

Лекция 14

§14. Екстремуми

1. Квадратични форми. Функцията $\varphi(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, се нарича **квадратична форма** на променливите x_1, x_2, \dots, x_n , когато има вида

$$\varphi(\mathbf{x}) = \sum_{i,j=1}^n a_{ij} x_i x_j.$$

За коефициентите предполагаме, че $a_{ij} = a_{ji}$, $1 \leq i, j \leq n$. При $n=1$ имаме $\varphi(x) = ax^2$, при $n=2$, $\varphi(x, y) = ax^2 + 2bxy + cy^2$. Коефициентите на формата образуват симетрична матрица $A = \{a_{ij}\}_{i,j=1}^n$ и обратно, всяка симетрична матрица поражда квадратична форма. С помощта на скалярно произведение, квадратичната форма може да се запише във вида

$$(14.1) \quad \varphi(\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle,$$

където $A\mathbf{x}$ е векторът, получен от умножението на $n \times n$ матрицата A с вектор стълба на променливите \mathbf{x} . Чрез матрично умножение, скалярното произведение се записва $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T \mathbf{x}$, следователно

$$\varphi(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}.$$

Някои квадратични форми приемат стойности с един и същи знак при всяко $\mathbf{x} \neq \mathbf{0}$. Такива форми се наричат **знакоопределени**.

Определение 14.1. Казва се, че квадратичната форма $\varphi(\mathbf{x})$ е положително (отрицателно) определена, когато за всяко $\mathbf{x} \neq \mathbf{0}$ е изпълнено $\varphi(\mathbf{x}) > 0$ ($\varphi(\mathbf{x}) < 0$). Симетричната матрица A се нарича положително (отрицателно) определена, когато породената от нея квадратична форма е положително (отрицателно) определена.

Когато неравенствата не са строги, формата $\varphi(\mathbf{x})$ и свързаната с нея матрица се наричат **неотрицателно (неположително) определени**. От (14.1) следва, че ако матрицата A е положително определена, то матрицата $-A$ е отрицателно определена и обратно.

Твърдение 14.1. Нека $\varphi(\mathbf{x})$ е положително определена. Тогава съществува константа $m > 0$, за която $\varphi(\mathbf{x}) \geq m|\mathbf{x}|^2$.

Доказателство. Единичната сфера $S = \{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{x}| = 1\}$ е ограничено и затворено множество (компактно множество), следователно непрекъснатата функция $\varphi(\mathbf{x}): S \rightarrow \mathbb{R}$ е ограничена, при което достига в S най-малка стойност,

$$m = \min_{\mathbf{x} \in S} \varphi(\mathbf{x}) = \varphi(\mathbf{x}_{\min}),$$

за някое $\mathbf{x}_{\min} \in S$. По условие $\varphi(\mathbf{x}) > 0$, за $\mathbf{x} \neq \mathbf{0}$, следователно $m = \varphi(\mathbf{x}_{\min}) > 0$, понеже $\mathbf{x}_{\min} \neq \mathbf{0}$ ($|\mathbf{x}_{\min}| = 1$). От вида на $\varphi(\mathbf{x})$ следва, че за всяка константа λ е изпълнено

$\varphi(\lambda\mathbf{x}) = \lambda^2 \varphi(\mathbf{x})$. Нека сега $\mathbf{x} \neq \mathbf{0}$. Тогава векторът $\mathbf{y} = \frac{\mathbf{x}}{|\mathbf{x}|} \in S$, понеже

$$|\mathbf{y}| = \left| \frac{\mathbf{x}}{|\mathbf{x}|} \right| = \frac{1}{|\mathbf{x}|} |\mathbf{x}| = 1.$$

По определение за константата m имаме $\varphi(\mathbf{y}) \geq m$, следователно

$$\varphi\left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) \geq m \text{ или } \frac{1}{|\mathbf{x}|^2} \varphi(\mathbf{x}) \geq m,$$

за всяко $\mathbf{x} \neq \mathbf{0}$, което трябваше да докажем. ■

По същия начин се доказва, че ако $\varphi(\mathbf{x})$ е отрицателно определена, то съществува константа $m > 0$, за която $\varphi(\mathbf{x}) \leq -m|\mathbf{x}|^2$.

Положителната (отрицателната) определеност на матрицата A зависи от знаците на нейните главни минори

$$D_1 = a_{11}, D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, D_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots, D_n = \det A.$$

Теорема 14.1 (критерий на Силвестър). Симетричната матрица A е положително определена тогава и само тогава, когато всичките главни минори са положителни, $D_1 > 0, D_2 > 0, D_3 > 0, \dots, D_n > 0$. ■

Ако A е положително определена, то като разместим редовете и стълбовете по един и същи начин отново се получава положително определена матрица, понеже това разместване означава пренареждане на променливите, което не променя положителната определеност. В този случай от критерия на Силвестър следва, че всеки диагонален елемент на A е положителен и изобщо всеки минор, получен от пресичането на някои редове и стълбове с едни и същи номера е положителен.

Матрицата A е отрицателно определена, когато $-A$ е положително определена, следователно от критерия на Силвестър се получава верността на

Твърдение 14.2. Симетричната матрица A е отрицателно определена тогава и само тогава, когато за нейните главни минори е изпълнено $D_1 < 0, D_2 > 0, D_3 < 0, \dots, (-1)^n D_n > 0$. ■

Пример 14.1. Матрицата

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

е положително определена, понеже

$$D_1 = 3 > 0, D_2 = \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 5 > 0 \text{ и } D_3 = \det A = 2 > 0.$$

Това означава, че формата породена от A

$$\varphi(x, y, z) = 3x^2 + 2y^2 + z^2 + 2xy + 2yz > 0,$$

винаги когато поне една от променливите x, y или z е различна от нула.

От курса по линейна алгебра знаем, че за всяка симетрична матрица A може да се намери унитарна матрица H , която **диагонализира** A в следния смисъл,

$$H^T A H = \Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Матрицата H се нарича **ортогонална**, когато $H^T H = I_n$, където I_n е единичната матрица от ред n (за ортогоналните матрици $H^{-1} = H^T$). Тогава след смяна на променливите $\mathbf{x} = H\mathbf{y}$, формата (14.1) приема вида

$$\begin{aligned} \varphi(H^T \mathbf{y}) &= \langle A H \mathbf{y}, H \mathbf{y} \rangle = \langle H^T A H \mathbf{y}, \mathbf{y} \rangle = \langle \Lambda \mathbf{y}, \mathbf{y} \rangle, \\ (14.2) \quad \varphi(H^T \mathbf{y}) &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2. \end{aligned}$$

Смяната на променливите посредством ортогонална матрица запазва скаларното произведение, понеже, ако $\mathbf{x}_1 = H\mathbf{y}_1$ и $\mathbf{x}_2 = H\mathbf{y}_2$, то

$$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle H\mathbf{y}_1, H\mathbf{y}_2 \rangle = \langle H^T H \mathbf{y}_1, \mathbf{y}_2 \rangle = \langle \mathbf{y}_1, \mathbf{y}_2 \rangle,$$

следователно тази смяна запазва дължините на векторите и ъглите между тях и по този начин представлява преобразуване на еднаквост в \mathbb{R}^n (завъртане на координатната система около началото).

Числата $\lambda_1, \lambda_2, \dots, \lambda_n$ се наричат **собствени стойности** на матрицата A и са корени на характеристичния полином

$$\chi(\lambda) = \det(\lambda I_n - A) = \det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & -a_{nn} \end{pmatrix}$$

и по тази причина, в конкретен случай, могат да бъдат намерени без да има необходимост от познаване на матрицата H .

От (14.2) се получава друг критерий за положителна (отрицателна) определеност. Условието $\varphi(\mathbf{x}) > 0$, за всяко $\mathbf{x} \neq \mathbf{0}$, означава, че дясната страна на (14.2) е положителна, за всяко $\mathbf{y} \neq \mathbf{0}$, а последното е вярно тогава и само тогава, когато всичките собствени значения $\lambda_1, \lambda_2, \dots, \lambda_n$ на матрицата A са положителни. Следователно е вярна

Теорема 14.2. Симетричната матрица A е положително (отрицателно) определена тогава и само тогава, когато всичките собствени значения са положителни (отрицателни). ■

Когато всичките собствени значения на A са различни от нула и между тях има както положителни така и отрицателни се казва, че квадратичната форма има поведение от тип **седло**. В този случай променливите y_1, y_2, \dots, y_n могат да се разделят условно на две групи според знака на съответното λ в дясната страна на (14.2). Нека за определеност $\lambda_k > 0$, за $k = 1, \dots, s$, и $\lambda_k < 0$, за $k = s+1, \dots, n$. Тогава формата $\varphi(\mathbf{x}) > 0$, за всяко $\mathbf{x} = H^T \mathbf{y}$, където $\mathbf{y} = (y_1, \dots, y_s, 0, \dots, 0) \neq \mathbf{0}$ и $\varphi(\mathbf{x}) < 0$, за всяко $\mathbf{x} = H^T \mathbf{y}$, където $\mathbf{y} = (0, \dots, 0, y_{s+1}, \dots, y_n) \neq \mathbf{0}$. В новата координатна система по една група от променливите формата е положително определена, а по групата на останалите е отрицателно определена. Произведението на собствените числа на A е равно на детерминантата $\det A$, понеже

$$\begin{aligned} \lambda_1 \lambda_2 \dots \lambda_n &= \det \Lambda = \det(H^T A H) = \det H^T \det A \det H = \\ &= \det H^T \det H \det A = \det(H^T H) \det A = \det I_n \det A = \det A \end{aligned}$$

следователно **всичките собствени числа на A са различни от нула тогава и само тогава, когато $\det A \neq 0$.**

2. Екстремуми на функция на много променливи. Нека функцията $f(\mathbf{x})$ е определена в областта G и нека $\mathbf{x}^{(0)} \in G$. Казва се, че $f(\mathbf{x})$ има **локален минимум** в

точката $\mathbf{x}^{(0)}$, когато $f(\mathbf{x}) \geq f(\mathbf{x}^{(0)})$, за всяко \mathbf{x} от някаква δ -околност $B(\mathbf{x}^{(0)}, \delta)$. Ако $f(\mathbf{x}) > f(\mathbf{x}^{(0)})$ за $\mathbf{x} \in B(\mathbf{x}^{(0)}, \delta)$, $\mathbf{x} \neq \mathbf{x}_0$, то минимумът се нарича **строг**. Определенията за **локален максимум** и **строг локален максимум** са аналогични.

Нека $f(\mathbf{x})$ е непрекъснато диференцируема в $\mathbf{x}^{(0)}$, която е точка на локален екстремум и нека $\mathbf{n} = (l_1, l_2, \dots, l_n)$ е някакъв единичен вектор, $|\mathbf{n}| = 1$. Тогава функцията на една променлива

$$\phi(t) = f(\mathbf{x}^{(0)} + t\mathbf{n}) = f(x_1^{(0)} + tl_1, x_2^{(0)} + tl_2, \dots, x_n^{(0)} + tl_n)$$

е диференцируема в околност на точката $t_0 = 0$ и има екстремум от същия вид, следователно от теоремата на Ферма получаваме, че $\phi'(0) = 0$. От друга страна $\phi'(0)$ е точно производната на $f(\mathbf{x})$ по направление \mathbf{n} в точката $\mathbf{x}^{(0)}$. Това показва, че производната по всяко направление в точката $\mathbf{x}^{(0)}$ е равна на нула, следователно всичките частни производни на $f(\mathbf{x})$ в точката $\mathbf{x}^{(0)}$ са равни на нула и градиентът е нулевият вектор, $\nabla f(\mathbf{x}^{(0)}) = \mathbf{0}$. По този начин доказахме следното необходимо условие за екстремум, което очевидно е аналог на познатата теорема на Ферма за функция на една променлива.

Твърдение 14.3. Нека $f(\mathbf{x})$ е непрекъснато диференцируема в $\mathbf{x}^{(0)}$, която е точка на локален екстремум. Тогава $\nabla f(\mathbf{x}^{(0)}) = \mathbf{0}$ и производните по всички направления в точката $\mathbf{x}^{(0)}$ са равни на нула. ■

С помощта на пълен диференциал, заключението на твърдение 14.3 се записва $df(\mathbf{x}^{(0)}) = 0$. Точки, в които градиентът се анулира, се наричат **стационарни точки** за функцията $f(\mathbf{x})$. Твърдение 14.3 гласи, че всяка екстремална точка е стационарна. Обратното както знаем не е вярно даже за функция на една променлива.

Сега ще установим условия, при които една стационарна точка е точка на локален екстремум. За тази цел ще използваме формулата на Тейлър. Нека $f(\mathbf{x})$ е определена и има непрекъснати производни до трети ред в някаква околност $B(\mathbf{x}^{(0)}, \delta)$ на стационарната точка $\mathbf{x}^{(0)}$. Тогава, според формулата на Тейлър, имаме

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}) = f(\mathbf{x}^{(0)}) + \langle \nabla f(\mathbf{x}^{(0)}), \Delta\mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}, \Delta\mathbf{x} \rangle + o(|\Delta\mathbf{x}|^2), \quad |\Delta\mathbf{x}| < \delta,$$

следователно

$$(14.3) \quad f(\mathbf{x}^{(0)} + \Delta\mathbf{x}) - f(\mathbf{x}^{(0)}) = \frac{1}{2} \langle \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}, \Delta\mathbf{x} \rangle + |\Delta\mathbf{x}|^2 \varepsilon(|\Delta\mathbf{x}|), \quad \lim_{|\Delta\mathbf{x}| \rightarrow 0} \varepsilon(|\Delta\mathbf{x}|) = 0,$$

където

$$\mathbf{H}(\mathbf{x}^{(0)}) = \begin{pmatrix} f''_{x_1x_1}(\mathbf{x}^{(0)}) & f''_{x_1x_2}(\mathbf{x}^{(0)}) & \dots & f''_{x_1x_n}(\mathbf{x}^{(0)}) \\ f''_{x_2x_1}(\mathbf{x}^{(0)}) & f''_{x_2x_2}(\mathbf{x}^{(0)}) & \dots & f''_{x_2x_n}(\mathbf{x}^{(0)}) \\ \dots & \dots & \dots & \dots \\ f''_{x_nx_1}(\mathbf{x}^{(0)}) & f''_{x_nx_2}(\mathbf{x}^{(0)}) & \dots & f''_{x_nx_n}(\mathbf{x}^{(0)}) \end{pmatrix}$$

е хесианът на $f(\mathbf{x})$, пресметнат в $\mathbf{x}^{(0)}$. Дали $\mathbf{x}^{(0)}$ е точка на локален екстремум зависи от поведението на разликата $f(\mathbf{x}^{(0)} + \Delta\mathbf{x}) - f(\mathbf{x}^{(0)})$, при достатъчно малко $\Delta\mathbf{x}$, за която разлика получихме равенството (14.3). Да предположим, че хесианът $\mathbf{H}(\mathbf{x}^{(0)})$ е положително определена матрица и нека $m > 0$ е константа, за която (твърдение 14.1)

$$\langle \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}, \Delta\mathbf{x} \rangle \geq m |\Delta\mathbf{x}|^2.$$

Нека освен това $\delta > 0$ е избрано достатъчно малко, че за величината $\varepsilon(|\Delta \mathbf{x}|)$ от (14.3) да бъде изпълнено $|\varepsilon(|\Delta \mathbf{x}|)| < \frac{m}{4}$ когато $0 < |\Delta \mathbf{x}| < \delta$. Тогава, за всяко $\Delta \mathbf{x}$, $0 < |\Delta \mathbf{x}| < \delta$,

$$f(\mathbf{x}^{(0)} + \Delta \mathbf{x}) - f(\mathbf{x}^{(0)}) \geq \frac{m}{2} |\Delta \mathbf{x}|^2 - \frac{m}{4} |\Delta \mathbf{x}|^2 = \frac{m}{4} |\Delta \mathbf{x}|^2 > 0,$$

което показва, че при направените предположения, функцията $f(\mathbf{x})$ има строг локален минимум в точката $\mathbf{x}^{(0)}$. Аналогично се показва, че ако хесианът $\mathbf{H}(\mathbf{x}^{(0)})$ е отрицателно определена матрица, то функцията $f(\mathbf{x})$ има строг локален максимум в точката $\mathbf{x}^{(0)}$. По този начин доказахме

Твърдение 14.4. Нека $f(\mathbf{x})$ е определена и има непрекъснати производни до трети ред в някаква околност $B(\mathbf{x}^{(0)}, \delta)$ на стационарната точка $\mathbf{x}^{(0)}$. Тогава, ако хесианът $\mathbf{H}(\mathbf{x}^{(0)})$ е положително (отрицателно) определена матрица, то функцията $f(\mathbf{x})$ има строг локален минимум (максимум) в точката $\mathbf{x}^{(0)}$. ■

Условията за положителна (отрицателна) определеност на хесианът понякога се означават с помощта на втория диференциал $d^2 f(\mathbf{x}^{(0)}) > 0$ ($d^2 f(\mathbf{x}^{(0)}) < 0$) при $|d\mathbf{x}| > 0$, понеже формално вторият диференциал задава същата квадратична форма, само че относно "променливите" dx_1, dx_2, \dots, dx_n ,

$$d^2 f(\mathbf{x}) = \sum_{i,j=1}^n \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} dx_i dx_j.$$

Чрез същите разсъждения може да се установи, че ако квадратичната форма, породена от хесиана $\mathbf{H}(\mathbf{x}^{(0)})$ има поведение от тип седло, то функцията $f(\mathbf{x})$ сигурно няма локален екстремум в стационарната точка $\mathbf{x}^{(0)}$, което означава, че могат да се намерят точки \mathbf{x} , произволно близки до $\mathbf{x}^{(0)}$, за които $f(\mathbf{x}) > f(\mathbf{x}^{(0)})$, както и произволно близки точки, за които $f(\mathbf{x}) < f(\mathbf{x}^{(0)})$. В този случай се казва, че $\mathbf{x}^{(0)}$ е **седлова точка** за функцията $f(\mathbf{x})$.

Като съпоставим направените изводи с критерия на Силвестър, получаваме верността на следната

Теорема 14.3. Нека $f(\mathbf{x})$ е определена и има непрекъснати производни до трети ред в околност $B(\mathbf{x}^{(0)}, \delta)$ на стационарната точка $\mathbf{x}^{(0)}$ и да разгледаме главните минори на хесиана $\mathbf{H}(\mathbf{x}^{(0)})$

$$D_1 = f''_{x_1 x_1}(\mathbf{x}^{(0)}), \quad D_2 = \begin{vmatrix} f''_{x_1 x_1}(\mathbf{x}^{(0)}) & f''_{x_1 x_2}(\mathbf{x}^{(0)}) \\ f''_{x_2 x_1}(\mathbf{x}^{(0)}) & f''_{x_2 x_2}(\mathbf{x}^{(0)}) \end{vmatrix},$$

$$D_3 = \begin{vmatrix} f''_{x_1 x_1}(\mathbf{x}^{(0)}) & f''_{x_1 x_2}(\mathbf{x}^{(0)}) & f''_{x_1 x_3}(\mathbf{x}^{(0)}) \\ f''_{x_2 x_1}(\mathbf{x}^{(0)}) & f''_{x_2 x_2}(\mathbf{x}^{(0)}) & f''_{x_2 x_3}(\mathbf{x}^{(0)}) \\ f''_{x_3 x_1}(\mathbf{x}^{(0)}) & f''_{x_3 x_2}(\mathbf{x}^{(0)}) & f''_{x_3 x_3}(\mathbf{x}^{(0)}) \end{vmatrix}, \dots, \quad D_n = \det \mathbf{H}(\mathbf{x}^{(0)}).$$

Тогава:

- 1) Ако $D_1 > 0, D_2 > 0, D_3 > 0, \dots, D_n > 0$, то функцията $f(\mathbf{x})$ има строг локален минимум в точката $\mathbf{x}^{(0)}$.
- 2) Ако $D_1 < 0, D_2 > 0, D_3 < 0, \dots, (-1)^n D_n > 0$, то функцията $f(\mathbf{x})$ има строг локален максимум в точката $\mathbf{x}^{(0)}$.

3) Ако $D_n = \det \mathbf{H}(\mathbf{x}^{(0)}) \neq 0$, но не са изпълнени условията за строг локален минимум от пункт 1) нито условията за строг локален максимум от пункт 2), то $\mathbf{x}^{(0)}$ е седлова точка за функцията $f(\mathbf{x})$. ■

Пример 14.2. Да намерим точките на екстремум на функцията
 $f(x, y, z) = x^2 + 2xy + 4xz + 8yz + 5y^2 + 9z^2$.

Стационарните точки на тази функция определяме от системата

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x + 2y + 4z = 0, \\ \frac{\partial f}{\partial y} &= 2x + 10y + 8z = 0, \\ \frac{\partial f}{\partial z} &= 4x + 8y + 18z = 0. \end{aligned}$$

За детерминантата на тази хомогенна система имаме

$$\begin{vmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{vmatrix} = 8 \begin{vmatrix} 1 & 1 & 2 \\ 1 & 5 & 4 \\ 2 & 4 & 9 \end{vmatrix} = 8 \cdot 16 \neq 0,$$

следователно единственото решение е точката $M_0(0,0,0)$. Хесианът на $f(x, y, z)$ във всяка точка има вида

$$\mathbf{H} = \begin{pmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 8 \end{pmatrix},$$

откъдето намираме

$$D_1 = 2 > 0, \quad D_2 = \begin{vmatrix} 2 & 2 \\ 2 & 10 \end{vmatrix} = 16 > 0, \quad D_3 = \begin{vmatrix} 2 & 2 & 4 \\ 2 & 10 & 8 \\ 4 & 8 & 18 \end{vmatrix} = 8 \cdot 16 > 0,$$

следователно, съгласно теорема 14.3, функцията $f(x, y, z)$ има строг локален минимум в точката $M_0(0,0,0)$.

Да разгледаме подробно случая на функция на две променливи $f(x, y)$ и стационарна точка $M_0(x_0, y_0)$. За хесиана имаме

$$\mathbf{H}(x_0, y_0) = \begin{pmatrix} f''_{xx}(x_0, y_0) & f''_{xy}(x_0, y_0) \\ f''_{yx}(x_0, y_0) & f''_{yy}(x_0, y_0) \end{pmatrix}.$$

Да положим $A = f''_{xx}(x_0, y_0)$, $B = f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$, $C = f''_{yy}(x_0, y_0)$. Тук имаме $D_1 = A$ и $D_2 = AC - B^2$. И в двата случая на екстремум е налице $AC - B^2 > 0$, а в кой от случаите се намираме зависи от знака на A . Когато $AC - B^2 < 0$ е налице седлова точка. Открит остава въпросът само когато $AC - B^2 = 0$. Следващият пример показва колко сложни могат да бъдат нещата, когато $AC - B^2 = 0$.

Пример 14.3. Да разгледаме функцията

$$f(x, y) = (y - x^2)(y - 2x^2)$$

в околност на стационарната точка $M_0(0,0)$. Производната на тази функция по всяко направление в точката $M_0(0,0)$ е равна на нула, по-точно, по всяко направление функцията има локален минимум, но въпреки това, $M_0(0,0)$ не е точка на локален

минимум, понеже $f(x, y)$ е отрицателна в ивицата между параболите $y = x^2$ и $y = 2x^2$ и положителна извън тази ивица (по самите параболи е нула).

3. Условен екстремум. Множители на Лагранж. Да разгледаме отначало случая на функция на две променливи $f(x, y)$ при наличие на едно условие $g(x, y) = 0$. Нека функцията $f(x, y)$ е непрекъснато диференцируема в областта G . Нека $E \subset G$ е множеството, определено от условието $g(x, y) = 0$, където $g(x, y)$ е непрекъснато диференцируема в G . Казва се, че функцията $f(x, y)$ има **условен локален минимум** в точката $M_0 \in E$, $g(x_0, y_0) = 0$, $|g'_x(x_0, y_0)| + |g'_y(x_0, y_0)| > 0$, когато съществува околност $B((x_0, y_0), \delta)$ такава, че $f(x, y) \geq f(x_0, y_0)$, за всяко $(x, y) \in E \cap B((x_0, y_0), \delta)$. Ако $f(x, y) > f(x_0, y_0)$ за $(x, y) \in E \cap B((x_0, y_0), \delta)$, $(x, y) \neq (x_0, y_0)$, то минимумът се нарича **строг**. Определенията за условен локален максимум и строг условен локален максимум са аналогични.

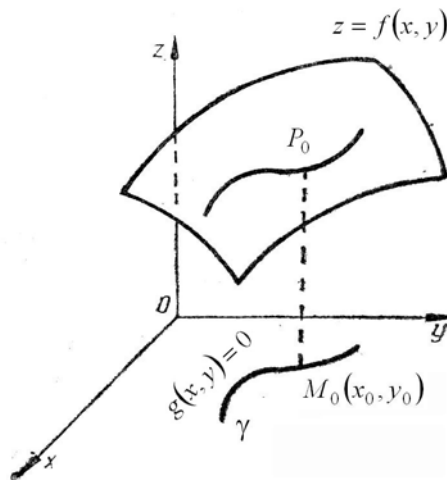


Рис. 14.1.

Условието $g(x, y) = 0$ задава някаква крива γ , съдържаща точката M_0 (Рис. 14.1). Когато сравняваме стойностите на $f(x, y)$ със стойността $f(x_0, y_0)$, точката (x, y) се мени в някаква малка околност на (x_0, y_0) , оставайки по кривата γ .

Ако M_0 е точка на обикновен (безусловен) екстремум, то очевидно M_0 е точка на условен екстремум при всякакво условие. От друга страна една функция може да няма безусловен екстремум в дадена точка, но при различни условия да има условни екстремуми от различен вид, както се вижда от следния пример.

Пример 14.4. Да разгледаме функцията $f(x, y) = x^2 - y^2$, за която стационарната точка $M_0(0, 0)$ е седлова. При условието $y = 0$, функцията $f(x, y) = x^2$ има локален условен минимум в M_0 , а при условието $x = 0$, функцията $f(x, y) = -y^2$ има локален условен максимум в M_0 .

Нека $g'_y(x_0, y_0) \neq 0$. Тогава от теоремата за неявните функции следва, че променливата y може да се представи като неявна функция $y = \varphi(x)$ в околност $(x_0 - \delta, x_0 + \delta)$, $\delta > 0$, т.е. $g(x, \varphi(x)) = 0$ за $x \in (x_0 - \delta, x_0 + \delta)$ и $\varphi(x_0) = y_0$. Нека $f(x, y)$ има условен локален екстремум в $M_0(x_0, y_0)$. Това означава, че функцията на една променлива $\varphi(x) = f(x, \varphi(x))$ има обикновен екстремум от същия вид в x_0 и по

теоремата на Ферма $\phi'(x_0) = 0$. От друга страна $\phi'(x) = f'_x(x, \phi(x)) + f'_y(x, \phi(x))\phi'(x)$, следователно

$$(14.4) \quad f'_x(x_0, y_0) + f'_y(x_0, y_0)\phi'(x_0) = 0.$$

За производната на неявната функция имаме също

$$(14.5) \quad g'_x(x_0, y_0) + g'_y(x_0, y_0)\phi'(x_0) = 0.$$

Да разгледаме сега **функцията на Лагранж**

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y),$$

където константата λ е **множител на Лагранж**, който ще определим от изискването

$$(14.6) \quad L'_y(x_0, y_0, \lambda) = f'_y(x_0, y_0) + \lambda g'_y(x_0, y_0) = 0, \lambda = \lambda_0 = -\frac{f'_x(x_0, y_0)}{g'_x(x_0, y_0)}.$$

Ще покажем, че при този избор на λ имаме също $L'_x(x_0, y_0, \lambda) = 0$. Наистина, като умножим (14.5) с λ и съберем почленно със (14.4) получаваме

$$[f'_x(x_0, y_0) + \lambda g'_x(x_0, y_0)] + [f'_y(x_0, y_0) + \lambda g'_y(x_0, y_0)]\phi'(x_0) = 0,$$

което заедно със (14.6) показва, че $L'_x(x_0, y_0, \lambda) = f'_x(x_0, y_0) + \lambda g'_x(x_0, y_0) = 0$. Очевидно $L'_\lambda(x_0, y_0, \lambda_0) = g(x_0, y_0) = 0$. По този начин доказахме

Теорема 14.4. Нека $M_0(x_0, y_0)$ е точка на локален условен екстремум за функцията $f(x, y)$ с условие $g(x, y) = 0$, при което $f(x, y)$ и $g(x, y)$ са непрекъснато диференцируеми в околност на M_0 и $|g'_x(x_0, y_0)| + |g'_y(x_0, y_0)| > 0$. Тогава съществува константа λ_0 (множител на Лагранж) такава, че точката (x_0, y_0, λ_0) е стационарна за функцията на Лагранж $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$. ■

Теорема 14.4 ни учи, че условните екстремуми трябва да се търсят между стационарните точки на функцията на Лагранж, т.е. като решения на системата

$$L'_x(x, y, \lambda) = f'_x(x, y) + \lambda g'_x(x, y) = 0,$$

$$L'_y(x, y, \lambda) = f'_y(x, y) + \lambda g'_y(x, y) = 0,$$

$$L'_\lambda(x, y, \lambda) = g(x, y) = 0.$$

Последното уравнение е всъщност самото условие.

Пример 14.5. Да намерим условните екстремуми на функцията $f(x, y) = e^{xy}$ при ограничението $g(x, y) = x^3 + y^3 + x + y - 4 = 0$. Тук функцията на Лагранж е

$$L = e^{xy} + \lambda(x^3 + y^3 + x + y - 4),$$

а системата за определяне на стационарните точки има вида

$$L'_x = ye^{xy} + \lambda(3x^2 + 1) = 0,$$

$$L'_y = xe^{xy} + \lambda(3y^2 + 1) = 0,$$

$$L'_\lambda = x^3 + y^3 + x + y - 4 = 0.$$

Умножаваме първото равенство с x , второто с y и изваждаме. Получаваме

$$(14.7) \quad \lambda(3x^3 - 3y^3 - x + y) = \lambda(x - y)(3x^2 + 3xy + 3y^2 + 1) = 0.$$

Ако $\lambda = 0$, то ще имаме $x = y = 0$, което не удовлетворява третото уравнение (условието). Множителят $3x^2 + 3xy + 3y^2 + 1 > 0$ за всеки x и y , следователно (14.7) ни дава единствено възможността $x = y$. Сега замествайки в третото уравнение намираме

$x^2 + x = 2$, следователно $x = y = 1$. При така намерените стойности, за λ получаваме $\lambda = -\frac{e}{4}$. Единствената стационарна точка е $\left(1, 1, -\frac{e}{4}\right)$.

За да определим дали в тази точка функцията има локален условен екстремум и вида на екстремума ще пресметнем втория диференциал по променливите x и y на функцията на Лагранж

$$d^2L(1,1,\lambda_0) = e \left[(dx + dy)^2 + 2dxdy - \frac{3}{2}(dx^2 + dy^2) \right].$$

Тук обаче променливите dx и dy не са независими, понеже са свързани от условието. Тази връзка определяме след като пресметнем пълния диференциал на връзката $g'_x(1,1)dx + g'_y(1,1)dy = 0$, откъдето намираме $dx + dy = 0$. Сега като заместим, получаваме $d^2L(1,1,\lambda_0) = -5edx^2 < 0$, което означава, че става дума за условен локален максимум.

В общия случай се постъпва аналогично. Нека $\mathbf{x}^{(0)} \in G$ е точка на локален условен екстремум за функцията $f(\mathbf{x})$, която е непрекъснато диференцируема в областта G , при наличие m на брой връзки ($m < n$) $g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_m(\mathbf{x})$, които също се предполагат непрекъснато диференцируеми в точката $\mathbf{x}^{(0)}$, при което се иска още рангът на **якобиана (матрицата производна)**

$$J(\mathbf{x}^{(0)}) = \begin{pmatrix} \frac{\partial g_1(\mathbf{x}^{(0)})}{\partial x_1} & \frac{\partial g_1(\mathbf{x}^{(0)})}{\partial x_2} & \dots & \frac{\partial g_1(\mathbf{x}^{(0)})}{\partial x_n} \\ \frac{\partial g_2(\mathbf{x}^{(0)})}{\partial x_1} & \frac{\partial g_2(\mathbf{x}^{(0)})}{\partial x_2} & \dots & \frac{\partial g_2(\mathbf{x}^{(0)})}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_m(\mathbf{x}^{(0)})}{\partial x_1} & \frac{\partial g_m(\mathbf{x}^{(0)})}{\partial x_2} & \dots & \frac{\partial g_m(\mathbf{x}^{(0)})}{\partial x_n} \end{pmatrix}$$

да бъде максимален, $r(J(\mathbf{x}^{(0)})) = m$. Тогава могат да се намерят константи $\lambda_1^{(0)}, \lambda_2^{(0)}, \dots, \lambda_m^{(0)}$ такива, че точката $(\mathbf{x}^{(0)}, \boldsymbol{\lambda}^{(0)})$ да бъде стационарна за **функцията на Лагранж**

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \lambda_1 g_1(\mathbf{x}) + \lambda_2 g_2(\mathbf{x}) + \dots + \lambda_m g_m(\mathbf{x}).$$

Изследването на откритите след решаване на съответната система стационарни точки се провежда по аналогичен начин.

Пример 14.6. Да намерим условните екстремуми на линейната функция $u = x - 2y + 2z$ върху сферата $x^2 + y^2 + z^2 = 1$. Тук функцията на Лагранж има вида

$$L = x - 2y + 2z + \lambda(x^2 + y^2 + z^2 - 1).$$

Нейните стационарни точки определяме от системата

$$L'_x = 1 + 2\lambda x = 0,$$

$$L'_y = -2 + 2\lambda y = 0,$$

$$L'_z = 2 + 2\lambda z = 0,$$

$$L'_\lambda = x^2 + y^2 + z^2 - 1 = 0.$$

Като изключим x, y и z получаваме $\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 + \left(\frac{1}{\lambda}\right)^2 = 1$, откъдето намираме

$\lambda_1 = \frac{3}{2}$ и $\lambda_2 = -\frac{3}{2}$. Така за функцията на Лагранж получихме две стационарни точки

$$M_1\left(-\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}, \frac{3}{2}\right) \text{ и } M_2\left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}, -\frac{3}{2}\right).$$

Пресмятаме $d^2L(M_1) = 3(dx^2 + dy^2 + dz^2) > 0$, следователно M_1 е точка на условен локален минимум за функцията u , и $d^2L(M_2) = -3(dx^2 + dy^2 + dz^2) < 0$, следователно M_2 е точка на условен локален максимум за функцията u .