# Optimizavimo metodai. Paskaitų konspektas.

**Rimantas Grigutis**

## 1 paskaita. Optimizavimo uždavinio formulavimas. Globalaus ir lokalaus ekstremumo apibrėžimas.

### Optimizavimo uždavinio apibrėžimas.
### Globalaus ir lokalaus ekstremumo apibrėžimas.

Optimizavimo uždavinys - tai funkcijos, kurią čia vadiname tikslo funkcija, minimumo arba maksimumo radimas kurioje nors galimų reikšmių srityje.

Tegu $f(x)$, čia $x = (x_1, ..., x_n)$, yra tikslo funkcija apibrėžta Euklido erdvėje $\mathbf{R}^n$. Šios funkcijos reikšmės nusako tikslo, kuriam ir formuluojamas pats uždavinys, pasiekimo laipsnį. Tegu $X \subseteq \mathbf{R}^n$ yra galimų sprendinių sritis, kurioje ir yra vykdoma sprendinio paieška. Reikia rasti tokį x$^*$, priklausantį galimų sprendinių sričiai, kad tikslo funkcija šiame taške įgytų mažiausią ( didžiausią) reikšmę:

$$f(x^*) = \min_{x \in X} f(x)$$

**Apibrėžimas 1.1** Taškas $x^* \in X$ vadinamas funkcijos $f(x)$ globalaus minimumo tašku srityje X, jei

$$f(x^*) \leq f(x), \qquad \forall x \in X.$$

**Apibrėžimas 1.2** Taškas $x^* \in X$ vadinamas funkcijos $f(x)$ lokalaus minimumo tašku srityje X, jei egzistuoja toks $\varepsilon > 0$, kad

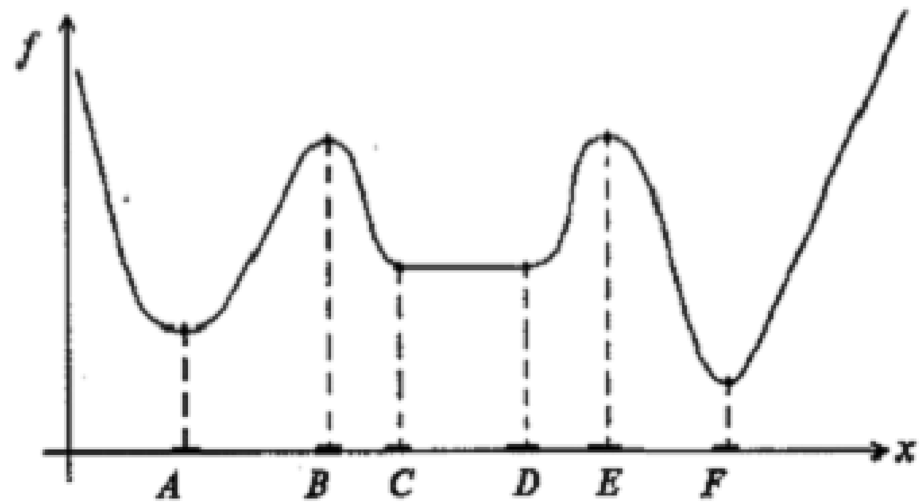
$$f(x^*) \leq f(x),$$

kai

$$\|x - x^*\| < \varepsilon \text{ ir} \qquad x \in X.$$

Čia $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}-$ vektoriaus $x$ ilgis Euklido erdvėje $\mathbf{R}^n$.

**Pavyzdys 1.3** Tegu $X = \mathbf{R}$ ir funkcijos $f(x)$ grafikas pavaizduotas piešinyje




Piešinyje yra išskirtos taškų $A, B, C, D, E, F$ $\varepsilon$-sritys. Taškas A yra lokalaus minimumo taškas, B - lokalaus maksimumo taškas, atkarpos CD taškai yra lokalaus minimumo taškai, taškas F yra lokalaus ir globalaus minimumo taškas. Funkcija neturi globalaus maksimumo taško.

**Funkcijos gradiento apibrėžimas**
**Hesse matricos apibrėžimas**
**Teigiamai( neigiamai) apibrėžta kvadratinė forma( apibrėžimai)**
**Neneigiamai( neteigiamai) apibrėžta kvadratinė forma( apibrėžimai)**

**Apibrėžimas 1.4**
Tolydžiai diferencijuojamos funkcijos f($x$) gradientu $\nabla f\,(x)$ vadiname jos dalinių išvestinių reikšmes taške x:

$$\nabla f\,(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \cdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}$$

Funkcijos gradiento kryptis sutampa su normalės vektoriaus, kuris yra statmenas liečiančiai plokštumai, einančiai per tašką x, kryptimi, kuri rodo didžiausią funkcijos pokytį duotame taške.

**Apibrėžimas 1.5**

Dukart tolydžiai diferencijuojamos funkcijos f($x$) taške x Hesse matrica vadinama antros eilės dalinių išvestinių reikšmių duotame taške matrica:

$$\mathrm{H}(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{pmatrix}$$

čia $h_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, i, j = 1, ..., n.$

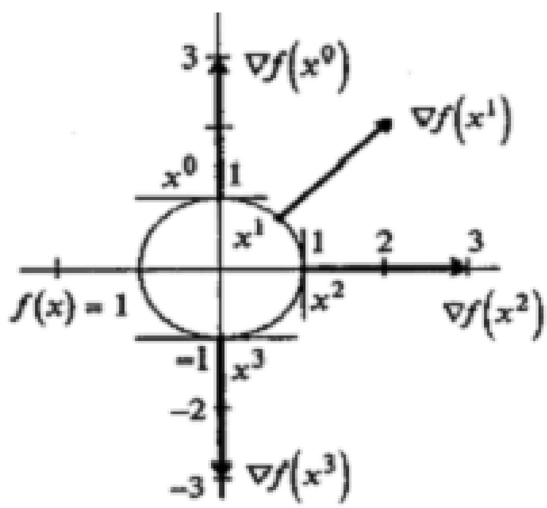
**Pavyzdys 1.6**

Funkcijai $f(x) = x_1^2 + x_2^2$ rasime gradientą taškuose $\mathrm{x}^0 = (0,1)^T$, $x^1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$, $x^2 = (1,0)^T$, $x^3 = (0,-1)^T$ ir rasime Hesse matricą.

Turime

$$\begin{gathered} \nabla f(x) = (2x_1; 2x_2)^T, \qquad \nabla f(x^0) = (0;2)^T, \\ \nabla f(x^1) = \left(\sqrt{2}; \sqrt{2}\right), \qquad \nabla f(x^2) = (2;0)^T \\ \nabla f(x^3) = (0;-2)^T, \qquad H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \end{gathered}$$

Pastebėkime, kad Hesse matrica nepriklauso nuo x. Žemiau esančiame paveikslėlyje pavaiduoti rasti gradientai



**Apibrėžimas 1.6**.

Nagrinėjamas Hesse matricos H($x^*$) stacionariame taške $x^*$ determinantas

$$\det H\left(x^*\right) = \begin{vmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{vmatrix}.$$

1. Determinantai $\Delta_1 = h_{11}, \Delta_2 = \begin{vmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{vmatrix}, ..., \Delta_n = \begin{vmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{vmatrix}$ vadinami kampiniais minorais.

2. m-osios eilės determinantai, gauti iš matricos H($x^*$) determinanto išbraukus kurias nors $(n-m)$ eilutes ir $(n-m)$ stulpelius su tais pačiais numeriais, vadinami pagrindiniais minorais.

**Apibrėžimai 1.7**

Hesse matrica $H\left(x^*\right)$ vadinama teigiamai apibrėžta, jei visų jos kampiniai minorai yra teigiami:

$$\Delta_1 > 0, \Delta_2 > 0, ..., \Delta_n > 0.$$

Hesse matrica $H\left(x^*\right)$ vadinama neigiamai apibrėžta, jei visų jos kampinių minorų ženklai kas kart keičiasi pradedant neigiamui:

$\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, ..., (-1)^n \Delta_n > 0.$

Hesse matrica $H\left(x^*\right)$ vadinama pusiau teigiamai apibrėžta, jei visų jos pagrindiniai minorai yra neneigiami.

Hesse matrica $H\left(x^*\right)$ vadinama pusiau neigiamai apibrėžta, jei visų jos lyginės eilės pagrindiniai minorai yra neneigiami, o nelyginės eilės pagrindiniai minorai - neteigiami.

**Pavyzdys 1.8**

Funkcijos $f\left(x\right) = x_1^2 + x_2^2$ Hesse matrica yra $H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$.Ši matrica yra teigiamai apibrėžta.

**Apibrėžimas 1.9**

Hesse matricos $H\left(x^*\right)$ tikrinėmis reikšmėmis $\lambda_i, i = 1, ..., n$ vadiname charakteristinio polinomo( tai n-ojo laipsnio polinomas):

$$
\det(H\left(x^{*}\right) - \lambda I_n) = \begin{vmatrix} h_{11}-\lambda & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22}-\lambda & \cdots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn}-\lambda \end{vmatrix} = 0
$$

šaknis.

**Pastaba 1.10**

Realiosios simetrinės Hesse matricos $H\left(x^{*}\right)$ visos tikrinės reikšmės yra realieji skaičiai.

**Teiginys 1.11**

Hesse matrica yra teigiamai apibrėžta tada ir tik tada, kai visos jos tikrinės reikšmės yra teigiamos: $\lambda_i > 0$ su visais $i$.

Hesse matrica yra neigiamai apibrėžta tada ir tik tada, kai visos jos tikrinės reikšmės yra neigiamos: $\lambda_i < 0$ su visais $i$.

Hesse matrica yra pusiau teigiamai apibrėžta tada ir tik tada, kai visos jos tikrinės reikšmės yra neneigiamos: $\lambda_i \geq 0$ su visais $i$.

Hesse matrica yra pusiau neigiamai apibrėžta tada ir tik tada, kai visos jos tikrinės reikšmės yra neteigiamos: $\lambda_i \leq 0$ su visais $i$.