad \frac{df}{d\vec{e}}\left(x^0\right) \ = \ \sum_{i=1}^{n} e_i \frac{df}{dx_i}\left(x^0\right).$$

Дефиницията на производна по направление обобщава понятието частна производна, дефинирано в предния параграф. Наистина, ако в ролята на $\vec{e}$ вземем i-тият координатен вектор $\vec{e^i}$ (това е вектор с координата единица на i-тото място и нули на останалите места), то от формулата $(*)$ получаваме:

$$\frac{df}{d\vec{e^i}}\left(x^0\right) \ = \ \frac{df}{dx_i}\left(x^0\right).$$

Забележка. По-нататък ние ще използваме дефиницията за производна по направление и формулата $(*)$ за произволен n-мерен вектор $\vec{h}$, без да изискваме условието $\left|\left|\vec{h}\right|\right| = 1$. Тогава определението губи директен геометричен смисъл, но основните му свойства остават в сила. За по-нататъшни цели ще въведем означението:

$$D_{\vec{h}} f(x) = \frac{df}{d\vec{h}}\left(x\right) = \sum_{i=1}^{n} h_i \frac{df}{dx_i}\left(x\right).$$

Ако означим с D_i операцията на диференциране по променливата x_i, е в сила равенството между оператори

$$D_{\vec{h}} = \sum_{i=1}^{n} h_i D_i.$$

Градиент на функция. Да разгледаме, при фиксирана точка x^0, зависимостта на производната по направление $f'_{\vec{e}}(x^0)$ от единичния вектор $\vec{e}$. С други думи, пита се по кое направление производната на функцията е най-голяма, респ. най-малка. За целта ще използваме понятието *скаларно произведение* на вектори от $\mathbb{R}^n$, което може да бъде дефинирано по аналитичен и геометричен път. Ще напомним, че за два вектора $\vec{a} = (a_1, \ldots, a_n)$, $\vec{b} = (b_1, \ldots, b_n)$ от $\mathbb{R}^n$ скаларното произведение се дефинира с формулата

$$\left\langle \vec{a}, \vec{b} \right\rangle \; = \; \sum_{i=1}^{n} a_i b_i.$$

В сила е равенството

$$\left\langle \vec{a}, \vec{b} \right\rangle \; = \; ||\vec{a}|| \left|\left|\vec{b}\right|\right| \cos \angle \left(\vec{a}, \vec{b}\right),$$

където $\angle\left(\vec{a}, \vec{b}\right)$ означава ъгъла между векторите $\vec{a}$ и $\vec{b}$.

Определение. *Под градиент на функцията на n променливи $f\left(x_1, \ldots, x_n\right)$ в точката x^0 ще разбираме n-мерен вектор с координати, равни на частните производни в тази точка:*

$$\overrightarrow{\mathrm{grad}}\, f\left(x^0\right) \; = \; \left(\frac{df}{dx_1}\left(x^0\right), \frac{df}{dx_2}\left(x^0\right), \ldots, \frac{df}{dx_n}\left(x^0\right)\right).$$

Тогава формулата $(*)$ може да се напише и така:

$$(**) \quad \frac{df}{d\vec{e}}\left(x^0\right) = \left\langle \overrightarrow{\mathrm{grad}}\, f\left(x^0\right), \vec{e} \right\rangle = \left|\left|\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)\right|\right| \cos \angle \left(\overrightarrow{\mathrm{grad}}\, f\left(x^0\right), \vec{e}\right),$$

тъй като $||\vec{e}|| = 1$.

Знаем, че най-голямата стойност на косинуса на един ъгъл е равна на единица и се достига за ъгъл, равен на нула. Оттук получаваме:

Теорема 1. *Производната по направление* $\frac{df}{d\vec{e}}\left(x^0\right)$ *в точката* $\left(x^0\right)$ *взема най-голямата си стойност, когато вектора* $\vec{e}$ *е еднопосочен с вектора* $\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)$, *т.е.*

$$\vec{e} \;=\; \frac{\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)}{\left\|\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)\right\|}.$$

В този случай имаме

$$\frac{df}{d\vec{e}}\left(x^0\right) \;=\; \left\|\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)\right\|.$$

С други думи, имаме следната геометрична интерпретация на вектора $\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)$: посоката му съвпада с посоката, в която $f(x)$ расте най-бързо, а големината му е равна на наклона на графиката на функцията в тази посока.

Ще посочим още една геометрична характеризация на вектора $\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)$. От формулата $(**)$ се вижда, че $\frac{df}{d\vec{e}}\left(x^0\right) = 0$ точно тогава, когато $\vec{e}$ е перпендикулярен на $\overrightarrow{\mathrm{grad}}\, f\left(x^0\right)$. Иначе казано, ако се движим по "хоризонталата", направлението на движението във всеки момент ще бъде перпендикулярно на градиента в съответната точка. Така, засега на интуитивно ниво, стигаме до твърдението (виж чертежа).

Теорема 2. *Градиентът на една функция в дадена точка е перпендикулярен на линията на ниво на функцията, минаваща през тази точка.*

Точната формулировка на това твърдение е следната: градиентът е перпендикулярен на допирателната към линията на ниво в точката. (За допирателна към параметрично зададена крива виж I.2.12.) Параметризирането на линията на ниво, както и доказателството на горната теорема, ще бъде направено в следващите параграфи ???????.

Забележка. Градиентът на дадена функция беше дефиниран при наличие на декартова координатна система в пространството $\mathbb{R}^n$. Теореми 1 и 2 обаче показват, че вектора $\overrightarrow{\mathrm{grad}}\, f(x)$ има инвариантен геометричен смисъл, т.е не зависи от избора на координатната система, а само от скаларното произведение в $\mathbb{R}^n$. За точната формулировка на това наблюдение виж зад. ???.

Формула за крайните нараствания за функции на много променливи. Ще обобщим една от най-важните формули на диференциалното смятане за една променлива – формулата за крайните нараствания, доказана в I.§2.3 (и използвана в предния параграф). Нека $x^0 \in \mathbb{R}^n$, $\vec{h} = (h_1, \ldots, h_n)$ (вектор на нарастването), и нека функцията $f(x)$ е дефинирана и диференцируема във всички точки от отсечката, свързваща точките x^0 и $x^0 + \vec{h}$ в $\mathbb{R}^n$.

Теорема 3. *При горните условия съществува число $\theta \in (0, 1)$ такова, че*

$$f\left(x^0 + \vec{h}\right) - f\left(x^0\right) \;=\; \sum_{i=1}^{n} h_i \frac{df}{dx_i}\left(x^0 + \theta\vec{h}\right)$$

Доказателство. Да въведем, както по-горе, помощната функция на една променлива $\varphi(t) = f\left(x_1^0 + th_1, \ldots, x_n^0 + th_n\right)$. Тогава от условието на теоремата следва, че тя е диференцируема в интервала $[0, 1]$. Прилагайки едномерната теорема за крайните нараствания, получаваме съществуването на $\theta \in (0, 1)$ такова, че $\varphi(1) - \varphi(0) = \varphi'(\theta)$. Тогава теоремата следва от очевидните равенства

$$\varphi(0) = f\left(x^0\right),\ \varphi(1) = f\left(x^0 + \vec{h}\right),\ \varphi'(t) = \sum_{i=1}^{n} h_i \frac{df}{dx_i}\left(x^0 + t\vec{h}\right). \blacksquare$$

Следствие 4. Да предположим, че производните на $f(x)$ удовлетворяват неравенството

$$\sqrt{\sum_{i=1}^{n} \left(\frac{df}{dx_i}(x)\right)^2} \leq C.$$

Тогава е в сила неравенството

$$\left|f\left(x^0 + \vec{h}\right) - f\left(x^0\right)\right| \;\leq\; C\,||h||\,.$$

Доказателство. Горното условие всъщност означава, че $\left\|\overrightarrow{\text{grad}}\, f(x)\right\| \leq C$. Твърдението на теоремата може да се напише във вида

$$f\left(x^0 + \vec{h}\right) - f\left(x^0\right) \;=\; \left\langle \vec{h}, \overrightarrow{\text{grad}}\, f\left(x^0 + \theta\vec{h}\right)\right\rangle$$

и следователно

$$\left|f\left(x^0 + \vec{h}\right) - f\left(x^0\right)\right| \;\leq\; \|h\| \left\|\overrightarrow{\text{grad}}\, f\left(x^0 + \theta\vec{h}\right)\right\| \;\leq\; C \|h\| . \blacksquare$$

1.6 Производни от по-висок ред

Досега ние разглеждахме само производни от първи ред. Както за функции на едно променливо, ние можем да продължим да диференцираме и да получим производни от по-висок ред. Така, с $\dfrac{\partial^2 f}{\partial x_i\, \partial x_j}(x)$ или с $f_{x_i x_j}(x)$ означаваме производната на функцията $\dfrac{\partial f}{\partial x_i}$ по променливата x_j в точката x.

На пръв поглед картината изглежда доста сложна: ако $f(x_1, \ldots, x_n)$ е функция на n променливи, то първите и производни са n на брой. Всяка от тях може да бъде диференцирана по n променливи, и така получаваме общо n^2 различни втори производни, n^3 - трети, и n^k производни от k-ти ред. Всъщност картината е малко по-оптимистична: в този параграф ние ще покажем, че при някои леки ограничения стойността на производните не зависи от реда, в който е извършено диференцирането. Така ситуацията се опростява – например за функция на две променливи различните производни от ред k не са 2^k, а само $k+1$ на брой.

Ще докажем теоремата за независимостта на висшите производни от реда на диференцирането най-напред за функция на две променливи.

Теорема 1. *Нека $f(x, y)$ е функция на две променливи, дефинирана в околност на точката (x_0, y_0). Да предположим, че първите производни f'_x, f'_y, и смесените втори производни f''_{xy} и f''_{yx} съществуват в околност на (x_0, y_0). Ако f''_{xy} и f''_{yx} са непрекъснати в тази точка, то стойностите им в нея са равни.*

Доказателство. Ще апроксимираме смесените втори производни с "диференчни частни от втори ред". Нека h и k са достатъчно малки числа. Да означим

$$W(h, k) = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0).$$

Да разгледаме помощната функция

$$\varphi(x) = f(x, y_0 + k) - f(x, y_0),$$

дефинирана за стойности на x, достатъчно близки до x_0. Групирайки в израза за $W(h, k)$ първото с третото, и второто с четвъртото събираемо, получаваме равенството

$$W(h,k) = \varphi(x_0+h) - \varphi(x_0) = h\varphi'(x_0+\theta_1 h),$$

за подходящо $\theta_1 \in (0,1)$ (по теоремата за крайните нараствания за функцията φ). Тъй като $\varphi'(x) = f'_x(x, y_0+k) - f'_x(x, y_0)$, то

$$W(h,k) = hf'_x(x_0+\theta_1 h, y_0+k) - f'_x(x_0+\theta_1 h, y_0) = hkf''_{xy}(x_0+\theta_1 h, y_0+\theta_2 k),$$

като последното равенство се получава отново чрез прилагане на теоремата за крайните нараствания, този път за функцията $f'_x(x_0+\theta_1 h, y)$ на променливата y. Тук отново $\theta_2 \in (0,1)$ (разбира се, числата θ_1 и θ_2 зависят от h и k). Оттук

$$\frac{W(h,k)}{hk} = f''_{xy}(x_0+\theta_1 h, y_0+\theta_2 k),$$

и следователно при $h \to 0$, $k \to 0$ имаме

$$\frac{W(h,k)}{hk} \to f''_{xy}(x_0, y_0).$$

Ясно е, че в горното разсъждение местата на x и y могат да бъдат разменени. Да въведем друга помощна функция $\psi(y) = f(x_0+h, y) - f(x_0, y)$. Тогава $W(h,k) = \psi(y_0+k) - \psi(y_0) = k\psi'(y_0+\theta_3 k)$, откъдето както по-горе получаваме

$$\frac{W(h,k)}{hk} = f''_{yx}(x_0+\theta_4 h, y_0+\theta_3 k) \text{ и } \lim_{h\to 0,\, k\to 0} \frac{W(h,k)}{hk} = f''_{yx}(x_0, y_0),$$

откъдето $f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$. ■

Следствие 2. *Ако функцията $f(x_1, \dots, x_n)$ притежава непрекъснати частни производни до ред k в дадена околност, то техните стойности не зависят от реда на диференциране.*

Доказателство. За удобство ние ще въведем нови означения. Да означим с D_i операцията на диференциране по променливата x_i:

$$D_i f(x) \;=\; \frac{\partial f}{\partial x_i}(x).$$

Тогава производните от по-висок ред се получават чрез последователно прилагане на съответните операции:

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x) \;=\; D_{i_1} D_{i_2} \dots D_{i_k} f(x).$$

При тези означения твърдението на теорема 1 означава, че при направените в нея предположения тези операции комутират, т.е.

$$D_i D_j f(x) = D_j D_i f(x)$$

за всеки два индекса i, j, намиращи се между 1 и n.

Нека сега изберем система $i_1, \dots, i_k$ от индекси между 1 и n, и разгледаме производната $\dfrac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}}(x)$. Трябва да докажем, че стойността на тази производна не се променя при произволно разместване на индексите $i_1, \dots, i_k$. От казаното дотук се вижда, че тази стойност няма да се промени, ако разменим местата на две съседни диференцирания, например на тези, отговарящи на индексите i_p и i_{p+1}. Наистина,

$$D_{i_1} \dots D_{i_{p+1}} D_{i_p} \dots D_{i_k} f(x) = D_{i_1} \dots D_{i_{p-1}} \Big(D_{i_{p+1}} D_{i_p} \Big(D_{i_{p+2}} \dots D_{i_k} f(x) \Big) \Big) =$$

$$= D_{i_1} \dots D_{i_{p-1}} \Big(D_{i_p} D_{i_{p+1}} \Big(D_{i_{p+2}} \dots D_{i_k} f(x) \Big) \Big) = D_{i_1} \dots D_{i_p} D_{i_{p+1}} \dots D_{i_k} f(x).$$

Както е известно от алгебрата, всяко разместване на индексите $i_1, \dots, i_k$ може да се получи като произведение на размествания, разменящи местата на два съседни елемента и не засягащи останалите. Тъй като при всяко разместване на два съседни индекса резултата не се променя, той няма да се промени и при произволно разместване на индексите. ■

Определение. *Ако една функция притежава непрекъснати частни производни до ред k включително във вътрешността на дефиниционната си област D, че казваме, че тя е <u>k-кратно гладка</u> в D. Пространството от всички такива функции ще бележим с $\mathcal{C}^k(D)$.*

Важен пример. В този раздел ще намерим висшите производни на помощната функция $\varphi(t)$, използвана в предния параграф. Ще припомним нейната дефиниция: нека функцията $f(x_1, \ldots, x_n)$ е дефинирана в околност на точката $x^0 = (x_1^0, \ldots, x_n^0)$ и притежава непрекъснати частни производни до ред m включително (т.е. $f \in \mathcal{C}^m(D)$), и нека $\vec{h} = (h_1, \ldots, h_n) \in \mathbb{R}^n$. Функцията $\varphi(t)$ е определена в околност на точката $t = 0$ с формулата

$$\varphi(t) = f\left(x_1^0 + th_1, \ldots, x_n^0 + th_n\right).$$

Както в §5, формулата за диференциране на съставна функция ни дава

$$\varphi'(t) = D_{\vec{h}} f\left(x^0 + t\vec{h}\right) = \sum_{i=1}^{n} h_i D_i f\left(x^0 + t\vec{h}\right).$$

Прилагайки последователно към горното равенство формулата за диференциране на съставна функция, получаваме

$$\varphi''(t) = D_{\vec{h}}\left(D_{\vec{h}} f\left(x^0 + t\vec{h}\right)\right) = \left(D_{\vec{h}}\right)^2 f\left(x^0 + t\vec{h}\right)$$

и по-общо

$$\varphi^{(k)}(t) = \left(D_{\vec{h}}\right)^k f\left(x^0 + t\vec{h}\right) \text{ за } k = 1, \ldots, m.$$

(Ще поясним, че с $\left(D_{\vec{h}}\right)^k f$ се означава резултата от k последователни прилагания на диференциалния оператор $D_{\vec{h}}$ към функцията f.)

Нашата задача се сведе към следната: да се изразят степените $\left(D_{\vec{h}}\right)^k$ на оператора $D_{\vec{h}}$ чрез операторите на частно диференциране $D_1, \ldots, D_n$. Ще припомним, че $D_{\vec{h}} = \sum_{i=1}^{n} h_i D_i$. Важно е да се отбележи, че, както беше показано по-горе, в това представяне отделните събираеми $h_i D_i$, $h_j D_j$ комутират помежду си.

От алгебрата знаем, че за всеки n елементи $a_1, \ldots, a_n$ на даден комутативен пръстен, и за всяко естествено k, е в сила формулата

$$(a_1 + \ldots + a_n)^k = \sum_{\substack{1 \le k_1, \ldots, k_n \le n \\ k_1 + \ldots + k_n = k}} \binom{k}{k_1, \ldots, k_n} a_1^{k_1} \ldots a_n^{k_n},$$

където $\binom{k}{k_1, \ldots, k_n}$ означава т.н. полиномен коефициент:

$$\binom{k}{k_1, \ldots, k_n} = \frac{k!}{k_1! \ldots k_n!}.$$

Прилагайки тази формула за комутиращите оператори $h_1D_1, \dots h_nD_n$, получаваме

$$\left(D_{\vec{h}}\right)^k = \sum_{\substack{1 \le k_1, \dots, k_n \le n \\ k_1 + \dots + k_n = k}} \binom{k}{k_1, \dots, k_n} h_1^{k_1} \dots h_n^{k_n} \, D_1^{k_1} \dots D_n^{k_n}.$$

В крайна сметка ние получихме:

Теорема 3. *За дефинираната по-горе функция $\varphi(t)$ е в сила равенството*

$$\varphi^{(k)}(t) = \sum_{\substack{1 \le k_1, \dots, k_n \le n \\ k_1 + \dots + k_n = k}} \binom{k}{k_1, \dots, k_n} h_1^{k_1} \dots h_n^{k_n} \frac{\partial^k f}{\partial^{k_1} x_1 \dots \partial^{k_n} x_n} \left(x^0 + t\vec{h}\right).$$

Често горното равенство формално се записва във вида

$$\varphi^{(k)}(t) = \left(h_1D_1 + \dots + h_nD_n\right)^k f\left(x^0 + t\vec{h}\right).$$

За яснота ние отделно ще напишем тази формула за функции на две променливи. Нека $f(x, y)$ е m-кратно диференцируема функция на променливите x и y, дефинирана в околност на точката (x_0, y_0), $h, k \in \mathbb{R}$, и $\varphi(t) = f\left(x_0 + th, y_0 + tk\right)$. Тогава за всяко n между 1 и m е изпълнено равенството*

$$\varphi^{(n)}(t) = \left(h\frac{\partial}{\partial\, x} + k\frac{\partial}{\partial\, y}\right)^n f\left(x_0 + th, y_0 + tk\right) =$$

$$= \sum_{p=0}^{n} \binom{n}{p} h^p k^{n-p} \frac{\partial^n f}{\partial^p x\ \partial^{n-p} y} \left(x_0 + th, y_0 + tk\right),$$

където $\binom{n}{p}$ е биномният коефициент

$$\binom{n}{p} = \frac{n!}{p\,!\ (n-p)!}.$$

---

*Читателят може сам да докаже това равенство чрез индукция по n, използвайки свойствата на биномните коефициенти (виж част I, стр. 148). Доказателството дословно повтаря доказателството на бинома на Нютон.

Упражнения.

1. Нека $f(x,y)$ е функцията, определена с равенствата

$$f(x,y) = xy\,\frac{x^2-y^2}{x^2+y^2} \text{ при } (x,y) \neq (0,0)\,,\; f(0,0)=0.$$

Докажете, че $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ съществуват и са непрекъснати навсякъде в $\mathbb{R}^2$ и че смесените производни в нулата съществуват. Покажете, че

$$\frac{\partial^2 f}{\partial x\,\partial y}(0,0) \neq \frac{\partial^2 f}{\partial y\,\partial x}(0,0).$$

Обяснете защо този факт не противоречи на теорема 1.

2. Да предположим, че първите производни f'_x, f'_y и смесената производна f''_{xy} съществуват в околност на точката (x_0,y_0), и f''_{xy} е непрекъсната в тази точка. Тогава $f''_{yx}\,(x_0,y_0)$ също съществува, и имаме $f''_{yx}\,(x_0,y_0) = f''_{xy}\,(x_0,y_0)$.

Използвайте, че в означенията на теорема 1 имаме

$$\frac{f'_y\,(x_0+h,y_0) - f'_y\,(x_0,y_0)}{h} = \lim_{k\to 0}\frac{W(h,k)}{hk},$$

и доказаното в тази теорема равенство

$$\lim_{h\to 0,k\to 0}\frac{W(h,k)}{hk} = f''_{xy}\,(x_0,y_0)\,.$$

1.7 Формула на Тейлор за функции на много променливи. Локални екстремуми.

Формула на Тейлор. В началото ще припомним формулата на Тейлор за функции на една променлива.(Виж част I, §2.7.) Ако функцията $f(x)$ е дефинирана в околност на точката x_0 и притежава производни до ред n в тази околност, за достатъчно малки нараствания $h = \Delta x$ е в сила формулата

$$f\left(x_0+h\right) \; = \; \sum_{k=0}^{n} \frac{f^{(k)}\left(x_0\right)}{k!}\left(x-x_0\right)^k \; + \; R_n(h)$$

(ще напомним, че под нулева производна на една функция се разбира самата функция). За остатъка $R_n(h)$ е в сила оценката $R_n(h) = o(h)$. Ако $f(x)$ притежава и $n+1$-ва производна в x_0, е в сила по-точна оценка; най-често за остатъка се използва формулата на Лагранж

$$R_n(h) \; = \; \frac{f^{(n+1)}\left(x_0+\theta h\right)}{k!}\left(x-x_0\right)^{n+1}$$

за подходящо избрано $\theta \in (0,1)$.

В този параграф ние ще изведем аналогична формула за функция на много променливи. Ще използваме същия похват, както и в предишните параграфи: свеждане на въпроса към функция на една променлива.

Нека $f\left(x_1,\ldots,x_n\right)$ е m-кратно диференцируема функция на n променливи, дефинирана в околност на точката $x^0 = (x_1^0,\ldots,x_n^0)$. Нека $\vec{h} = \left(h_1,\ldots,h_n\right)$ е вектора на нарастването. Да предположим, че отсечката, свързваща точките x^0 и $x^0+\vec{h}$, се съдържа в дефиниционната област на функцията f. Тогава ние можем да образуваме познатата вече помощна функция

$$\varphi(t) = f\left(x_1^0+th_1,\ldots,x_n^0+th_n\right),$$

определена за $t \in [0,1]$. Както видяхме в предния параграф, тя притежава производни до ред m включително, и ние можем за всяко

$k \leq m-1$ да напишем за нея формулата на Тейлор с $t_0 = 0$ и $\Delta t = 1$:

$$\varphi(1) = \sum_{p=0}^{k} \frac{\varphi^{(p)}(0)}{p!} + R_k,$$

с $R_k = \frac{\varphi^{(k+1)}(\theta)}{(k+1)!}$ за някое $\theta \in (0,1)$. Тъй като $\varphi(0) = f\left(x^0\right)$ и $\varphi(1) = f\left(x^0 + \vec{h}\right)$, то припомняйки си формулите от предходния параграф за производните на $\varphi(t)$, получаваме *:

Формула на Тейлор с остатъчен член във вид на Лагранж. *Ако $f(x)$ притежава частни производни до ред $k+1$ включително в някаква околност на x^0, в сила е равенството*

$$f\left(x^0 + \vec{h}\right) = \sum_{p=0}^{k} \frac{1}{p!} \left(h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n} \right)^p f\left(x^0\right) \; + \; R_k\left(\vec{h}\right),$$

където

$$R_k\left(\vec{h}\right) \; = \; \frac{1}{(k+1)!} \left(h_1 \frac{\partial}{\partial x_1} + \ldots + h_n \frac{\partial}{\partial x_n} \right)^{k+1} f\left(x^0 + \theta \vec{h}\right)$$

за подходящо $\theta \in (0,1)$.

Забележка. За остатъчния член може да бъде използвана всяка от другите познати от част I формули - формата на Шлемилх-Рош или интегралната формула от зад. 3 на §4.3; разликите не са съществени.

Следствие (остатъчен член във формата на Пеано). *Нека $f(x)$ притежава частни производни до ред k включително, непрекъснати в някаква околност на x^0. Тогава за остатъка $R_k\left(\vec{h}\right)$ е в сила съотношението $R_k\left(\vec{h}\right) = o\left(\left\|\vec{h}\right\|^k\right)$, т.е.*

$$\lim_{\vec{h} \to 0} \frac{R_k\left(\vec{h}\right)}{\left\|\vec{h}\right\|^k} = 0.$$

*Тук за краткост използваме формулите за производните в съкратен вид. Точният смисъл на тези съкратени означения беше даден при тяхното извеждане в §6.

Доказателство. Да положим

$$\alpha_{k_1\ldots k_n}(\vec{h}) = \frac{\partial^k f}{\partial^{k_1}x_1\ldots\partial^{k_n}x_n}\left(x^0+\vec{h}\right) - \frac{\partial^k f}{\partial^{k_1}x_1\ldots\partial^{k_n}x_n}\left(x^0\right).$$

От непрекъснатостта на k-тите производни следва, че $\alpha_{k_1\ldots k_n}(\vec{h}) \to 0$ при $\vec{h}\to 0$.

От формулата на Лагранж за $k-1$-вия остатък $R_{k-1}\left(\vec{h}\right)$ имаме

$$R_{k-1}\left(\vec{h}\right) =$$

$$\sum_{k_1+\ldots+k_n=k}\binom{k}{k_1,\ldots,k_n} h_1^{k_1}\ldots h_n^{k_n}\left(\frac{\partial^k f}{\partial^{k_1}x_1\ldots\partial^{k_n}x_n}\left(x^0\right)+\alpha_{k_1\ldots k_n}(\theta\vec{h})\right),$$

откъдето получаваме

$$R_k\left(\vec{h}\right) = \sum_{k_1+\ldots+k_n=k}\binom{k}{k_1,\ldots,k_n} h_1^{k_1}\ldots h_n^{k_n}\,\alpha_{k_1\ldots k_n}(\theta\vec{h}).$$

Следователно

$$\frac{R_k\left(\vec{h}\right)}{\left|\left|\vec{h}\right|\right|^k} = \sum_{k_1+\ldots+k_n=k}\binom{k}{k_1,\ldots,k_n}\left(\frac{h_1}{\left|\left|\vec{h}\right|\right|}\right)^{k_1}\ldots\left(\frac{h_n}{\left|\left|\vec{h}\right|\right|}\right)^{k_n}\alpha_{k_1\ldots k_n}(\theta\vec{h}).$$

Тъй като $\alpha_{k_1\ldots k_n}(\vec{h}) \to 0$ при $\vec{h}\to 0$, и $h_i/\,||h|| \le 1$, това доказва твърдението. ■

Локални екстремуми. Определението на локален екстремум на функция на много променливи звучи по същия начин, както и за функция на една променлива:

Определение. *Нека функцията $f(x) = f\left(x_1,\ldots,x_n\right)$ е дефинирана в областта $D\subset\mathbb{R}^n$. Точката $x^0\in D$ се нарича точка на* <u>*локален максимум*</u> *за $f(x)$, ако*

а/ x^0 е вътрешна за D, и

б/ съществува $\varepsilon>0$ такова, че $f(x)\le f\left(x^0\right)$ за всяко $x\in D$, за което $||x-x^0||<\varepsilon$.

Ако в б/ вместо $f(x) \leq f(x^0)$ пишем $f(x) \geq f(x^0)$, получаваме определението на локален минимум.

Локалният максимум или минимум се нарича строг, ако при $x \neq x^0$ неравенството е строго.

Накрая, под точка на локален екстремум ще разбираме точка, която е или локален минимум, или локален максимум.

Ще напомним резултатите в едномерния случай (виж част I, §2.9): за да бъде точката x_0, вътрешна за дефиниционната област на диференцируемата функция $f(x)$, точка на локален екстремум, имаме

– необходимо условие: ако x_0 е локален екстремум, то $f'(x_0) = 0$ (теорема на Ферма).

– достатъчно условие: ако $f'(x_0) = 0$ и $f''(x_0) > 0$, то x_0 е локален минимум; ако $f'(x_0) = 0$ и $f''(x_0) < 0$, то x_0 е локален максимум. (доказва се чрез развитие по Тейлор от ред 2 на $f(x)$ около точката x_0.)

Елементарни примери показват, че нито необходимото условие е достатъчно, нито достатъчното условие е необходимо.

За функции на няколко променливи ситуацията е аналогична. Ще изложим необходимото условие:

Теорема 1. (необходимо условие за локален екстремум). *Нека точката $x^0 = (x_1^0, \ldots, x_n^0)$ е локален екстремум за функцията $f(x_1, \ldots, x_n)$. Ако в x^0 тази функция притежава частни производни, то всички те са равни на нула, т.е.*

$$\frac{\partial f}{\partial x_1}\left(x^0\right) = \ldots = \frac{\partial f}{\partial x_n}\left(x^0\right).$$

Доказателство. Ще използваме съответните частични функции на една променлива. По-подробно, за дадено i между 1 и n да означим чрез

$$\varphi_i(t) = f\left(x_1^0, \ldots, x_{i-1}^0, t, x_{i+1}^0, \ldots, x_n^0\right)$$

функцията, получена чрез фиксиране на всички други променливи освен променливата x_i. Очевидно функцията на едно променливо $\varphi_i(t)$ е дефинирана в околност на точката $t = x_i^0$ и притежава локален екстремум в тази точка (докажете!). Тогава според теоремата на Ферма за функции на едно променливо имаме

$$\frac{\partial f}{\partial x_i}\left(x^0\right) = \varphi'\left(x_i^0\right) = 0. \blacksquare$$

Удобно е да въведем следното понятие:

Определение. *Точка, вътрешна за дефиниционната област, в която всички частни производни на функцията f се анулират, наричаме* <u>*критична точка*</u> *за дадената функция.*

От горната теорема веднага следва едно полезно наблюдение:

Следствие 2. *Ако $f(x)$ е еднократно гладка непрекъсната функция, дефинирана върху компактното множество $D \subset \mathbb{R}^n$, то нейните най-голяма и най-малка стойности се достигат или върху контура на D, или в някоя от критичните точки на функцията.*

Доказателство. Наистина, според втората теорема на Вайерщрас минималната и максималната стойност на $f(x)$ се достигат в някои точки на D, които могат да бъдат вътрешни или контурни. Ако глобалният максимум се достига във вътрешна точка на D, то тази точка е също и локален максимум и следователно е критична точка. Разбира се, същото е вярно и за глобалния минимум. $\blacksquare$

За намиране на достатъчното условие ще изследваме члена от развитието по Тейлор, съдържащ вторите производни на функцията. Както ще видим, този член представлява квадратична функция на координатите на нарастването $h_1, \ldots, h_n$. Ще припомним някои понятия от линейната алгебра, засягащи такива функции (наричани обикновено квадратични форми). Нека $\mathcal{A}\vec{h}$ е квадратична форма в $\mathbb{R}^n$, дефинирана с формулата

$$A\vec{h} \; = \; \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij} \; h_i h_j.$$

Определение. *Квадратичната форма $\mathcal{A}$ се нарича*

– <u>*положително определена*</u> *(съотв.* <u>*отрицателно определена*</u>*), ако за всеки ненулев вектор $\vec{h} \in \mathbb{R}^n$, $\vec{h} \neq \vec{0}$, имаме $\mathcal{A}\vec{h} > 0$ (съотв. $\mathcal{A}\vec{h} < 0$).*

– знакопроменлива, ако съществуват вектори $\vec{h}^1$, $\vec{h}^2 \in \mathbb{R}^n$, така че $\mathcal{A}\vec{h}^1 > 0$, $\mathcal{A}\vec{h}^2 < 0$.

– положително (съотв. отрицателно) полуопределена, ако за всеки $\vec{h} \in \mathbb{R}^n$ имаме $\mathcal{A}\vec{h} \geq 0$ (съотв. $\mathcal{A}\vec{h} \leq 0$).

Примери. В пространството $\mathbb{R}^2$ с координати x, y можем да разгледаме квадратичните форми, зададени с формулите*:

- $x^2 + y^2$ – положително определена,
- $-x^2 - y^2$ – отрицателно определена,
- $x^2 - y^2$ – знакопроменлива,
- x^2 – положително полуопределена,
- $-x^2$ – отрицателно полуопределена.

Критерий за определеност. Нека $\mathcal{A}$ е написаната по-горе квадратична форма с коефициенти A_{ij}, $i, j = 1, \ldots, n$, и нека означим с Δ_k k-тият главен минор на матрицата от коефициентите:

$$\Delta_k = \det \{A_{ij}\}_{i,j=1,\ldots,k}$$

Тогава критерият на Силвестър гласи, че формата $\mathcal{A}$ е положително определена, когато $\Delta_k > 0$, и отрицателно определена, ако $(-1)^k \Delta_k > 0$, за $k = 1, \ldots, n$. Ако поне един от минорите с четни номера $\Delta_{2k} < 0$, то формата $\mathcal{A}$ е знакопроменлива.

Сега можем да формулираме достатъчното условие за екстремум.

Теорема 3. (достатъчно условие за локален екстремум.) *Нека $f(x_1, \ldots, x_n)$ е двукратно гладка функция, дефинирана в околност на точката x^0. Нека освен това x^0 е критична точка, т.е.*

$$\frac{\partial f}{\partial x_1}\left(x^0\right) = \ldots = \frac{\partial f}{\partial x_n}\left(x^0\right).$$

*В алгебрата се доказва, че всяка квадратична форма в $\mathbb{R}^2$ се привежда към една от изброените чрез линейна смяна на променливите.

Тогава

а/ ако квадратичната форма

$$\mathcal{D}f\left(x^0\right)\vec{h} \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i\,\partial x_j}\left(x^0\right)\;h_i h_j$$

е положително дефинирана, x^0 е строг локален минимум за f.

б/ ако $\mathcal{D}f\left(x^0\right)\vec{h}$ е отрицателно дефинирана, x^0 е строг локален максимум за f.

в/ ако $\mathcal{D}f\left(x^0\right)\vec{h}$ е знакопроменлива, x^0 не е локален екстремум за f.

Забележка. Нищо не може да се твърди в случаите, когато $\mathcal{D}f\left(x^0\right)\vec{h}$ е полуопределена: този случай може да има или да няма локален екстремум. Така, ако разгледаме функциите

$$f(x,y) = x^2 + y^4 \quad \text{и} \quad g(x,y) = x^2 - y^4$$

в $\mathbb{R}^2$, то в критичната точка $(0,0)$ имаме

$$\mathcal{D}f\left((0,0)\right)(h,k) = \mathcal{D}g\left((0,0)\right)(h,k) = h^2,$$

но първата функция притежава локален екстремум в тази точка, а втората – не.

В частност, оттук се вижда, че достатъчното условие от теоремата не е необходимо.

Доказателство на достатъчното условие. Ще ни е нужно помощно твърдение за квадратичните форми:

Лема. *Нека $\mathcal{A}\vec{h}$ е положително определена квадратична форма в $\mathbb{R}^n$. Тогава съществуват константи c и C, $0 < c \le C$, така че за всяко $\vec{h} \in \mathbb{R}^n$ имаме*

$$c\left|\left|\vec{h}\right|\right|^2 \le \mathcal{A}\vec{h} \le C\left|\left|\vec{h}\right|\right|^2.$$

Доказателство на лемата. Нека S е единичната сфера в $\mathbb{R}^n$, т.е. множеството от тези вектори $\vec{h} \in \mathbb{R}^n$, за които $\left|\left|\vec{h}\right|\right| = 1$. Очевидно множеството S е компактно, и по теоремите на Вайерщрас непрекъснатата функция $\mathcal{A}\vec{h}$ достига върху него своята най-малка и най-голяма стойност. Да ги означим съответно с c и C, $c \leq C$. От положителната определеност следва, че $c > 0$ (нулевият вектор не принадлежи на S). Нека сега $\vec{h}$ е произволен ненулев вектор. Тогава $\frac{\vec{h}}{||\vec{h}||} \in S$, и следователно

$$\mathcal{A}\left(\frac{\vec{h}}{\left|\left|\vec{h}\right|\right|}\right) = \frac{1}{\left|\left|\vec{h}\right|\right|^2}\mathcal{A}\vec{h} \in [c, C]. \blacksquare$$

Числата c и C, чието съществуване се установява от лемата, се наричат съответно <u>долна и горна граница на квадратичната форма</u>.

За доказателство на достатъчното условие ще напишем развитието на $f(x)$ около точката x^0 по Тейлор до втория член с остатъчен член във формата на Пеано. Тъй като първите производни се анулират в x^0, съответният член в тейлоровото развитие отсъствува. За достатъчно малки по норма вектори $\vec{h}$ получаваме

$$\Delta f = f\left(x^0 + \vec{h}\right) - f\left(x^0\right) \ = \ \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \, \partial x_j}\left(x^0\right) \ h_i h_j + o\left(\left|\left|\vec{h}\right|\right|^2\right),$$

откъдето получаваме

$$\frac{\Delta f}{\left|\left|\vec{h}\right|\right|^2} - \frac{1}{\left|\left|\vec{h}\right|\right|^2}\mathcal{D}f\left(x^0\right)\vec{h} \to_{\vec{h}\to 0} 0.$$

Да предположим, че квадратичната форма $\mathcal{D}f\left(x^0\right)$ е положително определена, и да означим с c нейната долна граница. Тогава за достатъчно малки по норма $\vec{h}$ имаме

$$\frac{\Delta f}{\left|\left|\vec{h}\right|\right|^2} - \frac{1}{\left|\left|\vec{h}\right|\right|^2}\mathcal{D}f\left(x^0\right)\vec{h} > -\frac{c}{2} \text{ и следователно } \frac{\Delta f}{\left|\left|\vec{h}\right|\right|^2} \geq c - \frac{c}{2} = \frac{c}{2} > 0,$$

т.е. $\Delta f > 0$ за достатъчно малки $\vec{h}$, което доказва точка а/. Точка б/ може да се докаже по същия начин, или като се приложи точка а/ за функцията $-f$.

Ако формата $\mathcal{D}f\left(x^0\right)$ е знакопроменлива, това означава, че съществуват вектори $\vec{h}^1, \vec{h}^2 \in \mathbb{R}^n$ такива, че $\mathcal{D}f\left(x^0\right)\vec{h}^1 > 0$ и

$\mathcal{D}f\left(x^0\right)\vec{h}^2 < 0$. Да заместим в горните формули вектора $\vec{h}$ с $t\vec{h}^1$. Имайки пред вид, че $\mathcal{D}f\left(x^0\right)t\vec{h}^1 = t^2\mathcal{D}f\left(x^0\right)\vec{h}^1$, за малки t имаме

$$\frac{f\left(x^0 + t\vec{h}^1\right) - f\left(x^0\right)}{t^2\left\|\vec{h}^1\right\|} = \frac{1}{\left\|\vec{h}^1\right\|}\mathcal{D}f\left(x^0\right)\vec{h}^1 + \frac{o\left(\left\|t\vec{h}^1\right\|\right)}{\left\|t\vec{h}^1\right\|^2}.$$

Тъй като второто събираемо клони към нула при $t \to 0$, разликата $f\left(x^0 + t\vec{h}^1\right) - f\left(x^0\right)$ е положителна при достатъчно малки стойности на t. По същия начин се вижда, че $f\left(x^0 + t\vec{h}^2\right) - f\left(x^0\right)$ е отрицателна при достатъчно малки стойности на t, и следователно в x^0 не може да има нито локален минимум, нито локален максимум. ■

Отново ще изложим отделно случая на функция на две променливи. Нека $f(x, y)$ е двукратно гладка функция в околност на критичната точка (x_0, y_0). Да дадем на x и на y нараствания съответно h и k. По формулата на Тейлор

$$\Delta f = f\left(x_0 + h, y_0 + k\right) - f\left(x_0, y_0\right) =$$

$$= \frac{1}{2}\left(h^2 f''_{xx}\left(x_0, y_0\right) + 2hk f''_{xy}\left(x_0, y_0\right) + k^2 f''_{yy}\left(x_0, y_0\right)\right) + o\left(\left(\sqrt{x^2 + y^2}\right)^2\right),$$

тъй като по предположение първите производни на функцията се анулират. Според теорема 3, знака на разликата отляво при достатъчно малки h и k съвпада със знака на първия член отдясно. Да означим за простота $f''_{xx}\left(x_0, y_0\right) = A$, $f''_{xy}\left(x_0, y_0\right) = B$, $f''_{yy}\left(x_0, y_0\right) = C$. Да предположим, че $A \neq 0$*. Имаме

$$Ah^2 + 2Bhk + Ck^2 = A\left(h^2 + 2\frac{B}{A}hk + \frac{C}{A}k^2\right) = A\left(\left(h + \frac{B}{A}k\right)^2 + \frac{AC - B^2}{A^2}k^2\right).$$

Оттук получаваме:

а/ При $AC - B^2 > 0$ и $A > 0$ имаме $\Delta f \geq 0$, т.е. (x_0, y_0) е локален минимум.

б/ При $AC - B^2 > 0$ и $A < 0$ имаме $\Delta f \leq 0$, т.е. (x_0, y_0) е локален минимум.

*Читателят лесно ще съобрази, че това предположение не намалява общността.

в/ Нека $AC - B^2 < 0$. Тогава за двойки (h,k), за които $h = -\frac{B}{A}k$, изразът в скобите е отрицателен, а за двойки от вида $(h,0)$ - положителен. Следователно нарастването Δf може да има различни знаци, и точката (x_0,y_0) не може да бъде точка на локален екстремум.

Забележка. Разбира се, доказаното по-горе веднага следва от теорема 3 и критерия на Силвестър. Наистина, в двумерния случай имаме

$$\Delta_1 = A \quad \text{и} \quad \Delta_2 = \begin{vmatrix} A & B \\ B & C \end{vmatrix} = AC - B^2.$$

Глобални екстремуми. В повечето случаи е важно да намерим най-голямата и на-малката стойности на дадена функция върху цялото дефиниционно множество, т.е. глобалният минимум и максимум на функцията. Да предположим, че дефиниционното множество D на функцията $f(x)$ е компактно множество в $\mathbb{R}^n$. Тогава по теоремата на Вайерщрас $f(x)$ достига най-малка и най-голямата си стойност в подходящи точки на D. Тогава получаваме следната алтернатива (например, за максималната стойност): глобалният максимум може да се достигне във вътрешна или контурна точка на D. Ако максимума се достига във вътрешна точка, то той е и локален, и ние можем да използваме развитата по-горе теория. Ние обаче все още не разполагаме със средства за откриване на екстремумите върху контура bD на дефиниционната област. За случая на "гладък"контур това ще бъде направено в §9.

В някои случаи подобни разсъждения могат да бъдат проведени и за некомпактни области – виж напр. задачи 2 и 3.

Приложение: Метод на най-малките квадрати. В този пункт ще дадем едно просто приложение на необходимото условие за локален екстремум, което обаче намира широко приложение в естествените науки. Нека x, y, z са някакви величини, и ние искаме да покажем, че между тях съществува линейна зависимост от вида

$$z = ax + by + c.$$

Проблема е да се намерят коефициентите a, b и c.

Разбира се, тяхното определяне трябва да се базира на направените измервания. Да предположим, че е проведена серия от n експеримента, в резултат на които са измерени величините $x_1, y_1, z_1, \ldots, x_n, y_n, z_n$. Трябва да се намерят такива стойности на коефициентите, така че експериментално измерените стойности z_k най-малко се различават от теоретично намерените по формулата $z_k = ax_k + by_k + c$. Метода на най-малките квадрати се състои в следното: да се минимизира сумата от квадратите на отклоненията. По-точно, ако означим

$$F(a,b,c) = \sum_{k=1}^{n} \left((ax_k + by_k + c) - z_k\right)^2,$$

да определим къде тази функция достига минималната си стойност. За критичните точки на тази функция получаваме уравненията:

$$F'_a = 2\sum_{k=1}^{n} x_k \left((ax_k + by_k + c) - z_k\right) = 0,$$

$$F'_b = 2\sum_{k=1}^{n} y_k \left((ax_k + by_k + c) - z_k\right) = 0,$$

$$F'_c = 2\sum_{k=1}^{n} \left((ax_k + by_k + c) - z_k\right) = 0.$$

Да въведем нови означения: Нека означим с $\vec{x}$, $\vec{y}$, $\vec{z}$ n-мерните вектори $\vec{x} = (x_1, \ldots, x_n)$, $\vec{y} = (y_1, \ldots, y_n)$, $\vec{z} = (z_1, \ldots, z_n)$, и нека $\vec{1}$ да означава n-мерният вектор, съставен само от единици. Тогава $F(a,b,c)$ се написва като

$$F(a,b,c) = \left|\left|\vec{z} - a.\vec{x} - b.\vec{y} - c.\vec{1}\right|\right|^2.$$

Уравненията на критична точка, след разкриване на скобите, са линейни относно a, b и c и имат вида:

$$a\left\langle \vec{x}, \vec{x}\right\rangle + b\left\langle \vec{x}, \vec{y}\right\rangle + c\left\langle \vec{x}, \vec{1}\right\rangle = \left\langle \vec{x}, \vec{z}\right\rangle,$$

$$a\left\langle \vec{y}, \vec{x}\right\rangle + b\left\langle \vec{y}, \vec{y}\right\rangle + c\left\langle \vec{y}, \vec{1}\right\rangle = \left\langle \vec{y}, \vec{z}\right\rangle,$$

$$a\left\langle \vec{1}, \vec{x}\right\rangle + b\left\langle \vec{1}, \vec{y}\right\rangle + c\left\langle \vec{1}, \vec{1}\right\rangle = \left\langle \vec{1}, \vec{z}\right\rangle.$$

В линейната алгебра се доказва, че ако векторите $\vec{x}$, $\vec{y}$ и $\vec{1}$ са линейно независими*, то детерминантата, съставена от коефициентите пред неизвестните, е различна от нула (т.н. детерминант на Грам). Следователно системата притежава единствено решение за a, b и c. Тъй като за функцията $F(a,b,c)$ е изпълнено

$$\lim_{||(a,b,c)||\to\infty} F(a,b,c) = +\infty,$$

то така намерената единствен критична точка е локален и глобален минимум на функцията (виж зад. 2).

Упражнения.

1. Нека в $\mathbb{R}^2$ е зададена функцията $f(x,y) = (y - x^2)\,(y - 3x^2)$. Докажете, че:

а/ $f(x,y)$ няма локален екстремум в точката $(0,0)$. (Упътване: разгледайте точките от вида $(0,y)$ и $(x, 2x^2)$).

б/ Ограничението на функцията върху произволна права през точката $(0,0)$ притежава строг локален минимум в тази точка.

2. Нека $f(x)$, $x \in \mathbb{R}^n$, е непрекъсната функция такава, че $\lim_{||x||\to\infty} f(x) = +\infty$. Докажете, че минималната стойност на $f(x)$ се достига в някоя точка.

3. Нека $f(x)$ е еднократно гладка функция в $\mathbb{R}^n$, удовлетворяваща условията на предната задача. Докажете, че ако $f(x)$ притежава само една критична точка, то в нея се достига глобалният минимум.

---

*Ако векторите $\vec{x}$, $\vec{y}$ и $\vec{1}$ са линейно зависими, това означава, че една от величините x, y се изразява линейно чрез другата и е излишна в дадения модел.

1.8 Теорема за неявната функция.

Първата теорема, разглеждана в този параграф, се отнася до решаване на уравнения от вида $F(x,y) = 0$ спрямо една от променливите, например y. Другата променлива - x - се разглежда като параметър, и очевидно решението трябва да зависи от нея. Да запишем това решение във вида $y = f(x)$. За да проверим дали така дефинираното y е решение на търсеното уравнение, трябва да го заместим в уравнението. Стигаме до равенството

$$F(x, f(x)) \equiv 0.$$

Определение. *Ако функцията $f(x)$ удовлетворява тъждествено горното равенство за всяко x от дефиниционната си област, ще казваме, че тя е неявна функция, определена от уравнението $F(x,y) = 0$.*

Геометрически това може да се каже така: множеството от точки в равнината, чиито координати удовлетворяват равенството $F(x,y) = 0$, да се представи като графика на функцията $f(x)$.

По-долу ние разглеждаме въпроса за съществуване и единственост на неявната функция. Разбира се, ние ще се интересуваме от неявни функции с хубави свойства, в частност диференцируеми. Ако $f(x)$ е такава функция, то, диференцирайки горното равенство по x и използвайки теоремата за диференциране на съставни функции, получаваме

$$F'_x(x, f(x)) + f'(x) F'_y(x, f(x)) \equiv 0,$$

откъдето

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

Оттук се вижда. че е естествено да наложим на функцията F условието $F'_y(x,y) \neq 0$.

Ще разгледаме един прост пример. Да разгледаме уравнението

$$F(x,y) = x^2 + y^2 - R^2 = 0$$

определящо окръжност с център в началото и радиус R в равнината $\mathbb{R}^2$. Ясно е, че за да има уравнението решение, трябва $x \in [-R, R]$. В крайните точки $x = \pm R$ се нарушава условието $F'_y(x, y) \neq 0$. Лесно се вижда, че в тези точки окръжността не може да бъде представена като графика на диференцируема функция (тангентата и става вертикална).

За всяко x от отворения интервал $(-R, R)$ това уравнение има точно две решения относно y:

$$y = \pm\sqrt{R^2 - x^2}.$$

Следователно всяка неявна функция, определена от горното уравнение, има вида

$$f(x) = \varepsilon(x)\sqrt{R^2 - x^2},$$

където $\varepsilon(x)$ е произволна функция, вземаща стойности плюс или минус единица. Такива функции има безбройно много.

Ако се интересуваме само от непрекъснати функции, получаваме само две неявни функции: $y = \sqrt{R^2 - x^2}$ и $y = -\sqrt{R^2 - x^2}$, т.е. отново няма еднозначност. За да се избегне нееднозначността, нещата трябва да се разглеждат локално – да се фиксира едно решение в дадена точка (тя може да лежи върху горната или долната полуокръжност), и да се разгледа непрекъснатата функция, вземаща съответната стойност в тази точка. Тогава, в зависимост от избраното решение в началната точка, ще получим уравнението на горната или долната полуокръжност.

Сега вече сме подготвени да дадем точната формулировка на теоремата:

Теорема 1. (Теорема за неявната функция, случай на едно уравнение). *Нека $F(x, y)$ е непрекъсната функция в $\mathbb{R}^2$, притежаваща непрекъсната производна по y. Нека (x_0, y_0) е точка от дефиниционното и множество, удовлетворяваща условията*

$$F(x_0, y_0) = 0\,,\ F'_y(x_0, y_0) \neq 0.$$

Тогава:

а/ Съществува интервал $(x_0 - \delta, x_0 + \delta)$ и дефинирана в него функция непрекъсната $f(x)$, удовлетворяващи условията

1/ $f(x_0) = y_0$, и

2/ $F(x, f(x)) \equiv 0$ за всяко $x \in (x_0 - \delta, x_0 + \delta)$.

б/ $f(x)$ е единствената непрекъсната функция в интервала $(x_0 - \delta, x_0 + \delta)$, удовлетворяваща условията 1/ и 2/.

в/ Ако $F(x, y)$ е диференцируема в точката (x_0, y_0), то $f(x)$ е диференцируема в x_0, като

$$f'(x_0) = -\frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}.$$

Ако $F(x, y)$ е k-кратно гладка в околност на точката (x_0, y_0), то същото е вярно и за $f(x)$ в околност на x_0.

Доказателство. За определеност можем да предположим, че $F'_y(x_0, y_0) > 0$. Тогава неравенството $F'_y(x, y) > 0$ е вярно и в някаква околност U на (x_0, y_0) в $\mathbb{R}^2$.

Да разгледаме $F(x_0, y)$ като функция на y; от горното следва, че тя е строго монотонно растяща в някаква околност на y_0. Тъй като $F(x_0, y_0) = 0$, то при достатъчно малки стойности на $\varepsilon > 0$ ще имаме

$$F(x_0, y) > 0 \,\text{при}\, y \in (y_0, y_0 + \varepsilon]\,, \quad \text{и} \quad F(x_0, y) < 0 \,\text{при}\, y \in [y_0 - \varepsilon, y_0)\,.$$

Да прекараме през точките $(x_0, y_0 - \varepsilon)$ и $(x_0, y_0 + \varepsilon)$ хоризонтални отсечки: тогава знакът на $F(x, y)$ се запазва в някаква околност на тези точки (виж чертежа, на който са означени знаците на $F(x, y)$ в съответните точки). Тогава можем да изберем (достатъчно малко) $\delta > 0$, така че

– за всяко $x \in (x_0 - \delta, x_0 + \delta)$ имаме $F(x, y_0 - \varepsilon) < 0$, $F(x, y_0 + \varepsilon) > 0$.

– правоъгълникът $\Delta = [x_0 - \delta, x_0 + \delta] \times [y_0 - \varepsilon, y_0 + \varepsilon]$ се съдържа в околността U (т.е. в него $F'_y > 0$).

Да фиксираме $x_1 \in (x_0 - \delta, x_0 + \delta)$; тогава функцията $F(x_1, y)$ е строго монотонно растяща в интервала $[y_0 - \varepsilon, y_0 + \varepsilon]$, взема отрицателна стойност в левия му край и положителна - в десния. Следователно съществува *единствено* $y_1 \in (y_0 - \varepsilon, y_0 + \varepsilon)$ такова, че $F(x_1, y_1) = 0$. Полагайки $f(x_1) = y_1$, получаваме функция, дефинирана в интервала $(x_0 - \delta, x_0 + \delta)$ и удовлетворяваща условията 1/ и 2/ от точка а/.

Ще отбележим едно следствие от горната конструкция: ако (x, y) е произволна точка от Δ такава, че $F(x, y) = 0$, то $y = f(x)$.

Ще докажем непрекъснатостта на така дефинираната функция. Най-напред ще докажем непрекъснатостта в точката x_0. Горната конструкция може да се изложи по следния начин: за всяко достатъчно малко $\varepsilon > 0$ ние намерихме $\delta > 0$, така че за всяко x, за което $|x - x_0| < \delta$, ще имаме $|f(x) - f(x_0)| < \varepsilon$; това е точно дефиницията на Коши за непрекъснатост.

Да вземем сега произволно $x_1 \in (x_0 - \delta, x_0 + \delta)$, и нека $y_1 = f(x_1)$; тогава ние можем да повторим горната конструкция, избирайки достатъчно малко $\varepsilon_1 > 0$ и зависеща от него $\delta_1 > 0$, и конструирайки неявната функция $\tilde{f}(x)$ за $x \in (x_1 - \delta_1, x_1 + \delta_1)$. Както доказахме, функцията $\tilde{f}(x)$ е непрекъсната в x_1.

Ние можем да изберем ε_1 и δ_1 толкова малки, че правоъгълникът $\Delta_1 = [x_1 - \delta_1, x_1 + \delta_1] \times [y_1 - \varepsilon_1, y_1 + \varepsilon_1]$ да се съдържа в Δ. Тогава, поради отбелязаната по-горе единственост, имаме $\tilde{f}(x) = f(x)$ за $x \in (x_1 - \delta_1, x_1 + \delta_1)$, и следователно $f(x)$ е непрекъсната в x_1. С това подточка а/ е доказана.

Подточка б/: Нека $\tilde{f}(x)$ е непрекъсната функция, дефинирана в околност на x_0 и удовлетворяваща условията 1/ и 2/. Тогава за стойности на x, достатъчно близки до x_0, точката с координати $\left(x, \tilde{f}(x)\right)$ ще принадлежи на правоъгълника Δ. От гореказаното се вижда, че тогава $\left(x, \tilde{f}(x)\right) = (x, f(x))$, т.е. $\tilde{f}(x)$ и $f(x)$ ще съвпадат в някаква околност на точката x_0.

Да означим сега с x_1 точната горна граница на точките, за които $\tilde{f}(x) = f(x)$; ако допуснем, че x_1 е в дефиниционната област на $\tilde{f}(x)$ и $f(x)$, то повтаряйки същите разсъждения за x_1 вместо за x_0, получаваме, че $\tilde{f}(x)$ и $f(x)$ съвпадат в околност на x_1, т.е. противоречие. С това б/ е доказано.

Доказателство на в/: Ако ни е известно, че неявната функция $f(x)$ е диференцируема, то, както беше показано в началото на параграфа, формулата за $f'(x_0)$ се получава чрез диференциране на равенството $F(x, f(x)) \equiv 0$.

За да докажем диференцируемостта на $f(x)$, ще използваме формулата за нарастването за диференцируемата функция $F(x, y)$. Нека x е достатъчно близко до x_0, $y = f(x)$, $\Delta y = y - y_0 = f(x) - f(x_0)$. Тогава

$$\Delta F = F(x, y) - F(x_0, y_0) = 0 - 0 = 0,$$

и формулата за нарастването дава

$$0 = \Delta F = F'_x(x_0, y_0)\,\Delta x + F'_y(x_0, y_0)\,\Delta y + \alpha(\Delta x, \Delta y)\,\Delta x + \beta(\Delta x, \Delta y)\,\Delta y,$$

където $\alpha(\Delta x, \Delta y), \beta(\Delta x, \Delta y) \to 0$ при $\Delta x, \Delta y \to 0$. Тогава

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = - \lim_{\Delta x \to 0} \frac{F'_x(x_0, y_0) + \alpha(\Delta x, \Delta y)}{F'_y(x_0, y_0) + \beta(\Delta x, \Delta y)} = - \frac{F'_x(x_0, y_0)}{F'_y(x_0, y_0)}.$$

Оттук следва диференцируемостта на $f(x)$ в x_0, както и формулата за производната и; разбира се, същата формула е валидна за всяко x от дефиниционния интервал.

Да допуснем, че $F(x, y)$ притежава производни до ред k: тогава, диференцирайки доказаната формула за $f'(x)$, получаваме k-кратната диференцируемост на $f(x)$. ■

Геометрична интерпретация. Нека $F(x, y)$ е еднократно гладка функция в $\mathbb{R}^2$. Да означим с M множеството на нулите на $F(x, y)$, т.е. от точките (x, y) в равнината, за които $F(x, y) = 0$. Ще предположим, че в нито една точка от M двете първи частни производни $F_x(x, y)$ и $F_y(x, y)$ не се анулират едновременно. Тогава доказаната по-горе теорема допуска следната геометрична интерпретация:

Теорема 2. *При горното условие множеството M локално (т.е. в някаква околност на всяка своя точка) се представя като графика на гладка функция.*

Наистина, да фиксираме някаква точка $(x_0, y_0) \in M$. Ако имаме $F'_y(x_0, y_0) \neq 0$, и $f(x)$ е съответната неявна функция, от горното доказателство се вижда, че в някаква околност на (x_0, y_0) равенствата $F(x, y) = 0$ и $y = f(x)$ са еквивалентни. Аналогично, ако $F'_x(x_0, y_0) \neq 0$, то около (x_0, y_0) множеството M се представя с уравнението $x = f(y)$.

Забележка. Както знаем, графиките на гладките функции са частен случай на регулярни параметрично зададени криви (виж част I, §2.12). Следователно локално множеството M притежава регулярна параметризация. Може да се докаже, че в такъв случай локалните параметризации могат да бъдат "слепени"и да се получи регулярна параметризация на цялата крива.

В крайна сметка геометричният смисъл на теорема 2 може да се формулира по следния начин:

Нека $F(x, y)$ е еднократно гладка функция на две променливи, като за всички (x, y) имаме $\overrightarrow{\mathrm{grad}}\, F(x, y) \neq \vec{0}$. Тогава множеството

M, определено с уравнението $F(x,y) = 0$, представлява регулярна еднократно гладка крива в равнината.

Доказателство на теорема 2 от §5. Сега ние можем да докажем теорема 2 от §5, която твърди, че градиентът на функция на две променливи в дадена точка е перпендикулярен на линията на ниво, минаваща през тази точка.

Да уточним формулировката: Нека $\Gamma \subset \mathbb{R}^2$ е регулярна параметрично зададена крива, $(x_0, y_0) \in \Gamma$, и $\vec{v}$ е вектор с начало в точката (x_0, y_0). Ще казваме, че *$\vec{v}$ е перпендикулярен на* Γ, ако $\vec{v}$ е перпендикулярен на допирателната права към Γ в (x_0, y_0).

Нека е дадена еднократно гладката функция $F(x,y)$ на две променливи и (x_0, y_0) е точка от дефиниционната и област. Теоремата има смисъл , ако $\overrightarrow{\operatorname{grad}} F(x_0, y_0) \neq \vec{0}$. Да предположим, че $F'_y(x_0, y_0) \neq 0$. Нека $F(x_0, y_0) = a$; тогава линията на ниво L_a, минаваща през (x_0, y_0), се представя с уравнението $F(x,y) - a = 0$. Очевидно производната по y на лявата страна на това равенство съвпада с $F'_y(x_0, y_0)$ и е различна от нула. Ако означим с $f(x)$ неявната функция, определена от това равенство, то в околност на (x_0, y_0) множеството L_a съвпада с графика на $f(x)$. Допирателната към тази графика в (x_0, y_0) е колинеарна с вектора $\vec{l}(x_0) = (1, f'(x_0))$. Така ортогоналността на векторите $\vec{l}(x_0)$ и $\overrightarrow{\operatorname{grad}} F(x_0, y_0)$ се свежда до равенството

$$F'_x(x_0, y_0) + f'(x_0) F'_y(x_0, y_0) = 0,$$

което беше получено по-горе чрез диференциране на тъждеството $F(x, f(x)) \equiv 0$.

Теорема за неявната функция за едно уравнение и няколко параметъра. В следващия по сложност случай, който ще разгледаме, отново имаме едно уравнение $F(x,y) = 0$, което трябва да бъде решено относно променливата y, но в този случай параметъра x е вече векторна - n-мерна - променлива, т.е. $x = (x_1, \ldots, x_n)$. Формулировката и доказателството в този случай почти напълно съвпадат с дадените по-горе, и ние само ще формулираме теоремата.

Теорема 3. (Теорема за неявната функция, случай на едно уравнение и няколко параметъра). *Нека $F(x,y) = F(x_1, \ldots, x_n, y)$*

е непрекъсната функция в $\mathbb{R}^{n+1}$, притежаваща непрекъсната частна производна по y. Нека $(x^0, y_0) = (x_1^0, \ldots, x_n^0, y_0)$ е точка от дефиниционното и множество, удовлетворяваща условията

$$F\left(x^0, y_0\right) = 0\,,\; F'_y\left(x^0, y_0\right) \neq 0.$$

Тогава:

а/ За достатъчно малко $\delta > 0$ съществува функция $f(x) = f(x_1, \ldots, x_n)$, дефинирана и непрекъсната в кълбото $B(x^0, \delta) \subset \mathbb{R}^n$ с център x^0 и радиус δ, и удовлетворяваща условията:

1/ $f(x^0) = y_0$, и

2/ $F(x, f(x)) \equiv 0$ за всяко $x \in B(x^0, \delta)$.

б/ $f(x)$ е единствената непрекъсната функция в кълбото $B(x^0, \delta)$, удовлетворяваща условията 1/ и 2/.

в/ Ако $F(x, y)$ е диференцируема в точката (x^0, y_0), то $f(x)$ е диференцируема в x^0, като

$$\frac{\partial f}{\partial x_i}\left(x^0\right) = -\frac{\frac{\partial F}{\partial x_i}(x_0, y_0)}{\frac{\partial F}{\partial y}(x_0, y_0)},\; i = 1, \ldots, n.$$

Теорема за неявната функция (общ случай). Тук ще формулираме и докажем теорема, аналогична на дадената по-горе, за случая на n уравнения с n неизвестни и произволен брой параметри. Като начало ще разгледаме случая $n = 2$. Нека имаме уравненията

$$F(x, y, u, v) = 0,$$

$$G(x, y, u, v) = 0,$$

където F и G са еднократно гладки функции на четири променливи, и нека нашата цел е да ги разрешим относно променливите u и v, т.е. да изразим u и v чрез параметрите x и y. По-точно, ние искаме да намерим функции $u(x, y)$ и $v(x, y)$ такива, че при заместването им в уравненията да получим тъждества относно x и y:

$$F(x, y, u(x, y), v(x, y)) \equiv 0,$$

$$G(x, y, u(x, y), v(x, y)) \equiv 0.$$

Да предположим, че диференцируемите функции $u(x, y)$ и $v(x, y)$ са вече намерени, да се опитаме да пресметнем техните производни, например, по x. Диференцирайки по x горните тъждества, получаваме

$$F'_x + u'_x F'_u + v'_x F'_v \equiv 0,$$

$$G'_x + u'_x G'_u + v'_x G'_v \equiv 0.$$

Тези равенства могат да се разглеждат като система от линейни уравнения относно неизвестните u'_x, v'_x. Както знаем от линейната алгебра, ако детерминантата от коефициентите пред неизвестните не се анулира, те имат единствено решение, зададено с формулите на Крамер. С други думи, ако предположим, че

$$\begin{vmatrix} F'_u & F'_v \\ G'_u & G'_v \end{vmatrix} = \frac{D(F, G)}{D(u, v)} \neq 0,$$

то частните производни на $u(x, y)$ и $v(x, y)$ относно x ще се задават с формулите

$$u'_x = -\frac{\begin{vmatrix} F'_x & F'_v \\ G'_x & G'_v \end{vmatrix}}{\begin{vmatrix} F'_u & F'_v \\ G'_u & G'_v \end{vmatrix}},\ v'_x = -\frac{\begin{vmatrix} F'_u & F'_x \\ G'_u & G'_x \end{vmatrix}}{\begin{vmatrix} F'_u & F'_v \\ G'_u & G'_v \end{vmatrix}}.$$

Ако се опитаме да намерим u'_y, v'_y, отново ще стигнем до условието $\dfrac{D(F, G)}{D(u, v)} \neq 0$; ясно е, че това условие замества условието $F'_y \neq 0$, налагано в случая на едно уравнение. Имайки това пред вид, вече можем да формулираме общия вид на теоремата - случая на n уравнения с n неизвестни и m параметъра:

Теорема 4. (Теорема за неявната функция - общ случай.) *Нека*

$$F_1(x_1, \ldots, x_m, y_1, \ldots, y_n), \ldots, F_n(x_1, \ldots, x_m, y_1, \ldots, y_n)$$

са n еднократно гладки функции, дефинирани в отворено подмножество на $\mathbb{R}^{n+m}$. Нека $(x^0, y^0) = (x_1^0, \ldots, x_m^0, y_1^0, \ldots, y_n^0)$ е точка от дефиниционното *им множество, така че*

$$F_1\left(x^0, y^0\right) = \ldots = F_n\left(x^0, y^0\right) = 0, \quad \frac{D\left(F_1, \ldots, F_n\right)}{D\left(y_1, \ldots, y_n\right)}\left(x^0, y^0\right) \neq 0.$$

Тогава:

а/ За достатъчно малко $\delta > 0$ съществуват функции $f_1(x) = f_1\left(x_1, \ldots, x_m\right), \ldots, f_n(x) = f_n\left(x_1, \ldots, x_m\right)$, дефинирани и непрекъснати в кълбото $B\left(x^0, \delta\right) \subset \mathbb{R}^m$ с център x^0 и радиус δ, и удовлетворяващи условията:

1/ $f_1\left(x^0\right) = y_1^0, \ldots, f_n\left(x^0\right) = y_n^0$, и

2/ $F_j\left(x_1, \ldots, x_m, f_1(x), \ldots, f_n(x)\right) \equiv 0$ за $j = 1, 2, \ldots, n$ и за всяко $x = \left(x_1, \ldots, x_m\right) \in B\left(x^0, \delta\right)$.

б/ $f_1(x), \ldots, f_n(x)$ е единствената система от n функции в кълбото $B\left(x^0, \delta\right)$, удовлетворяваща условията 1/ и 2/.

в/ Функциите $f_1(x), \ldots, f_n(x)$ са диференцируеми в точката x^0.

Доказателството на теоремата се извършва чрез индукция по броя на уравненията n. За да обясним обаче идеята по-добре, ще изложим отделно доказателството на частния случай, разгледан по-горе.

Доказателство в случая $n = 2$. Трябва да решим уравненията $F(x, y, u, v) = 0$, $G(x, y, u, v) = 0$ относно u и v в околност на точката (x_0, y_0, u_0, v_0) при условие, че $F\left(x_0, y_0, u_0, v_0\right) = G\left(x_0, y_0, u_0, v_0\right) = 0$ и $\frac{D(F,G)}{D(u,v)} \neq 0$. Ще следваме обичайния метод на решаване на системи от две уравнения - ще определим едната от неизвестните величини от едното уравнение и ще я заместим в другото, като получим в резултат едно уравнение с едно неизвестно.

Да отбележим най-напред, че от условието на теоремата следва, че поне една от частните производни F'_u, F'_v, G'_u, G'_v не се анулира в точката (x_0, y_0, u_0, v_0); може да предположим, че $G'_v\left(x_0, y_0, u_0, v_0\right) \neq 0$. Следователно може да приложим теоремата за неявната функция за уравнението

$$G(x, y, u, v) = 0$$

и да го решим относно променливата v. Получаваме функция $\varphi(x, y, u)$ със свойствата

$$G(x, y, u, \varphi(x, y, u)) \equiv 0, \quad \varphi\left(x_0, y_0, u_0\right) = v_0.$$

Чрез диференциране на горното тъждество по u получаваме

$$\varphi_u'(x,y,u) = -\frac{G_u'(x,y,u,\varphi(x,y,u))}{G_v'(x,y,u,\varphi(x,y,u))}.$$

Да заместим променливата v в първото уравнение с $\varphi(x,y,u)$; получаваме уравнението

$$\widetilde{F}(x,y,u) \overset{def}{=} F(x,y,u,\varphi(x,y,u)) = 0.$$

По теоремата за диференциране на съставни функции получаваме

$$\widetilde{F}_u' = F_u' + \varphi_u' F_v' = F_u' - \frac{G_u'}{G_v'} F_v' = \frac{\frac{D(F,G)}{D(u,v)}}{G_v'},$$

което по условие не се анулира в точката (x_0, y_0, u_0, v_0). Следователно можем да приложим теоремата за неявната функция към уравнението $\widetilde{F}(x,y,u) = 0$. Да означим получената функция с $u(x,y)$, и нека $v(x,y) = \varphi(x,y,u(x,y))$. Лесно се вижда, че така построената двойка функции $u(x,y)$, $v(x,y)$ удовлетворява исканите условия, т.е.

$$F(x,y,u(x,y),v(x,y)) \equiv 0, \quad G(x,y,u(x,y),v(x,y)) \equiv 0,$$

$$u(x_0,y_0) = u_0, \quad v(x_0,y_0) = v_0. \blacksquare$$

Доказателство в общия случай. Доказателството се извършва чрез индукция по броя на уравненията n (броят на параметрите не е от значение). Отново се използва методът на заместването. От условието $\frac{D(F_1,\ldots,F_n)}{D(y_1,\ldots,y_n)}(x^0,y^0) \neq 0$ следва, че поне един от елементите на съответната матрица не се анулира. Сменяйки, ако е необходимо, номерацията на уравненията и неизвестните, може да считаме, че

$$\frac{\partial F_n}{\partial y_n}\left(x^0, y^0\right) \neq 0.$$

Следователно към уравнението

$$F_n(x_1,\ldots,x_m,y_1,\ldots,y_n) = 0$$

може да бъде приложена теоремата за неявната функция относно променливата y_n. Получаваме функция $y_n = \varphi(x_1, \dots, x_m, y_1, \dots, y_{n-1})$ такава, че равенството

$$F_n(x_1, \dots, x_m, y_1, \dots, y_{n-1}, \varphi(x_1, \dots, x_m, y_1, \dots, y_{n-1})) = 0$$

е тъждествено изпълнено по $x_1, \dots, x_m, y_1, \dots, y_{n-1}$. Чрез диференциране получаваме равенствата

$$\frac{\partial F_n}{\partial y_i} + \frac{\partial \varphi}{\partial y_i} \cdot \frac{\partial F_n}{\partial y_n} = 0 \quad \text{за} \quad i = 1, \dots, n-1.$$

Да въведем сега функциите

$$\widetilde{F}_j(x_1, \dots, x_m, y_1, \dots, y_{n-1}) \stackrel{def}{=}$$

$$\stackrel{def}{=} F_j(x_1, \dots, x_m, y_1, \dots, y_{n-1}, \varphi(x_1, \dots, x_m, y_1, \dots, y_{n-1})) \text{ за } j = 1, \dots, n-1.$$

Отново основната трудност е в доказателството на равенството

$$\frac{D\left(\widetilde{F}_1, \dots, \widetilde{F}_{n-1}\right)}{D(y_1, \dots, y_{n-1})}\left(x_1^0, \dots, x_m^0, y_1^0, \dots, y_{n-1}^0\right) \neq 0.$$

За $i, j = 1, \dots, n-1$ имаме

$$\frac{\partial \widetilde{F}_j}{\partial y_i} = \frac{\partial F_j}{\partial y_i} + \frac{\partial \varphi}{\partial y_i} \cdot \frac{\partial F_n}{\partial y_n}.$$

Да разгледаме сега детерминантата

$$\frac{D(F_1, \dots, F_n)}{D(y_1, \dots, y_n)} = \left|\begin{array}{ccc|c} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_{n-1}} & \frac{\partial F_1}{\partial y_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_{n-1}}{\partial y_1} & \dots & \frac{\partial F_{n-1}}{\partial y_{n-1}} & \frac{\partial F_{n-1}}{\partial y_n} \\ \hline \frac{\partial F_n}{\partial y_1} & \dots & \frac{\partial F_n}{\partial y_{n-1}} & \frac{\partial F_n}{\partial y_n} \end{array}\right|$$

Да преобразуваме тази детерминанта, като вземем последния стълб, умножен по $\frac{\partial \varphi}{\partial y_1}$ и го прибавим към първия, умножен по $\frac{\partial \varphi}{\partial y_2}$ го прибавим към втория,... , умножен по $\frac{\partial \varphi}{\partial y_{n-1}}$ го прибавим към $n-1$-вия.

Както знаем, при тези операции стойността на детерминантата не се изменя. Като вземем пред вид получените по-горе равенства, имаме

$$\frac{D(F_1,\ldots,F_n)}{D(y_1,\ldots,y_n)} = \left|\begin{array}{ccc|c} \frac{\partial \widetilde{F}_1}{\partial y_1} & \cdots & \frac{\partial \widetilde{F}_1}{\partial y_{n-1}} & \frac{\partial F_1}{\partial y_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \widetilde{F}_{n-1}}{\partial y_1} & \cdots & \frac{\partial \widetilde{F}_{n-1}}{\partial y_{n-1}} & \frac{\partial F_{n-1}}{\partial y_n} \\ \hline 0 & \cdots & 0 & \frac{\partial F_n}{\partial y_n} \end{array}\right| = \frac{D\left(\widetilde{F}_1,\ldots,\widetilde{F}_{n-1}\right)}{D(y_1,\ldots,y_{n-1})} \cdot \frac{\partial F_n}{\partial y_n},$$

откъдето получаваме търсеното неравенство $\frac{D\left(\widetilde{F}_1,\ldots,\widetilde{F}_{n-1}\right)}{D(y_1,\ldots,y_{n-1})} \neq 0$.

Сега според индуктивното предположение можем да решим системата от уравнения $\widetilde{F}_1 = 0, \ldots, \widetilde{F}_{n-1} = 0$, относно неизвестните $y_1, \ldots, y_{n-1}$, т.е. да получим функции $f_1(x), \ldots, f_{n-1}(x)$, които, заместени в тези уравнения, да дават тъждества по x. За да намерим неизвестното y_n, да положим

$$f_n(x_1,\ldots,x_n) = \varphi\left(x_1,\ldots,x_n, f_1(x_1,\ldots,x_n),\ldots,f_{n-1}(x_1,\ldots,x_n)\right).$$

От конструкцията е ясно, че замествайки в уравненията $F_1 = 0, \ldots, F_n = 0$ неизвестните $y_1, \ldots, y_n$ с функциите $f_1(x), \ldots, f_n(x)$, получаваме тъждества по x. Лесно се проверяват и равенствата $f_i\left(x_1^0,\ldots,x_n^0\right) = y_i^0$, $i = 1,\ldots,n$. Теоремата за неявната функция е доказана. ■

Геометрична интерпретация. Както вече направихме в частния случай на едно уравнение, теоремата за неявната функция може да бъде използвана за описание на някои геометрични обекти. По-специално, нека $F_1(x_1,\ldots,x_{n+k}), \ldots, F_n(x_1,\ldots,x_{n+k})$ са n на брой еднократно гладки функции, зададени в някаква област в $\mathbb{R}^{n+k}$. Нека M е подмножеството на $\mathbb{R}^{n+k}$, определено с уравненията

$$F_1(x_1,\ldots,x_{n+k}) = 0,$$

$$\cdots\cdots\cdots\cdots$$

$$F_n(x_1,\ldots,x_{n+k}) = 0.$$

Без допълнителни условия върху функциите $F_i(x)$ структурата на множеството M трудно може да бъде описана; може да се случи например измежду тези функции да има повтарящи се, или някои от тях да са тъждествено нула. Затова върху функциите обикновено се налага следното:

Условие за регулярност: Казваме, че функциите $F_1(x), \ldots, F_n(x)$ образуват регулярна система от функции , ако техните градиенти $\overrightarrow{\mathrm{grad}}\, F_1(x), \ldots, \overrightarrow{\mathrm{grad}}\, F_n(x)$ са линейно независими за всяко x.

При горното предположение множеството M може (поне локално) да бъде лесно описано:

Теорема 5. *Ако функциите $F_1, \ldots, F_n$ образуват регулярна система, то множеството M от техните общи нули може локално (в околност на всяка своя точка) да се представи като графика на изображение от $\mathbb{R}^k$ в $\mathbb{R}^n$.*

Доказателство. Да разгледаме матрицата с n реда и $n+k$ стълба, образувана от координатите на векторите $\overrightarrow{\mathrm{grad}}\, F_1\left(x^0\right), \ldots, \overrightarrow{\mathrm{grad}}\, F_n\left(x^0\right)$:

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_1}\left(x^0\right) & \frac{\partial F_1}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial F_1}{\partial x_{n+k}}\left(x^0\right) \\ \frac{\partial F_2}{\partial x_1}\left(x^0\right) & \frac{\partial F_2}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial F_2}{\partial x_{n+k}}\left(x^0\right) \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial F_n}{\partial x_1}\left(x^0\right) & \frac{\partial F_n}{\partial x_2}\left(x^0\right) & \ldots & \frac{\partial F_n}{\partial x_{n+k}}\left(x^0\right) \end{pmatrix}$$

От линейната алгебра знаем, че векторите на редовете на една матрица са линейно независими точно тогава, когато тя е от максимален ранг. Следователно горната матрица е от ранг n, т.е. тя притежава поне една различна от нула поддетерминанта от ред n. Всяка такава детерминанта се получава чрез избор на n различни стълба на матрицата; ако това са стълбовете с номера $i_1, \ldots, i_n$, $1 \leq i_1 < \ldots < i_n \leq n+k$, то получената детерминанта е равна на

$$\frac{D\left(F_1, \ldots, F_n\right)}{D\left(x_{i_1}, \ldots, x_{i_n}\right)}\left(x^0\right).$$

Поне една от тези детерминанти трябва да е различна от нула; размествайки, ако е необходимо, номерацията на променливите, можем да считаме, че това е детерминантата, образувана от първите n стълба, т.е.

$$\frac{D\left(F_1, \ldots, F_n\right)}{D\left(x_1, \ldots, x_n\right)}\left(x^0\right) \neq 0.$$

Сега можем да приложим теоремата за неявната функция и да решим уравненията $F_1(x) = \ldots = F_n(x) = 0$ относно променливите $x_1, \ldots, x_n$. Получаваме, че в някаква околност на x^0 координатите на точките от M удовлетворяват уравненията

$$x_1 = f_1\left(x_{n+1}, \ldots, x_{n+k}\right),$$
$$\ldots\ldots\ldots$$
$$x_n = f_n\left(x_{n+1}, \ldots, x_{n+k}\right),$$

т.е. локално множеството M се представя като графиката на изображението от $\mathbb{R}^k$ в $\mathbb{R}^n$, определено от функциите $f_1, \ldots, f_n$. ■

Сега ще докажем многомерния аналог на теорема 2 от §5. Ще използваме въведеното в §4 понятие за допирателно подпространство към графиката на изображение.

Теорема 6. *Допирателното подпространство към M в точката x^0 съвпада с пространството от всички вектори в $\mathbb{R}^{n+k}$, ортогонални на векторите $\overrightarrow{\operatorname{grad}} F_1\left(x^0\right), \ldots, \overrightarrow{\operatorname{grad}} F_n\left(x^0\right)$.*

С други думи, векторите $\overrightarrow{\operatorname{grad}} F_1\left(x^0\right), \ldots, \overrightarrow{\operatorname{grad}} F_n\left(x^0\right)$ образуват базис в ортогоналното допълнение $T_{x^0}M^{\perp}$ към допирателното пространство $T_{x^0}M$.

Доказателство. Нека, както по-горе, представим локално M като графика на изображението определено от функциите $f_1, \ldots, f_n$. С други думи, локално M се параметризира с променливите $x_{n+1}, \ldots, x_{n+k}$. За всяко $j = 1, \ldots, n$ е изпълнено тъждеството

$$F_j\left(f_1\left(x_{n+1}, \ldots, x_{n+k}\right), \ldots, f_n\left(x_{n+1}, \ldots, x_{n+k}\right), x_{n+1}, \ldots, x_{n+k}\right) \equiv 0.$$

Да диференцираме това тъждество по променливата x_{n+p}, където p е между 1 и n. Получаваме:

$$\sum_{i=1}^{n} \frac{\partial F_j}{\partial x_i} \frac{\partial f_i}{\partial x_{n+p}} + \frac{\partial F_j}{\partial x_{n+p}} \equiv 0.$$

От друга страна, в §4 видяхме, че допирателното пространство към M се поражда от векторите $\vec{l_p}$, $p = 1, \ldots, k$:

$$\vec{l_p} = \left(\frac{\partial f_1}{\partial x_{n+p}}, \ldots, \frac{\partial f_n}{\partial x_{n+p}}, 0, \ldots, 0, 1, 0, \ldots, 0\right),$$

като единицата е разположена на p-то място измежду последните k координати на вектора $\vec{l_p}$. * Тогава горното тъждество може да се напише във вида

$$\left\langle \vec{l_p}, \overrightarrow{\mathrm{grad}}\, F_j \right\rangle = 0.$$

Това равенство е изпълнено за всички $j = 1, \dots, n$ и $p = 1, \dots, k$. От условието на теоремата следва, че ортогоналното допълнение на векторите $\overrightarrow{\mathrm{grad}}\, F_1\left(x^0\right), \dots, \overrightarrow{\mathrm{grad}}\, F_n\left(x^0\right)$ има размерност k. Следователно то съвпада с линейното пространство, породено от векторите $\vec{l_1}, \dots, \vec{l_k}$, т.е. с допирателното пространство към M. ■

Забележка. От доказаната теорема се вижда, че допирателното пространство към M не зависи от начина на представянето му като графика, т.е от това, кои измежду променливите $x_1, \dots, x_{n+k}$ ще изберем за независими променливи. Нещо повече: от теоремата се вижда, че допирателното подпространство не зависи от избора на ортогонална координатна система в даденото евклидово пространство. Наистина, както беше показано в §4, градиентът на дадена функция има геометричен смисъл и не зависи от координатната система,а само от скаларното произведение в евклидовото пространство.

Теорема за обратното изображение. Едно от важните приложения на теоремата за неявната функция е теоремата за обратното изображение. Нека

$$f(x) = \left(f_1\left(x_1, \dots, x_n\right), \dots, f_n\left(x_1, \dots, x_n\right)\right)$$

е еднократно гладко изображение, дефинирано в отвореното множество $D \in \mathbb{R}^n$. Нека освен това

$$\frac{D\left(f_1, \dots, f_n\right)}{D\left(x_1, \dots, x_n\right)}(x) \neq 0 \quad \text{за всяко} \quad x \in D$$

(такова изображение ще наричане регулярно). Теоремата гласи, че всяко регулярно изображение локално е обратимо.

Теорема 7. (Теорема за обратното изображение). *Нека $f(x)$ е изображение, удовлетворяващо горните условия, и нека $x^0 \in D$ и*

---

* в сравнение с §4 тук са разменени местата на зависимите и независимите променливи

$y^0 = f(x^0) \in f(D)$. *Тогава съществуват околности* U *и* V *на точките* x^0 *и* y^0, *и еднократно гладко регулярно изображение* $g(y) : V \to U$, *обратно на изображението* $f(x)$ *(С други думи,* $f(g(y))$ *за всяко* $y \in V$.*)*

В частност, от теоремата следва, че образът $f(D)$ на изображението f е отворено множество в $\mathbb{R}^n$.

Доказателство. Да разгледаме графиката Γ_f на изображението $f(x)$, т.е. множеството от всички точки $(x, y) = (x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{R}^{2n}$, за които $y = f(x)$. Очевидно множеството Γ_f може да се зададе с уравненията

$$F_1(x_1, \ldots, x_n, y_1, \ldots, y_n) \stackrel{def}{=} f_1(x_1, \ldots, x_n) - y_1 = 0,$$

$$\ldots\ldots\ldots\ldots,$$

$$F_n(x_1, \ldots, x_n, y_1, \ldots, y_n) \stackrel{def}{=} f_n(x_1, \ldots, x_n) - y_n = 0.$$

Да вземем точката $(x^0, y^0) \in \Gamma_f$. Имаме

$$\frac{D(F_1, \ldots, F_n)}{D(x_1, \ldots, x_n)}\left(x^0, y^0\right) = \frac{D(f_1, \ldots, f_n)}{D(x_1, \ldots, x_n)}\left(x^0\right) \neq 0$$

и следователно около тази точка можем да приложим теоремата за неявната функция, считайки $y_1, \ldots, y_n$ за независими променливи и изразявайки чрез тях $x_1, \ldots, x_n$. С други думи, в някаква околност V на точката y^0 съществуват единствени функции $g_1(y_1, \ldots, y_n), \ldots, g_n(y_1, \ldots, y_n)$, и определено от тях изображение $g(y) : V \to \mathbb{R}^n$, удовлетворяващи условията

1/ $g(y^0) = x^0$, и

2/ $F_k(g(y), y) = 0$, $k = 1, \ldots, n$. Равенствата 2/ могат да се напишат във вида

$$f_k(g_1 y_1, \ldots, y_n, \ldots, g_n y_1, \ldots, y_n) = y_k, \quad k = 1, \ldots, n,$$

т.е. $f(g(y)) = y$. Да означим $U = g(V)$; тогава, тъй като V е отворено, множеството U съдържа всяка своя точка заедно с нейна околност, т.е. U е също отворено.

Накрая, регулярността на изображението $g(y)$ следва от основното свойство на функционалните детерминанти. Наистина, т.6 от §4 за производна на обратно изображение показва, че

$$\frac{D\,(g_1,\ldots,g_n)}{D\,(y_1,\ldots,y_n)} \neq 0. \quad \blacksquare$$

Забележка. Ако предварително се знае, че изображението f е взаимно еднозначно, то обратното му изображение е също такова и изобразява $f(D)$ върху D. Според доказаната теорема то е еднократно гладко и регулярно навсякъде.

В общия случай обаче локалната обратимост не влече след себе си глобална. Да разгледаме добре познатата полярна смяна на координатите. По-точно, нека $D = \{(\rho,\theta) : \rho > 0\} \subset \mathbb{R}^2$ и изображението от D в $\mathbb{R}^2 \setminus \{(0,0)\}$ е определено с формулите

$$x = \rho\cos\theta,\ y = \rho\sin\theta.$$

Лесно се смята, че

$$\frac{D(x,y)}{D(\rho,\theta)}(\rho,\theta) = \begin{vmatrix} x'_\rho & x'_\theta \\ y'_\rho & y'_\theta \end{vmatrix} = \begin{vmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{vmatrix} = \rho > 0$$

и следователно за изображението е приложима доказаната теорема, т.е. то е локално обратимо. Очевидно обаче то не е глобално обратимо, тъй като всяка точка от $\mathbb{R}^2 \setminus \{(0,0)\}$ има безкрайно много праобрази (със стойности на θ, различаващи се със целочислени кратни на 2π.)

1.9 Обвивки на фамилии от криви.

В този параграф ще изложим едно приложение на развитата по-горе техника, отнасящо се за фамилии от криви в равнината. В теорема 2 от предния параграф и в последвалата я забележка беше показано, че гладките криви в равнината могат да бъдат зададени с уравнения от вида

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

където $F(x, y)$ е еднократно гладка функция на x, y и $\overrightarrow{\mathrm{grad}}\, F(x, y) \neq \vec{0}$. Тук ще разглеждаме *фамилии* от еднократно гладки криви в равнината. С други думи, въвежда се параметър a, изменящ се в даден интервал, така че на всяка негова стойност съответства дадена крива Γ_a. По-точно, имаме

Определение. *Нека е дадена еднократно гладката функция на три променливи $F(x, y, a)$, като $(x, y) \in D \subset \mathbb{R}^2$, $a \in \Delta = (a_1, a_2)$, и нека навсякъде да е изпълнено равенството*

$$|F'_x| + \left|F'_y\right| > 0.$$

Тогава казваме, че функцията F поражда фамилията от криви $\{\Gamma_a\}_{a\in\Delta}$, като при всяка фиксирана стойност на а кривата $\Gamma_a \subset \mathbb{R}^2$ се състои от точките (x, y) в равнината, удовлетворяващи равенството

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

.

Ще въведем понятието обвивка на фамилия от криви. Първото понятие, което ще ни трябва, е понятието допиране на криви:

Определение. *Ще казваме, че две регулярни еднократни криви, минаващи през дадена точка, се допират в тази точка, ако допирателните им прави в тази точка съвпадат.*

Определение. *Регулярната и еднократно гладка крива Γ ще наричаме обвивка на фамилията $\{\Gamma_a\}_{a\in\Delta}$, ако през всяка точка от Γ преминава точно една крива от фамилията $\{\Gamma_a\}$, и в тази точка двете криви се допират.*

Ще изведем параметричните уравнения на обвивката на дадена фамилия от криви (ако такава обвивка съществува). Ще изберем за параметър индекса a на кривата Γ_a от фамилията, минаваща през съответната точка. По-точно, нека

$$x = x(a),\ y = y(a)\,, a \in \Delta$$

са параметричните уравнения на Γ. Ще искаме кривите Γ и Γ_a да минават през точката $(x(a), y(a))$ и да се допират в нея.

Първото условие ни дава, че $(x(a), y(a)) \in \Gamma_a$, т.е.

$$F\left(x(a), y(a), a\right) = 0.$$

Диференцирайки това равенство по a, получаваме

$$x'(a)F'_x + y'(a)F'_y + F'_a = 0,$$

като производните на F се вземат в точката $(x(a), y(a), a)$.

Да изразим и условието за допиране. Както знаем, допирателният вектор към Γ в тази точка е равен на $(x'(a), y'(a))$. От друга страна, доказаната в предния параграф т.2 от §5 показва, че допирателният вектор към Γ_a е ортогонален на вектора $\left(F'_x, F'_y\right)$ в съответната точка, и следователно това е вярно и за допирателния вектор към Γ. Получаваме, че равенството

$$x'(a)F'_x + y'(a)F'_y = 0$$

е необходимо и достатъчно за допирането на кривите Γ и Γ_a.

Изваждайки това равенство от полученото по-горе, получаваме. че $F'_a = 0$. В крайна сметка за параметричните функции на търсената обвивка получихме системата уравнения

$$F\left(x(a), y(a), a\right) = 0$$

$$F_a\left(x(a), y(a), a\right) = 0.$$

Системи от подобен вид бяха разгледани в предния параграф. Така стигаме до следното твърдение:

Теорема. *Нека функцията $F(x, y, a)$ е двукратно гладка и удовлетворява условията*

$$\begin{vmatrix} F'_x & F'_y \\ F''_{ax} & F''_{ay} \end{vmatrix} \neq 0, \quad F''_{aa} \neq 0.$$

Тогава породената от нея фамилия от криви в равнината притежава обвивка.

Доказателство. Като приложим теоремата за неявната функция към уравненията

$$F(x, y, a) = 0,\ F'_a(x, y, a) = 0$$

относно неизвестните x и y и имайки пред вид първото от горните условия, получаваме съществуването на двукратно гладки функции $x(a), y(a)$, удовлетворяващи горните уравнения. Като диференцираме второто от уравненията по a, получаваме

$$x' F''_{ax} + y' F''_{ay} = -F''_{aa} \neq 0,$$

т.е. $x'(a)$ и $y'(a)$ не се анулират едновременно.

Да означим с Γ параметрично зададената крива, определена с параметричното представяне $x = x(a)$, $y = y(a)$. Току-що видяхме, че това параметрично представяне е регулярно. Диференцирайки първото уравнение по a и изваждайки от него второто, получаваме равенството $x'(a)F'_x + y'(a)F'_y = 0$, от което, както видяхме, следва допирането на кривите Γ и Γ_a. ∎

Забележка. Условията на теоремата не са задължителни за съществуването на обвивка. В конкретните задачи обикновено те не се проверяват, а задачата се свежда до решаване на уравненията $F(x, y, a) = 0$, $F'_a(x, y, a) = 0$ относно неизвестните x и y.

Примери. 1. Нека разгледаме фамилията от окръжности с даден радиус r, чиито център се движи по абцисната ос. Ако означим с a координатата на центъра на окръжността, виждаме, че съответната окръжност се дава с уравнението

$$F(x, y, a) = (x - a)^2 + y^2 - r^2 = 0.$$

Диференцирайки по a, получаваме второто уравнение за обвивката:

$$F'_a(x,y,a) = -2\,(x-a) = 0.$$

Оттук виждаме, че обвивката на фамилията се състои от правите $y=r$ и $y=-r$.

2. Нека е дадена отсечка с дължина l, чиито краища се плъзгат съответно по абсцисната и ординатната ос. Да се определи вида на фигурата в равнината, запълнена от всички възможни положения на отсечката.

Да разгледаме случая, когато отсечката се намира в първи квадрант. Ако единият край на отсечката има координати $(a,0)$, то координатите на другия и край са $\left(0,\sqrt{l^2-a^2}\right)$. Припомняйки си отрезовото уравнение на права линия, получаваме уравнението на получената фамилия от прави:

$$F(x,y,a) = \frac{x}{a} + \frac{y}{\sqrt{l^2-a^2}} - 1 = 0, \quad a \in [0,l].$$

Имаме

$$F_a(x,y,a) = -\frac{x}{a^2} + \frac{ay}{\left(l^2-a^2\right)^{3/2}} = 0.$$

Решавайки уравненията относно x и y, получаваме $x = \frac{a^3}{l^2}$, $y = \frac{\left(l^2-a^2\right)^{3/2}}{l^2}$. Изключвайки a, получаваме уравнението на обвивката:

$$x^{2/3} + y^{2/3} = l^{2/3}.$$

В останалите три квадранта се получава същото уравнение, и ние получаваме частен случай на кривата, наречена астроида. В крайна сметка търсената фигура е оградена от получената крива и се състои от всички точки $(x,y) \in \mathbb{R}^2$, чиито координати удовлетворяват неравенството

$$|x|^{2/3} + |y|^{2/3} \leq l^{2/3}.$$

3. Парабола на безопасността. Нека си представим оръдие, разположено в началото на координатите и стрелящо с определена начална скорост, но под различни ъгли. Да се определи множеството от точките в равнината, които могат да бъдат достигнати от снарядите.

Ако снарядът е изстрелян с начална скорост v под ъгъл α към хоризонта, уравненията на движението му имат вида

$$x = t\,v\cos\alpha, \quad y = t\,v\sin\alpha - \frac{g}{2}\,t^2,$$

където g е земното ускорение. Изключвайки времето t, получаваме уравнението на траекторията на движението

$$y = -\frac{g}{2v^2\cos^2\alpha}\,x^2 + x\,\mathrm{tg}\,\alpha.$$

Да положим $a = \mathrm{tg}\,\alpha$; тогава при вариране на ъгъла α съответната фамилия от параболи се поражда от функцията

$$F(x,y,a) = -k\left(1+a^2\right)x^2 + ax - y = 0.$$

(за момента сме положили $k = \frac{g}{2v^2}$). Следващото уравнение от системата е

$$F_a(x,y,a) = x - 2kax^2 = 0,$$

откъдето получаваме параметричните уравнения на обвивката

$$x = \frac{1}{2ka}, \quad y = \frac{1}{4k}\left(\frac{1}{a^2} - 1\right).$$

Изключвайки параметъра a, получаваме обвивката на получената система от параболи, която се оказва също парабола (т.н. парабола на безопасността), с уравнение:

$$y = \frac{1}{4k} - k\,x^2 = \frac{v^2}{2g} - \frac{g}{2v^2}\,x^2.$$

В крайна сметка зоната на безопасността, т.е. множеството от онези точки (x,y) в равнината, които не могат да бъдат достигнати при никакъв ъгъл на изстрелване, се описва с неравенството:

$$y > \frac{v^2}{2g} - \frac{g}{2v^2}\,x^2.$$

4. От горните три примера читателят може да остане с впечатлението, че обвивката винаги огражда частта от равнината, запълнена от фамилията криви. Ще дадем пример, когато това не е така. Да вземем графиката на функцията $y = x^3$ и да я плъзгаме успоредно на абцисната ос. Получаваме фамилия от криви, определена с уравнението

$$F(x,y,a) = (x-a)^3 - y = 0.$$

Решавайки получената система, виждаме, че единствената обвивка на тази фамилия съвпада с абсцисната ос, докато кривите от фамилията запълват цялата равнина.

В този пример се вижда също, че обвивката съществува, макар че условията на теоремата не са удовлетворени (покажете!).

5. Всяка регулярна крива съвпада с обвивката на фамилията от собствените си тангенти. За улеснение ще докажем това в случая, когато кривата е явно зададена, т.е. е представена като графика на гладка функция. Нека Γ се задава с уравнението

$$y = f(x), \; x \in \Delta.$$

Ще предполагаме, че $f(x)$ е двукратно гладка и $f''(x) \neq 0$. Допирателната линия към Γ, минаваща през точката $(a, f(a))$, има уравнение $y = f(a) + f'(a)(x-a)$, и следователно фамилията от тангентите се задава чрез уравнението

$$F(x,y,a) = -y + f(a) + f'(a)(x-a) = 0.$$

Тъй като $F_a(x,y,a) = f''(a)(x-a)$, то за обвивката получаваме уравненията $x = a$, $y = f(a)$, т.е. тя съвпада с Γ.

6. Тук ще разгледаме обвивката на фамилията от нормалите на дадена крива, и ще докажем, че тя съвпада с еволютата на кривата.*

Понятието *еволюта* на дадена двукратно гладка крива с ненулева кривина беше въведено в том I, §4.4, и означаваше геометричното място на всички центрове на кривината на кривата. Нека двукратно гладката регулярна крива G е зададена с параметричните уравнения

$$x = \varphi(a), \; y = \psi(a), \; a \in \Delta.$$

*Този факт беше даден без обосновка при разглеждането на геометричния смисъл на еволютата, направено в I, §4.4.

Нормалата към кривата G в точката $(\varphi(a), \psi(a))$ е колинеарна с вектора $(-\psi'(a), \varphi'(a))$ и следователно се описва с параметричните уравнения

$$x = \varphi(a) - t\,\psi'(a),\ y = \psi(a) + t\,\varphi'(a),\ t \in \mathbb{R}.$$

Елиминирайки параметъра t, получаваме уравнението на нормалата във вида

$$F(x, y, a) = \psi'(a)\,(y - \psi(a)) + \varphi'(a)\,(x - \varphi(a)) = 0.$$

Оттук получаваме

$$F_a(x, y, a) = \psi''(a)\,(y - \psi(a)) - (\psi'(a))^2 + \varphi''(a)\,(x - \varphi(a)) - (\varphi'(a))^2\,.$$

Така уравненията на обвивката придобиват вида
$\psi'(a)\,(y - \psi(a)) + \varphi'(a)\,(x - \varphi(a)) = 0,$
$\psi''(a)\,(y - \psi(a)) + \varphi''(a)\,(x - \varphi(a)) = (\psi'(a))^2 + (\varphi'(a))^2\,.$

Ще напомним, че кривината $k(a)$ на кривата G в точката $(\varphi(a), \psi(a))$ се дава с формулата

$$k(a) \;=\; \frac{\varphi'(a)\psi''(a) - \varphi''(a)\psi'(a)}{(\varphi'(a)^2 + \psi'(a)^2)^{\frac{3}{2}}}\,.$$

Така, условието $k(a) \neq 0$ може да се напише във вида

$$\begin{vmatrix} \psi'(a) & \varphi'(a) \\ \psi''(a) & \varphi''(a) \end{vmatrix} \neq 0.$$

Следователно, ако разгледаме горните уравнения като линейни уравнения спрямо променливите $y - \psi(a)$, $x - \varphi(a)$, то те притежават единствено решение. Използвайки формулите на Крамер, получаваме решенията на системата:

$$x(a) = \varphi(a) - \frac{\varphi'(a)^2 + \psi'(a)^2}{\varphi'(a)\psi''(a) - \varphi''(a)\psi'(a)}\,\psi'(a),$$

$$y(a) = \psi(a) + \frac{\varphi'(a)^2 + \psi'(a)^2}{\varphi'(a)\psi''(a) - \varphi''(a)\psi'(a)}\,\varphi'(a).$$

Лесно е да се види, че получените формули съвпадат с параметричното представяне на еволютата на G, изведено в том I, §4.4 по съвсем различен начин.

Упражнения.

1. (задача, дадена за кандидат-студентски изпит за ФМИ на СУ през 1999 г.) Дадено е уравнението

$$3x^2 - (6a+1)x + a^2 + 6a - 3 = 0.$$

Да се намерят стойностите на a, за които даденото уравнение притежава корен u, така че стойността на $\left|u - \sqrt{2}\right|$ да е минимална.

Упътване. Разбира се, задачата притежава елегантно елементарно решение. За да решите задачата с помощта на теорията на обвивките, намерете обвивката на съответното семейство от параболи, и разгледайте пресечните и точки с абсцисната ос.

2. Дадена е елипсата

$$\frac{x^2}{p^2} + \frac{y^2}{q^2} = 1.$$

За всяка стойност на $a \in [-p, p]$ да прекараме вертикалната хорда на елипсата през точката $(a, 0)$, и да означим с Γ_a окръжността, имаща тази хорда за диаметър. Докажете, че обвивката на фамилията $\{\Gamma_a\}$ е елипса със същите оси на симетрия като дадената.

Забележка. От доказателството се вижда, че обвивката ще докосва само тези криви Γ_a, за които $|a| \leq \frac{p^2}{\sqrt{p^2+q^2}}$, т.е. обвивката съществува само за част от дадената фамилия.

3. (Характеристични точки). Нека е дадена фамилията $\{\Gamma_a\}$ от гладки криви. Нека фиксираме дадена стойност на a и за достатъчно малки стойности на нарастването h да означим с (x_h, y_h) някаква пресечна точка на кривите Γ_a и Γ_{a+h}. Да допуснем, че при $h \to 0$ пресечните точки (x_h, y_h) имат граница (x_0, y_0). Тогава точката (x_0, y_0) се нарича *характеристична точка* за фамилията $\{\Gamma_a\}$.

Докажете, че характеристичните точки лежат върху обвивката на фамилията.

Упътване. Очевидно координатите x_h, y_h на пресечната точка удовлетворяват уравненията

$$F(x_h, y_h, a) = 0, \quad F(x_h, y_h, a+h) = 0.$$

Използвайки теоремата за крайните нараствания, докажете, че тази система е еквивалентна на

$$F(x_h, y_h, a) = 0, \quad F'_a(x_h, y_h, a+\theta h) = 0, \quad \theta \in (0,1),$$

и направете граничен преход при $h \to 0$.

Забележка. От разгледания по-горе пример 4 се вижда, че не винаги точките от обвивката са характеристични.

1.10 Условни локални екстремуми. Множители на Лагранж.

В задачата за търсене на максимална и минимална стойности на функция на много променливи в област $D \subset \mathbb{R}^n$, разгледана в §7, остана открит следния проблем: да се намерят екстремалните стойности на функцията върху границата bD на областта D. Ако например D е кълбо, то възниква въпросът за търсене на екстремумите на функцията върху ограничаващата го сфера. На такава задача е посветен настоящият параграф.

Нека M е подмножество на $\mathbb{R}^n$ и $f(x)$ е функция, дефинирана върху M.

Определение. *Казваме, че точката $x^0 \in M$ е точка на <u>относителен</u> <u>локален</u> <u>максимум</u> на функцията $f(x)$ върху множеството M, ако съществува $\varepsilon > 0$, така че за за всяка точка $x \in M$, за която $||x - x^0|| < \varepsilon$, да имаме*

$$f(x) \leq f\left(x^0\right).$$

Както се вижда, това определение се различава от даденото в §5 определение на (безусловен) локален максимум по това, че се разглеждат само стойностите на $f(x)$ върху множеството M.

Ако заместим горното неравенство с противоположното му, получаваме дефиницията на <u>относителен</u> <u>локален</u> <u>минимум</u>. Най-сетне, двете дефиниции се обединяват в понятието <u>относителен</u> <u>локален</u> <u>екстремум</u>.

Ще разгледаме една елементарна геометрична задача: измежду всички правоъгълници с даден периметър да се намери този с максимално лице. Ако означим с x и y страните на правоъгълника, стигаме до следната формулировка:

Ако множеството $M \subset \mathbb{R}^2$ е множеството на всички точки $(x, y) \in \mathbb{R}^2$, удовлетворяващи условията $x > 0, y > 0, x + y = p$, да се намери максималната стойност върху M на функцията $f(x, y) = xy$.

Разбира се, задачата се решава елементарно, като се изрази y чрез x и след това се намери максимума на получената функция на една променлива. Ние обаче използваме задачата за илюстриране на развитите по-долу методи.

Множители на Лагранж - случай на едно уравнение в $\mathbb{R}^2$. За изясним идеята, в тази точка ще разгледаме случая, когато M е множеството от всички точки $(x, y) \in \mathbb{R}^2$, за които $F(x, y) = 0$. Тук $F(x, y)$ е еднократно гладка функция на 2 променливи, дефинирана в област в $\mathbb{R}^2$ и удовлетворяваща условието

$$\overrightarrow{\mathrm{grad}}\, F(x,y) \neq \vec{0},$$

т.е. първите и частни производни никъде не се анулират едновременно. Налице е следното необходимо условие за условен локален екстремум:

Теорема 1. *Нека еднократно гладката функция $f(x, y)$ достига условен локален екстремум върху M в точката $(x_0, y_0) \in M$. Тогава съществува константа λ такава, че*

$$\overrightarrow{\mathrm{grad}}\, f(x_0, y_0) = \lambda\, \overrightarrow{\mathrm{grad}}\, F(x_0, y_0).$$

Преди доказателството ще дадем геометрична интерпретация на твърдението на теоремата. Да си спомним теорема 2 от §5, чието доказателство дадохме в предния параграф. Теоремата гласи, че градиентът на функцията в дадена точка е перпендикулярен на линията на ниво на функцията, минаваща през същата точка. Според теорема 1, векторите $\overrightarrow{\mathrm{grad}}\, f(x_0, y_0)$ и $\overrightarrow{\mathrm{grad}}\, F(x_0, y_0)$ са колинеарни; това означава, че линиите на ниво на функциите $f(x, y)$ и $F(x, y)$, минаващи през точката (x_0, y_0), имат обща тангента, т.е. се допират помежду си. Очевидно линията на ниво на $F(x, y)$, минаваща през (x_0, y_0), съвпада с множеството M, описвано с уравнението $F(x, y) = 0$. Така стигаме до следната формулировка на теорема 1:

В точката (x_0, y_0) множеството M се допира до линията на ниво на $f(x, y)$, минаваща през тази точка.

На чертежа теоремата е илюстрирана за задачата за правоъгълниците, дадена по-горе; представени са множеството M (отсечка), и линиите на ниво на функцията $f(x, y) = xy$. Условният локален максимум се достига в точката $A = (p/2, p/2)$.

Можем да дадем и друга интерпретация на теоремата: да си представим, както в §5, планинска местност, като функцията $f(x,y)$ задава надморската височина в дадена точка, и множеството M - като пътека в тази местност. Тогава в най-високата си точка пътеката става хоризонтална, т.е. допира се до хоризонталата, минаваща през тази точка.

Доказателство на теорема 1. Трябва да докажем равенствата

$$\frac{\partial f}{\partial x}(x_0,y_0)=\lambda\,\frac{\partial F}{\partial x}(x_0,y_0)\,,\qquad \frac{\partial f}{\partial y}(x_0,y_0)=\lambda\,\frac{\partial F}{\partial y}(x_0,y_0)\,.$$

Ще използваме теоремата за неявната функция, за да параметризираме множеството M около точката (x_0,y_0). (Виж теорема 2 от предния параграф.) Тъй като $\overrightarrow{\mathrm{grad}}\,F\,(x_0,y_0)\neq\vec{0}$, можем да предположим например, че $\frac{\partial F}{\partial y}(x_0,y_0)\neq 0$. Тогава в някаква околност на точката x_0 съществува еднократно гладка функция $\varphi\,(x)$, удовлетворяваща условията

$$\varphi\,(x_0)=y_0,\quad F\,(x,\varphi(x))\equiv 0.$$

Диференцирайки последното равенство по x в точката x_0, получаваме

$$\frac{\partial F}{\partial x}(x_0,y_0)+\varphi'\,(x_0)\ \frac{\partial F}{\partial y}(x_0,y_0)=0.$$

От друга страна, заествайки y с $\varphi(x)$, получаваме функцията на едно променливо $g(x)=f\,(x,\varphi(x))$, определена в околност на x_0. Точките от вида $(x,\varphi(x))$ принадлежат на M; следователно, ако $f(x,y)$ притежава условен локален екстремум в точката (x_0,y_0), то $g(x)$ притежава локален екстремум от същия вид в точката x_0. Оттук следва, че производната и в тази точка се анулира:

$$g'\,(x_0)=\frac{\partial f}{\partial x}(x_0,y_0)+\varphi'\,(x_0)\ \frac{\partial f}{\partial y}(x_0,y_0)=0.$$

Да изразим $\varphi'\,(x_0)$ от предишното равенство и да го заместим тук; получаваме

$$\frac{\partial f}{\partial x}(x_0,y_0)-\frac{\partial F}{\partial x}(x_0,y_0)\left(\frac{\frac{\partial f}{\partial y}(x_0,y_0)}{\frac{\partial F}{\partial y}(x_0,y_0)}\right)=0.$$

Полагайки $\lambda = \dfrac{\frac{\partial f}{\partial y}\left(x_0, y_0\right)}{\frac{\partial F}{\partial y}\left(x_0, y_0\right)}$, виждаме, че и двете искани равенства са изпълнени. ■

Забележка. Лесно се вижда, че необходимото условие за локален екстремум, дадено в теоремата, не е достатъчно. Наистина, нека $F(x, y) = y$ (множеството M съвпада с абсцисната ос) и $f(x, y) = x^3 + y$. Тогава в точката $(0, 0)$ градиентите на функциите f и F са колинеарни, но в тази точка нямаме условен локален екстремум.

Множители на Лагранж - общ случай. Нека сега M е подмножество на $\mathbb{R}^n$, зададено с условията

$$F_1\left(x_1, \ldots, x_n\right) = F_2\left(x_1, \ldots, x_n\right) = \ldots = F_k\left(x_1, \ldots, x_n\right) = 0,$$

където $F_1\left(x_1, \ldots, x_n\right), \ldots, F_k\left(x_1, \ldots, x_n\right)$ е регулярна система от еднократно гладки функции, дефинирани в област в $\mathbb{R}^n$. Ще напомним, че условието за регулярност означава, че техните градиенти $\overrightarrow{\operatorname{grad}} F_1(x), \ldots, \overrightarrow{\operatorname{grad}} F_k(x)$ са линейно независими за всяко x от дефиниционната им област.*

Теорема 2. *Нека еднократно гладката функция $f\left(x_1, \ldots, x_n\right)$ достига условен локален екстремум върху M в точката $x^0 = (x_1^0, \ldots, x_n^0) \in M$. Тогава съществуват константи $\lambda_1, \ldots, \lambda_k$ такива, че*

$$\overrightarrow{\operatorname{grad}} f\left(x^0\right) = \sum_{i=1}^{k} \lambda_i \overrightarrow{\operatorname{grad}} F_i\left(x^0\right).$$

Геометрична интерпретация. В теорема 6 от предния параграф беше показано, че векторите $\overrightarrow{\operatorname{grad}} F_1(x^0), \ldots, \overrightarrow{\operatorname{grad}} F_k(x^0)$ образуват базис в ортогоналното допълнение на допирателното пространство към M в точката x^0. Така твърдението на теоремата може да се формулира по следния начин:

Векторът $\overrightarrow{\operatorname{grad}} f\left(x^0\right)$ е перпендикулярен към допирателното пространство към M в точката x^0.

*От условието за регулярност се вижда, че броят k на функциите не може да надвишава размерността n на пространството. Задачата е смислена само при $k < n$, което ще смятаме за изпълнено.

Казано по-свободно, линията на ниво на функцията f, минаваща през точката x^0, се допира към множеството M в тази точка.

Доказателство на теорема 2. Сменяйки, ако е нужно, номерата на променливите, можем да считаме, че

$$\frac{D\left(F_1, \ldots, F_k\right)}{D\left(x_1, x_2, \ldots, x_k\right)}\left(x^0\right) \neq 0.$$

Разсъждавайки както в теорема 5 от предния параграф, можем, в околност на точката x^0, можем да представим множеството M чрез равенствата

$$x_1 = \varphi_1\left(x_{k+1}, \ldots, x_n\right)$$

$$\ldots\ldots\ldots\ldots$$

$$x_k = \varphi_k\left(x_{k+1}, \ldots, x_n\right).$$

Вземайки пред вид теорема 6 от предния параграф, виждаме, че доказателството ще бъде извършено, ако докажем, че $\overrightarrow{\operatorname{grad}} f\left(x^0\right)$ е ортогонален на векторите $\vec{l_1}, \ldots, \vec{l}_{n-k}$, определени с формулите

$$\vec{l_p} = \left(\frac{\partial \varphi_1}{\partial x_{k+p}}\left(x^0\right), \ldots, \frac{\partial \varphi_k}{\partial x_{k+p}}\left(x^0\right), 0, \ldots, 0, 1, 0, \ldots, 0\right),$$

като единицата е разположена на p-то място измежду последните $n-k$ координати на вектора $\vec{l_p}$.

Да разгледаме функцията

$$g\left(x_{k+1}, \ldots, x_n\right) = f\left(\varphi_1\left(x_{k+1}, \ldots, x_n\right), \ldots, \varphi_k\left(x_{k+1}, \ldots, x_n\right), x_{k+1}, \ldots, x_n\right).$$

Тъй като $f(x)$ достига условен локален екстремум в точката x^0, то функцията $g\left(x_{k+1}, \ldots, x_n\right)$ достига локален екстремум от същия вид в точката $\left(x_{k+1}^0, \ldots, x_n^0\right)$. Следователно, в тази точка са изпълнени равенствата

$$\frac{\partial g}{\partial x_{k+1}} = \ldots = \frac{\partial g}{\partial x_n} = 0.$$

По теоремата за диференциране на съставни функции при $p = 1, \ldots, n-k$ имаме

$$\frac{\partial g}{\partial x_{k+p}}\left(x_{k+1}^0, \ldots, x_n^0\right) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}\left(x^0\right) \cdot \frac{\partial \varphi_i}{\partial x_{k+p}}\left(x_{k+1}^0, \ldots, x_n^0\right) + \frac{\partial f}{\partial x_{k+p}}\left(x^0\right),$$

или, с други думи, при $p = 1, \dots, n-k$ имаме

$$\left\langle \overrightarrow{\mathrm{grad}} f\left(x^0\right), \vec{l_p} \right\rangle = 0,$$

което доказва твърдението на теоремата. ■

От твърдението на теоремата лесно се извежда рецепта за откриване на точките, в които може да се очаква относителен локален екстремум. "Подозрителната"точка x трябва да удовлетворява уравненията

$$\frac{\partial f}{\partial x_j}(x) = \sum_{i=1}^{k} \lambda_i \frac{\partial F_i}{\partial x_j}(x), \quad j = 1, \dots, n.$$

Като добавим и уравненията за връзка $F_1(x) = \ldots = F_k(x) = 0$, получаваме $n+k$ уравнения за $n+k$ неизвестни $x_1, \dots, x_n, \lambda_1, \dots, \lambda_k$. По-нататък целта е да се елиминират променливите $\lambda_1, \dots, \lambda_k$ и да се намерят стойностите на $x_1, \dots, x_n$. По-нататък трябва да се провери дали така получената точка е наистина точка на екстремум, и от какъв вид е той; обикновено това се вижда непосредствено.

Необходимото условие за условен локален екстремум, което току-що доказахме, засяга първите производни на изследваната функция и е аналогично на необходимото условие за обикновен екстремум - анулиране на всички първи производни в екстремалната точка (т. 1 от §7). Сега ще докажем и достатъчно условие за условен локален екстремум, използващо вторите производни на функцията, т.е. аналог на теорема 3 от същия параграф. Трябва да отбележим обаче, че при решаване на конкретни задачи това условия рядко се използва.

Разбира се, достатъчните условия, които търсим, трябва да включват и доказаните по-горе необходими условия. Нека $f, F_1, \dots, F_k$ са както в теорема 2. Ще казваме, че точката x^0 е критична точка на функцията $f(x)$ върху множеството M, ако съществуват константи $\lambda_1, \dots, \lambda_k$ такива, че

$$\overrightarrow{\mathrm{grad}} f\left(x^0\right) = \sum_{i=1}^{k} \lambda_i \overrightarrow{\mathrm{grad}} F_i\left(x^0\right).$$

Теорема 3. (Достатъчно условие за условен локален екстремум.) *Нека $F_1(x), \dots, F_k(x)$ е регулярна система от двукратно гладки функции, определени в област в $\mathbb{R}^n$, $M \subset \mathbb{R}^n$ е множеството от техните общи нули, и $f(x)$ е двукратно гладка функция, определена в околност на M. Да предположим, че x^0 е критична точка на функцията $f(x)$ върху множеството M, и нека $\lambda_1, \dots, \lambda_k$ са съответните константи. Да определим функцията $\Phi(x)$ с формулата:*

$$\Phi(x) = f(x) - \sum_{i=1}^{k} \lambda_i F_i(x).$$

Тогава, ако ограничението върху подпространството $T_{x^0}M$ на квадратичната форма

$$\mathcal{D}\Phi\left(x^0\right)\vec{h} \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2\Phi}{\partial x_i\,\partial x_j}\left(x^0\right)\; h_i h_j$$

е положително (отрицателно) определена квадратична форма, то x^0 е точка на локален екстремум на $f(x)$ върху M, и този екстремум е минимум (максимум).

Доказателство. Най-напред ще отбележим,че върху M функциите $f(x)$ и $\Phi(x)$ съвпадат, и е достатъчно да се докаже, че x^0 е локален условен екстремум за $\Phi(x)$. От дефиницията на функцията $\Phi(x)$ следва равенството $\overrightarrow{\mathrm{grad}}\,\Phi\left(x^0\right)=\vec{0}$, т.е. всички първи частни производни на $\Phi(x)$ се анулират в точката x^0.*

В предишния параграф беше отбелязано, че допирателното протранство $T_{x^0}M$ не зависи от избора на ортогонална координатна система в $\mathbb{R}^n$. Разбира се, същото е вярно и за понятието положително (или отрицателно) определена квадратична форма. Това ни позволява да опростим доказателството на теоремата, избирайки подходяща координатна система. Като начало можем да смятаме, че точката x^0 съвпада с началото 0 на координатната система.

По-нататък, можем да изберем координатите така, че $n-k$-мерното подпространство $T_{x^0}M$ на $\mathbb{R}^n$ съвпада с координатното подпространство, съответстващо на координатите $x_{k+1},\ldots,x_n$, т.е. се представя с равенствата $x_1=\ldots=x_k=0$. Да означим $x'=(x_{k+1},\ldots,x_n)\in\mathbb{R}^{n-k}$. Тогава условието за положителна или отрицателна определеност означава, че квадратичната форма

$$\mathcal{D}\Phi\left(x'\right)=\sum_{i=k+1}^{n}\sum_{j=k+1}^{n}\frac{\partial^2\Phi}{\partial x_i\,\partial x_j}(0)\,x_i x_j$$

е положително (отрицателно) определена в $\mathbb{R}^{n-k}$.

При такъв избор на координати променливите $x_{k+1},\ldots,x_n$ могат да бъдат използвани за локално параметризиране на множеството M около точката 0 (докажете!). Както по-горе, локално M се представя с равенствата

$$x_1=\varphi_1\left(x'\right),\ldots,x_k=\varphi_k\left(x'\right),\quad \text{като}\quad \varphi_1(0)=\ldots=\varphi_k(0)=0.$$

Знаем, че използваните в доказателството на теорема 2 вектори $\vec{l_1},\ldots,\vec{l_{n-k}}$ пораждат допирателното пространство в точката 0. От избора на координатната система следва,че първите k координати на тези вектори се анулират, т.е. изпълнени са равенствата

$$\frac{\partial\varphi_i}{\partial x_{k+p}}(0)=0\quad \text{за}\quad i=1,\ldots,k,\, p=1,\ldots,n-k.$$

*Именно поради това в теоремата се разглежда $\Phi(x)$ вместо $f(x)$.

От формулата на Тейлор следват равенствата

$$\varphi_i\left(x'\right) = o\left(||x'||^2\right), \quad i = 1, \ldots, k, p = 1, \ldots, n-k.$$

Да въведем функцията

$$\Psi\left(x_{k+1}, \ldots, x_n\right) = \Phi\left(\varphi_1\left(x_{k+1}, \ldots, x_n\right), \ldots, \varphi_k\left(x_{k+1}, \ldots, x_n\right), x_{k+1}, \ldots, x_n\right).$$

Формулата на Тейлор от втори ред за функцията Φ ни дава представянето

$$\Delta\Phi = \Phi(x) - \Phi\left(0\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2\Phi}{\partial x_i\, \partial x_j}\left(0\right) x_i x_j + o\left(||x||^2\right).$$

За да получим оттук формула за нарастването на функцията Ψ, трябва да заместим $x_1, \ldots, x_k$ с $\varphi_1\left(x'\right), \ldots, \varphi_k\left(x'\right)$. Получаваме

$$\Delta\Psi = \sum_{i=k+1}^{n}\sum_{j=k+1}^{n} \frac{\partial^2\Phi}{\partial x_i\, \partial x_j}\left(0\right) x_i x_j + 2\sum_{i=k+1}^{n}\sum_{j=1}^{k} \frac{\partial^2\Phi}{\partial x_i\, \partial x_j}\left(0\right) x_i \varphi_j\left(x'\right) +$$

$$+\sum_{i=1}^{k}\sum_{j=1}^{k} \frac{\partial^2\Phi}{\partial x_i\, \partial x_j}\left(0\right) \varphi_i\left(x'\right)\varphi_j\left(x'\right) + o\left(||x||^2\right).$$

Лесно се доказва, че изразите от вида $x_i\varphi_j\left(x'\right)$, $\varphi_i\left(x'\right)\varphi_j\left(x'\right)$, $o\left(||x||^2\right)$ клонят към нула по-бързо от $||x'||^2$, т.е. могат да бъдат записани като $o\left(||x'||^2\right)$. Така горното представяне добива вида

$$\Delta\Psi = \sum_{i=k+1}^{n}\sum_{j=k+1}^{n} \frac{\partial^2\Phi}{\partial x_i\, \partial x_j}\left(0\right) x_i x_j + o\left(||x'||^2\right).$$

След като имаме това представяне за нарастването на Ψ, можем да използваме доказателството на теорема 3 от §7 , което показва, че $\Psi\left(x'\right)$ има локален екстремум в точката $0 \in \mathbb{R}^{n-k}$, и следователно $f(x)$ има локален екстремум в точката $0 \in \mathbb{R}^n$. ■

Упражнения.

1. Намерете максималната стойност на функцията $f\left(x_1, \ldots, x_n\right) = x_1 x_2 \ldots x_n$ върху множеството M, зададено с условията $x_1 \geq 0, \ldots, x_n \geq 0, x_1 + \ldots + x_n = 1$.

2. Центрираното уравнение на елипса в равнината има вида $ax^2 + 2bxy + cy^2 = 1$, като квадратичната форма от лявата страна е положително определена. Намерете голямата и малката полуос на елипсата.

Решение. Да означим елипсата с M, и нека $f(x,y) = x^2 + y^2$ (квадрата на разстоянието от точката (x,y) до началото на координатите). Тогава задачата се свежда към това, да се намерят максималната и минималната стойности на $\sqrt{f(x,y)}$ върху M. Уравненията на критичните точки добиват вида

$$2x = \lambda\,(2ax + 2by)$$

$$2y = \lambda\,(2bx + 2cy)$$

От линейната алгебра знаем, че получената система хомогенни уравнения има ненулево решение точно тогава, когато детерминантата от коефициентите пред неизвестните е равна на нула. Така за λ получаваме квадратното уравнение

$$\begin{vmatrix} a - \frac{1}{\lambda} & b \\ b & c - \frac{1}{\lambda} \end{vmatrix} = 0.$$

Нека μ_1, μ_2, $0 < \mu_1 < \mu_2$ да са собствените стойности на матрицата от коефициентите на квадратичната форма. Тогава решенията на горното уравнение са $\frac{1}{\mu_1}$, $\frac{1}{\mu_2}$. От друга страна, умножавайки първото уравнение по x, второто по y и събирайки, получаваме

$$x^2 + y^2 = \lambda\left(ax^2 + 2bxy + cy^2\right) = \lambda.$$

Така, малката и голямата полуос на елипсата (минималната и максимална стойности на разстоянието до началото) са равни на $\frac{1}{\sqrt{\mu_2}}$ и $\frac{1}{\sqrt{\mu_1}}$ съответно.

Забележка. Ние достигнахме по аналитичен път до процедурата, известна в алгебрата и в аналитичната геометрия като канонизиране на квадратична форма.

3. Нека елипсоидът M в $\mathbb{R}^n$ е зададен с уравнението

$$\sum_{i,j=1}^{n} a_{ij} x_i x_j = 1.$$

Докажете, че критичните точки на функцията $f(x) = ||x||^2$ върху M съвпадат с върховете на полуосите на елипсоида.

4. Нека е даден елипсоид в $\mathbb{R}^3$, зададен с каноничното си уравнение. Да се намерят полуосите на елипсата, получена като сечение на елипсоида с равнина в $\mathbb{R}^3$, минаваща през началото на координатите.

Упътване. Задачата се свежда към намирането на максималната и минимална стойност на функцията $f(x,y,z) = x^2 + y^2 + z^2$ върху множеството M, определено с уравненията

$$F_1(x,y,x) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0,$$

$$F_2(x,y,x) = \alpha x + \beta y + \gamma z = 0.$$

Приравнявайки на нула частните производни на функцията
$f - \lambda F_1 - \mu F_2$, получаваме равенствата

$$\lambda = f(x,y,z), \quad x = \mu\frac{\alpha a^2}{a^2 - \lambda}, \quad y = \mu\frac{\beta b^2}{b^2 - \lambda}, \quad z = \mu\frac{\gamma c^2}{c^2 - \lambda}.$$

Умножавайки второто, третото и четвъртото равенство съответно с α, β, γ и събирайки резултатите, получаваме за λ квадратното уравнение

$$\frac{\alpha^2 a^2}{a^2 - \lambda} + \frac{\beta^2 b^2}{b^2 - \lambda} + \frac{\gamma^2 c^2}{c^2 - \lambda} = 0.$$

Ако λ_1 и λ_2 са корените на това уравнение, то $\sqrt{\lambda_1}$ и $\sqrt{\lambda_2}$ са дължините на полуосите на търсената елипса.

1.11 Функционална независимост. Ранг на система от функции.

Функционална независимост на система от функции. Нека в областта $D \in \mathbb{R}^n$ са дефинирани еднократно гладките функции $f_1(x), \dots, f_k(x)$, където $k \leq n$, и $x = (x_1, \dots, x_n) \in D$.

Определение. *Ще казваме, че функциите $f_1, \dots, f_k$ са функционално зависими в областта D, ако съществува номер i , $1 \leq i \leq k$, и еднократно гладка функция g на $k-1$ променливи, така че за всяко $x \in D$ е изпълнено тъждеството*

$$f_i(x) = g\left(f_1(x), \dots, f_{i-1}(x), f_{i+1}(x), \dots, f_k(x)\right).$$

С други думи, поне една от функциите $f_1, \dots, f_k$ се изразява чрез останалите.

Ще казваме, че функциите $f_1, \dots, f_k$ са функционално независими в областта D, ако зависимост от горния вид не съществува.

Пример. Да разгледаме дефинираните в $\mathbb{R}^3$ функции

$$\begin{aligned} f_1(x, y, z) &= x + y + z, \\ f_2(x, y, z) &= x^2 + y^2 + z^2, \\ f_3(x, y, z) &= xy + yz + zx. \end{aligned}$$

Тогава очевидно тези функции удовлетворяват съотношението

$$f_2(x, y, z) = (f_2(x, y, z))^2 - 2f_3(x, y, z)$$

и следователно са функционално зависими.

Оказва се, че функционалната независимост на система от функции е свързана с ранга на матричната производна на изображението $f = (f_1, \dots, f_k)$:

$$Df\,(x) = \left\{\frac{\partial f_i}{\partial x_j}\,(x)\right\}_{\substack{i=1,\dots,k \\ j=1,\dots,n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j 1}\,(x) & \frac{\partial f_1}{\partial x_2}\,(x) & \dots & \frac{\partial f_1}{\partial x_n}\,(x) \\ \frac{\partial f_2}{\partial x_1}\,(x) & \frac{\partial f_2}{\partial x_2}\,(x) & \dots & \frac{\partial f_2}{\partial x_n}\,(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_k}{\partial x_1}\,(x) & \frac{\partial f_k}{\partial x_2}\,(x) & \dots & \frac{\partial f_k}{\partial x_n}\,(x) \end{pmatrix}$$

Теорема 1. *Ако матрицата $Df(x)$ е с максимален ранг (равен на k) в областта D, то функциите $f_1, \ldots, f_k$ са функционално независими.*

Доказателство. Да предположим противното: че една от функциите, например f_k, се изразява чрез останалите; с други думи, съществува функция $g(y_1, \ldots, y_{k-1})$, за която е изпълнено тъждеството

$$f_k(x) = g\left(f_1(x), \ldots, f_{k-1}(x)\right).$$

Чрез диференциране получаваме при $i = 1, \ldots, n$ равенствата

$$\frac{\partial f_k}{\partial x_i} = \sum_{j=1}^{k-1} \frac{\partial g}{\partial y_j} \frac{\partial f_j}{\partial x_i}.$$

С други думи, k-тият ред на матрицата $Df(x)$ се изразява като линейна комбинация на първите $k-1$ реда, и следователно тази матрица не може да има максимален ранг. ■

Ранг на система от функции. Нека, както по-горе, имаме еднократно гладките функции $f_1, \ldots, f_k$, и нека $f(x) = (f_1(x), \ldots, f_k(x))$ е породеното от тях изображение от D в $\mathbb{R}^k$. Имайки пред вид теорема 1, можем да си зададем въпроса: колко е най-малкият брой функции измежду $f_1, \ldots, f_n$, така че всички останали да се изразяват чрез тях. Оказва се, че този въпрос е тясно свързан с ранга на разгледаната по-горе матрица $Df(x)$.

Определение. *Казваме, че изображението $f(x)$ има ранг p в точката x, ако рангът на матрицата $Df(x)$ е равен на p.*

Ще казваме, че изображението $f(x)$ има ранг p в областта D, ако $\operatorname{rang} Df(x) \leq p$ за всяко $x \in D$, и ако съществува поне едно $x^0 \in D$, така че $\operatorname{rang} Df(x^0) = p$ (в такъв случай ще казваме, че рангът на $f(x)$ се достига в точката x^0).

Очевидно последното условие може да бъде формулирано така: съществуват индекси $1 \leq i_1 < \ldots < i_p \leq n$, $1 \leq j_1 < \ldots < j_p \leq k$, така че

$$\frac{D\left(f_{j_1}, \ldots, f_{j_p}\right)}{D\left(x_{i_1}, \ldots, x_{i_p}\right)}\left(x^0\right) \neq 0.$$

Теорема 2. *Нека изображението $f(x)$ има ранг p, $p < k$, в областта D, и нека този ранг се достига в точката x^0. Да предположим, че номерацията на функциите и независимите променливи е така избрана, че*

$$\frac{D\left(f_1,\ldots,f_p\right)}{D\left(x_1,\ldots,x_p\right)}\left(x^0\right) \neq 0.$$

Тогава съществуват $k-p$ на брой еднократно гладки функции

$$g_1\left(y_1,\ldots,y_p\right),\ldots,g_{k-p}\left(y_1,\ldots,y_p\right)$$

такива, че в някаква околност на x^0 са изпълнени тъждествата

$$f_{p+1}(x) = g_1\left(f_1(x),\ldots,f_p(x)\right),$$

$$\ldots\ldots\ldots,$$

$$f_k(x) = g_{k-p}\left(f_1(x),\ldots,f_p(x)\right).$$

Доказателство. Нека U е околност на x^0 в $\mathbb{R}^n$, в която продължава да бъде изпълнено равенството $\frac{D(f_1,\ldots,f_p)}{D(x_1,\ldots,x_p)}\left(x^0\right) \neq 0$. Доказателството се извършва с помощта на подходяща смяна на променливите в U. Нека изображението

$$u = \varphi(x) = \left(\varphi_1\left(x_1,\ldots,x_n\right),\ldots,\varphi_n\left(x_1,\ldots,x_n\right)\right)$$

е определено с формулите

$u_1 = \varphi_1\left(x\right) = f_1\left(x_1,\ldots,x_n\right),$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$
$u_p = \varphi_p\left(x\right) = f_p\left(x_1,\ldots,x_n\right),$
$u_{p+1} = \varphi_{p+1}\left(x\right) = x_{p+1},$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$
$u_n = \varphi_n\left(x\right) = x_n.$

Лесно се вижда равенството

$$\frac{D\left(\varphi_1,\ldots,\varphi_n\right)}{D\left(x_1,\ldots,x_n\right)}(x) = \frac{D\left(f_1,\ldots,f_p\right)}{D\left(x_1,\ldots,x_p\right)}(x) \neq 0.$$

Следователно, според теоремата за обратното изображение (т. 7 от §8) изображението $\varphi(x)$ притежава обратно изображение $\psi(u)$:

$$x = \psi(u) = \left(\psi_1\left(u_1, \dots, u_n\right), \dots, \psi_n\left(u_1, \dots, u_n\right)\right).$$

като

$$\frac{D\left(\psi_1, \dots, \psi_n\right)}{D\left(u_1, \dots, u_n\right)} = \left(\frac{D\left(\varphi_1, \dots, \varphi_n\right)}{D\left(x_1, \dots, x_n\right)}\right)^{-1} \neq 0.$$

Разглеждайки първите p координати на равенството $\varphi(\psi(u)) = u$, получаваме

$$f_i(\psi(u)) = u_i \qquad i = 1, \dots, p.$$

Да разгледаме изображението

$$g(u) = f\left(\psi(u)\right) : \mathbb{R}^n \to \mathbb{R}^k,$$

определено като суперпозиция на изображенията ψ и f. Според матричната форма на теоремата за диференциране на съставно изображение (т. 4 от §8) имаме

$$Dg(u) = Df(\psi(u)) \circ D\psi(u).$$

Тъй като квадратната матрица $D\psi(u)$ е обратима, то матриците $Dg(u)$ и $Df(\psi(u))$ имат еднакъв ранг, равен на p.

От друга страна, както току-що показахме, налице са равенствата

$$g_1\left(u_1, \dots, u_n\right) = u_1, \dots, g_p\left(u_1, \dots, u_n\right) = u_p.$$

Следователно, матричната производна на изображението g има вида

$$Dg(u) = \left(\begin{array}{ccc|ccc} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline \frac{\partial g_{p+1}}{\partial u_1} & \cdots & \frac{\partial g_{p+1}}{\partial u_p} & \frac{\partial g_{p+1}}{\partial u_{p+1}} & \cdots & \frac{\partial g_{p+1}}{\partial u_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial u_1} & \cdots & \frac{\partial g_k}{\partial u_1} & \frac{\partial g_k}{\partial u_{p+1}} & \cdots & \frac{\partial g_k}{\partial u_n} \end{array}\right)$$

Сега ще използваме факта, че матрицата $Dg(u)$ има ранг p и следователно всяка нейна поддетерминанта от ред $p+1$ се анулира. Да вземем произволни номера i и j такива, че $1 \leq i \leq n-p$, $1 \leq j \leq k-p$. Да изберем редовете с номера $1, 2, \ldots, p, p+i$ и стълбовете с номера $1, 2, \ldots, p, p+j$. Получената поддетерминанта от ред $p+1$ има вида

$$\frac{D\left(g_1, \ldots, g_p, g_{p+j}\right)}{D\left(u_1, \ldots, u_p, u_{p+i}\right)} = \left|\begin{array}{ccc|c} 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ \hline \frac{\partial g_{p+j}}{\partial u_1} & \cdots & \frac{\partial g_{p+j}}{\partial u_p} & \frac{\partial g_{p+j}}{\partial u_{p+i}} \end{array}\right| = \frac{\partial g_{p+j}}{\partial u_{p+i}}.$$

Така ние получихме равенствата

$$\frac{\partial g_{p+j}}{\partial u_{p+i}} = 0 \quad \text{за} \quad i, j = 1, 2, \ldots, n-p.$$

С други думи, функциите $g_{p+1}, \ldots, g_k$ зависят само от променливите $u_1, \ldots, u_p$.

Да заместим обратно променливите u с $\varphi(x)$. Имаме

$$g(\varphi(x)) = f(\psi(\varphi(x))) = f(x).$$

Вземайки в горното равенство координатните функции с номера $p+1, p+2, \ldots, k$, получаваме равенствата

$f_{p+1}(x) = g_{p+1}\left(f_1(x), \ldots, f_p(x)\right),$

$\ldots\ldots\ldots\ldots\ldots,$

$f_k(x) = g_k\left(f_1(x), \ldots, f_p(x)\right),$

които са тъждествено изпълнени за всяко $x \in U$. Сменяйки означенията и написвайки g_1 вместо g_{p+1}, g_2 вместо g_{p+2}, ..., g_{k-p} вместо g_k, получаваме твърдението на теоремата. ■

Геометричен смисъл на теоремата. Нека U е околността на x^0, в която е изпълнено твърдението на теорема 2. Тогава теоремата дава представа за геометрическата структура в $\mathbb{R}^k$ на образа $f(U)$ на U чрез изображението f. По-точно, нека V е образът в $\mathbb{R}^p$ на множеството U чрез функциите $f_1(x), \ldots, f_p(x)$. Поради регулярността на тази система от функции от т. 7 на §8 се вижда, че V е отворено в $\mathbb{R}^p$.

Тогава теорема 2 показва, че множеството $f(U)$ може да се определи в $\mathbb{R}^k$ с равенствата

$$y_{p+1} = g_1\left(y_1, \ldots, y_p\right), \ldots, y_k = g_{k-p}\left(y_1, \ldots, y_p\right),\ \left(y_1, \ldots, y_p\right) \in V.$$

С други думи, множеството $f(U)$ съвпада с графика на изображението $g = (g_1, \ldots, g_{k-p}) : V \to \mathbb{R}^{k-p}$.

Ако предположим, че навсякъде в дефиниционната си област D изображението f има ранг p, то образът му $f(D)$ притежава подобно представяне около всяка своя точка (като при това изборът на p независими променливи измежду променливите $y_1, \ldots, y_k$ може да бъде различен). Както ще видим по-нататък, подмножества с подобно локално представяне отговарят на интуитивната представа за "гладкост"и се наричат *p-мерни подмногообразия на* $\mathbb{R}^k$. Така стигаме до следната геометрична формулировка на теорема 2:

Теорема $2'$. *Нека еднократно гладкото изображение* $f(x) : D \to \mathbb{R}^k$ *има ранг* p, $p < k$, *навсякъде в областта* $D \subset \mathbb{R}^n$. *Тогава образът му* $f(D)$ *представлява p-мерно подмногообразие на* $\mathbb{R}^k$.

Забележка. Изискването за постоянство на ранга на изображението f, поставено в теоремата, е доста ограничително. В общия случай за ранга знаем само, че е полунепрекъсната отгоре функция на точката x (виж зад. 1). Образът на изображение с променлив ранг може да има сложна структура. Като прост пример можем да разгледаме изображението от $\mathbb{R}$ в $\mathbb{R}^2$, зададено с формулите $x = t^3$, $y = t^2$. Рангът на това изображение е равен на единица при $t \neq 0$ и на нула при $t = 0$. Образът на изображението е крива в $\mathbb{R}^2$, която има особеност от вида "рогова точка"в началото на координатите.

При пространства от по-висока размерност структурата на образа се усложнява значително и не се поддава на систематично описание.

Упражнения.

1. Докажете, че рангът на матрицата $Df\,(x)$ е полунепрекъсната отгоре функция на x (с други думи, ако редицата $\{x^n\}$ клони към точката x^0, то $\operatorname{rang} Df\,(x^0) \leq \liminf \operatorname{rang} Df\,(x^n)$). Дайте пример, в който тази функция не е непрекъсната.

Предметен указател

n-мерно евклидово пространство, 1
допиране на криви, 93
допирателен вектор, 41
допирателна равнина, 41
допирателно подпространство, 41
евклидова норма, 4
евклидово пространство, 1
евклидово разстояние, 3
евклидово разстояние в $\mathbb{R}^n$, 3
фамилия от криви, 93
функционална зависимост, 112
обвивка на фамилия от криви, 93
относителен локален екстремум, 102
произведение, 2
произведение на множества, 2
равнина, 2
регулярна система от функции, 88
регулярно изображение, 90
вектори, 2

Съдържание

1 Диференциално смятане **1**
1.1 Разстояние и норма в $\mathbb{R}^n$ 1
1.2 Топология и сходимост в $\mathbb{R}^n$ 10
1.3 Непрекъснатост на функции и изображения 22
1.4 Диференцируемост на функции и изображения 35
1.5 Производна по направление и градиент 52
1.6 Производни от по-висок ред 58
1.7 Формула на Тейлор и локални екстремуми. 64
1.8 Теорема за неявната функция. 76
1.9 Обвивки на фамилии от криви. 93
1.10 Множители на Лагранж 102
1.11 Функционална независимост. 112