Optimizavimo metodai. Paskaitų konspektas

Rimantas Grigutis

2 paskaita. Besąlyginio ekstremumo radimas

Besąlyginio ekstremumo radimo uždavinys

Duota dukart tolygiai diferencijuojama funkcija $f(x)$, apibrėžta aibėje $X = \mathbf{R}^n$. Reikia rasti funkcijos $f(x)$ ekstremumus, t.y. tokius taškus $x^* \in R^n$, kurie būtų nagrinėjamos funkcijos lokaliais minimumais ir maksimumais Euklido erdvėje $\mathbf{R}^n$:

$$f(x^*) = \min_{x \in \mathbf{R}^n} f(x); \qquad f(x^*) = \max_{x \in \mathbf{R}^n} f(x)$$

Sprendimo strategija.

Lokalūs ekstremumai x^* randami pirmosios ir antrosios eilės būtinomis sąlygomis, o taip pat besąlyginio lokalaus ekstremumo pakankama sąlyga. Vėliau surastuose taškuose skaičiuojamos funkcijos reikšmės $f(x^*)$.

Teiginys 2.1(pirmosios eilės būtinoji ekstremumo sąlyga).

Jei $x^* \in \mathbf{R}^n$ yra funkcijos $f(x)$ lokalaus minimumo(maksimumo) taškas ir funkcija $f(x)$ yra diferencijuojama taške x^*, tai funkcijos $f(x)$ gradientas taške x^* lygus nuliui:

$$\nabla f(x^*) = 0$$

arba

$$\frac{\partial f(x)}{\partial x_i} = 0, \qquad i = 1, ..., n. \qquad (2.1)$$

Apibrėžimas 2.2 Taškas x^* tenkinantis Teiginio 2.1 sąlygas vadinamas stacionariu tašku.

Teiginys 2.3(antrosios eilės būtinoji ekstremumo sąlyga).

Jei $x^* \in \mathbf{R}^n$ yra funkcijos $f(x)$ lokalaus minimumo(maksimumo) taškas ir funkcija $f(x)$ yra dukart diferencijuojama taške x^*, tai funkcijos $f(x)$ Hesse matrica taške x^*, $H(x^*)$, yra pusiau teigiamai apibrėžta(pusiau neigiamai apibrėžta):

$$H(x^*) \geq 0$$
$$(H(x^*) \leq 0).$$

Teiginys 2.4(pakankama ekstremumo sąlyga)

Jei funkcija $f(x)$ yra dukart diferencijuojama taške $x^* \in \mathbf{R}^n$, funkcijos $f(x)$ gradientas taške x^* lygus nuliui ir funkcijos $f(x)$ Hesse matrica taške $x^*, H(x^*)$, yra teigiamai apibrėžta(neigiamai apibrėžta):

$$\nabla f(x^*) = 0 \text{ ir } H(x^*) > 0 \ (H(x^*) < 0),$$

tai taškas x^* yra funkcijos $f(x)$ lokalaus minimumo(maksimumo) taškas aibėje $\mathbf{R}^n$.

Besąlyginio ekstremumo radimo algoritmas

1 žingsnis. Užrašyti pirmosios eilės būtinas ekstremumo sąlygas n lygčių sistema (2.1) ir išspendus šią sistemą rasti stacionarius taškus x^*.

2 žingsnis. Rastiems stacionariems taškams patikrinti pakankamas ekstremumo sąlygas ir jeigu šios sąlygos netenkinamos, tai patikrinti antrosios eilės būtinąją ekstremumo sąlygą: jei nėra išpildyta, tai nėra ekstremumo, jei išpildyta, tai reikia tęsti tyrimą.

3 žingsnis. Apskaičiuoti funkcijos reikšmes ekstremumo taškuose.

Pastabos 2.5

1. Sprendžiant praktinius uždavinius, kaip taisyklė, tyrimo tęsimas atliekant algoritmą nėra vykdomas.

2. Norint rasti globalius e kstremumo taškus reikia palyginti funkcijos reikšmes lokalių minimumų ir maksimumų taškuose ir išrinkti reikiamas.

3. Kai $n = 1$, tai algoritmo *2 žingsnis* yra toks:

Jeigu $f(x)$ ir jos išvestinės yra tolydžios funkcijos, tai taškas x^* yra ekstremumas tada ir tik tada, kai m - lyginis sveikas skaičius ir $f'(x^*) = f''(x^*) = \cdots = f^{(m-1)}(x^*) = 0$, bet $f^{(m)}(x^*) \neq 0$. Jei $f^{(m)}(x^*) > 0$, tai taškas x^* yra lokalusis minimumas, jei $f^{(m)}(x^*) < 0$, tai taškas x^* yra lokalusis maksimumas. Jeigu m - nelyginis, tai taške x^* ekstremumo nėra.

4. Praktiškai sprendžiant uždavinius, atskiruose taškuose, kuriuos randame skaitiniais metodais, tikriname pakankamas ir būtinas sąlygas. Tai daryti būtina, nes skaitiniais metodais randami tik stacionarūs taškai ir jų pobūdį būtina patikslinti.

Pavyzdys 2.6 Rasti funkcijos $f(x) = -x_1^2 - x_2^2 - x_3^2 - x_1 + x_1x_2 + 2x_3$ ekstremumą aibėje $\mathbf{R}^3$.

1. Užrašome pirmosios eilės būtinas ekstremumo sąlygas:
$\frac{df(x)}{dx_1} = -2x_1 - 1 + x_2 = 0, \qquad \frac{df(x)}{dx_2} = -2x_2 + x_1 = 0, \qquad \frac{df(x)}{dx_3} = -2x_3 + 2 = 0.$

Išsprendę šią sistemą gauname stacionarųjį tašką $x^* = \left(-\frac{2}{3}; -\frac{1}{3}; 1\right)^T$.

2. Patikrinsime pakankamas ekstremumo sąlygas

Pirmasis būdas. Funkcijos Hesse matrica yra

$$H(x^*) = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Ši matrica yra neigiamai apibrėžta, nes

$$\Delta_1 = -2 < 0, \Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 4 - 1 = 3 > 0, \Delta_3 = (-2) \cdot 3 = -6 < 0.$$

Taigi taškas x^* yra maksimumo taškas.

Antrasis būdas. Raskime Hesse matricos tikrines reikšmes spresdami lygtį:
$\det(H - \lambda I) = \begin{pmatrix} -2-\lambda & 1 & 0 \\ 1 & -2-\lambda & 0 \\ 0 & 0 & -2-\lambda \end{pmatrix} = \lambda^3 + 6\lambda^2 + 11\lambda + 6 = (\lambda+3)(\lambda+2)(\lambda+1) = 0.$

Visos tikrinės reikšmės yra neigiamos: $\lambda_1 = -1, \lambda_2 = -2, \lambda_3 = -3$. Taigi taškas x^* yra maksimumo taškas.

3. Funkcijos reikšmė lokalaus maksimumo taške yra $f(x^*) = \frac{4}{3}$.