Paskaita 1

Teorijos pagrindai (minimumas)

teorijos pagrindų

Nagrinėkime f-ją $y = f(x)$ apibrėžtą realiųjų skaičių tiesėje $\mathbb{R}$.

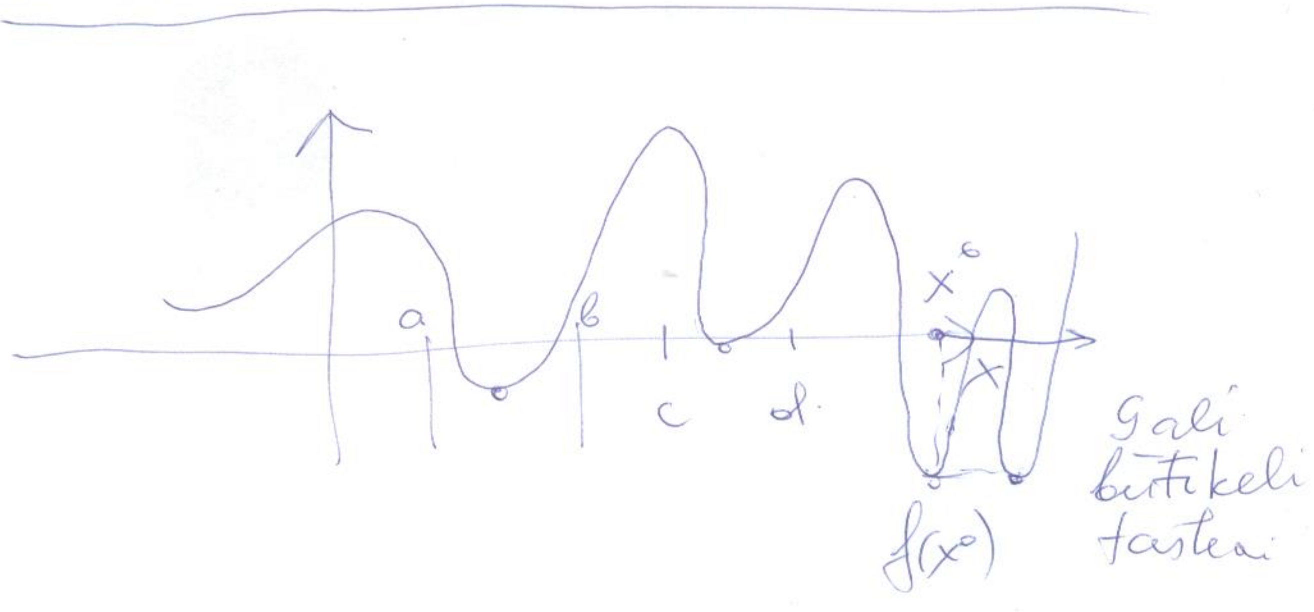
Ieškosime tokio taško $x^\circ \in \mathbb{R}$, kad

$$f(x^\circ) = \min f(x), \quad x \in \mathbb{R} \qquad (1^*)$$

(Tai reiškia, kad

$$f(x^\circ) \leq f(x), \quad x \in \mathbb{R}$$).

Tokį tašką x° vadinsime <u>globaliuoju minimumo tašku</u>



Lokalaus minimumo taškas.

Nagrinėjame taško x^{1} aplinką ε_2:

$$x^{1} \in [a, b]$$

x^{1} yra lokalaus minimumo taškas

jei

$$f(x^{1}) = \min_{x \in [a, b]} f(x)$$

$$\left(\text{t. y.} \quad f(x^{1}) \le f(x), \quad x \in [a, b] \right).$$

Abu apibrėžimai yra panašūs, bet jų sprendimo sudėtingumai yra labai skirtingi.

Globalaus optimizavimo uždavinys yra NP sudėtingumo. Jį galima spręsti taip:

Rasti minimumo taško uždaroje aprėžtoje aibėje $[a, b]$

1 Iš matematinės analizės paskaitų žinome pakankamų išsprendžiamumo sąlygą:

<u>Vejerstraso teorema</u>: Jeigu $\underline{f(x)}$ yra <u>tolydi funkcija</u> ir $x \in \bar{D}$, kur $\bar{D}$ yra <u>uždara</u>, <u>aprėžta</u> aibė, tai

a) $\exists m$ toks, kad $\boxed{f(x) \geq m, \ \forall x \in \bar{D}}$,

b) m yra <u>tiksliausi</u> apatiniai rėžis ir $\exists x^{\circ} \in \bar{D}$

$$m = f(x^{\circ})$$

Minimizavimo uždavinys turi sprendinį.

Kaip surasti x°, kai mes žinome, kad jis egzistuoja.

Galimas algoritmas (bet NP neapgausime)

1. Surandame visus lokalinio minimumo taškus

$$x_1^\circ, x_2^\circ, \dots, x_m^\circ \in D$$

$\tilde{x}^\circ$ – yra tas taškas, kuriame

$$f(\tilde{x}^\circ) \le f(x_j^\circ), \quad j = 1, \dots, m.$$

2. Surandame kraštinį tašką

$$\tilde{x}^1 \in \partial D$$

$$f(\tilde{x}^1) \le f(x), \quad x \in \partial D.$$

3. Palyginame

$$f(\tilde{x}^\circ) \text{ ir } f(\tilde{x}^1).$$

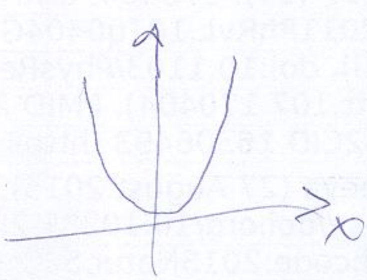
Būtinoji stacionarumo sąlyga

$$\frac{d f(x^0)}{dx} = 0$$

Toks taškas yra kandidatas būti minimumo tašku:

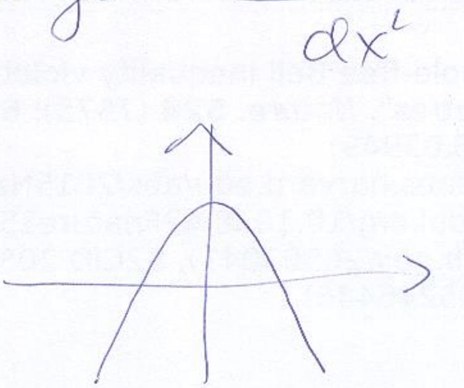
1. Pakankamoji sąlyga, kad turėtume minimumo (lokalaus) tašką:

$$\frac{d^2 f(x^0)}{dx^2} > 0$$



$$y = x^2$$

2. Jeigu $\frac{d^2 f(x^0)}{dx^2} < 0 \Rightarrow$ maksimumas (lokalus)



3) O kaip yra, jei

$$\frac{d^2 f(x^0)}{dx^2} = 0 \quad ?$$

$$y = x^3$$

Kai nagrinėjame $x \in R^n$, tai pakeičiame pateiktus rezultatus bendresniais: $f = f(x_1, \dots, x_n)$.

$$\frac{df}{dx} \to \text{grad} f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

$$\frac{d^2 f}{dx^2} \to H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)$$

Sąlyga $\frac{d^2 f}{dx^2} > 0$ yra keičiama kaip kvadratinės formos teigiamumas

$$y^T H(x^0)\, y > 0, \ \forall y \in R^n$$

y – vektorius stulpelis

Paskaita 1 (tęsinys)

Optimizavimo uždaviniai.

Optimizavimo uždaviniai visada buvo aktualūs – kiekvienoje veikloje siekiame rasti geresnį sprendimą, variantą, gaminio ar proceso charakteristiką, nei iki šiol naudojome ir naudoja konkurentai.

Šioje paskaitoje prisiminsime uždavinius, kuriuose optimizavimo žingsniai buvo svarbus etapas visame darbe.

Pradėkime nuo paprastos klasifikacijos (ją jau minėjome vieno kintamojo f-jos atveju):

a) lokalus ekstremumas

b) globalus ekstremumas.

Lokalus minimumas (maksimumas)

Def. Tarsime, kad X yra metrinės erdvės aibė (t.y galime apskaičiuoti atstumą tarp elementų). $\exists$ metrika $\rho(x,y)$)

Funkcijai (realaus kintamojo) taškas $x^* \in X$ yra lokalaus minimumo taškas (elementas, pvz. vektorius), jei $\exists\, \varepsilon > 0$:

$$f(x^*) \le f(x), \quad \forall x : \rho(x, x^*) \le \varepsilon.$$

Analogiškai apibrėžiame maksimumo sąvoką.

Aibės X savybės ypač svarbios, kai sudarysime minimizavimo uždavinių sprendimo algoritmus

a) atvira aibė X

b) uždara aibė X

c) iškiloji aibė X

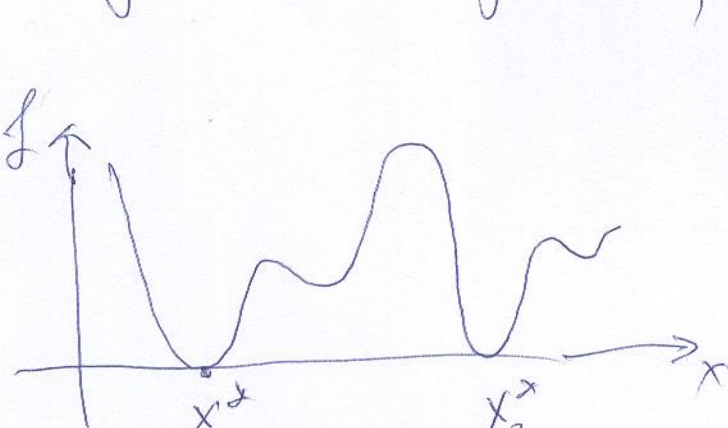
Globalus minimumo

$x^* \in X$ yra globalaus minimumo taškas (f funkcijos $f(x)$), jei

$$f(x^*) \le f(x), \quad \forall x \in X.$$

Ekstremumo taškas yra griežtas (strong) globalaus minimumo taškas, jei

$$f(x^*) < f(x), \quad \forall x \in X \text{ ir } x \neq x^*.$$



- nėra griežto (stipraus) globalaus ekstremumo taško

Globalaus ekstremumo taškų radimas yra daug sunkesnis uždavinys (skaičiavimų operacijų skaičius daug didesnis)

Todėl:

a) dažnai apsiribojame tik ekstremumo algoritmais, leidžiančiais rasti globalaus sprendinio artinį

b) supaprastiname uždavinį (sumažiname jo dimensiją), kad šio uždavinio globalaus ekstremumo taškus rastume daug efektyviau

Aptarsime abu variantus

a) Padaliname X į N sričių

$$X = X_1 \cup X_2 \cup \dots \cup X_N$$

Kiekvienoje srityje randame po lokalaus minimumo tašką

$$x_j^*$$

Tarp jų surandame mažiausią

$$\bar{x}^* = \min_{1 \le j \le N} x_j^*$$

Šio ~~euristinio~~ teorinio algoritmo pagrinda, yra teiginys, kad jeigu randame visus lokalius ekstremumo taškus (srityje X viduje ir ant krašto) tai tada galime tiksliai išspręsti ir globalaus ekstremumo uždavinį

b) sntyje X generuojame (dažniausiai atsitiktinai parenkame) N taškų $x_j \in X$, $j=1,\dots,N$.

Apskaičiuojame

$$f_j = f(x_j), \quad j=1,\dots,N$$

Randame globalaus minimumo artinį:

$\bar{x}^*$ yra randamas

$$\arg\min_{1\le j\le N} f_j$$

sprendinys

$$\bar{x}^* = x_k : \quad f_k \le f_j, \ j=1,\dots,N.$$

Šiuo metu žinomas geriausias artinys – o gal ir globalaus minimumo taškas

Optimizavimo uždavinių sprendimo algoritmų palyginimas:

1. Kiek veiksmų reikia atlikti, jei norime rasti sprendinį (tam tikru tikslumu).

2. Kokia informacija yra naudojama realizuojant algoritmą:

   2.1. $f(x)$ reikšmės

   2.2. $f(x)$, $f'(x)$ reikšmės

   2.3. $f(x)$, $f'(x)$, $f''(x)$ reikšmės

Dažnai skaičiuoti analizines išvestinių išraiškas yra sudėtinga (reikalaują daug veiksmų, ypač jei sprendžiame daugelio kintamųjų uždavinį.)

   2.4. Išvestines galime skaičiuoti aproksimuodami skaitiniais algoritmais

Stiprus argumentas, Absoliuti optimizavimo algoritmas yra šiuolaikinių - raktinių technologijų pavieniai

a) Mašininis mokymasis (ML)

b) DNT (ANN)

c) Dirbtinis Intelektas (AI)

Apmokymo metu minimizuojama netikties funkcija (loss function) ir nuo šio žingsnio tikrumo priklauso sudaryto Neuroninio tinklo galimybės