Paskaita 2

Vieno kintamojo funkcijų optimizav.

Nagrinėsime uždavinį, kai ieškome ~~užd~~

$$\min_{a\le x\le b} f(x) = f(x^*)$$

sprendinio. Tarsime, kad $f\in C[a,b]$ (sprendinio $\exists$ seka iš Vejerštraso teoremos)

yra tolydi ir x^* yra vienintelis lokalaus minimumo taškas (taigi x^* yra globalusis minimumo taškas).

Sąlygas, kada tokios prielaidos teisingos, nagrinėsime kitose paskaitose.

Naudosime informaciją tik apie $f(x)$ reikšmes (t.y. tik jas galime išmatuoti eksperimente)

Apibendrinsime paprastąją metodą, kurią naudojote spręsdami netiesines lygtis

$f(x) = 0$ (rasti $f(x^*) = 0$).

Tegul žinome, kad

$f(a) \cdot f(b) < 0$, (ženklai skirtingus galuose yra priešingi)

tada $\exists\, a < c < b$: $f(c) = 0$

Intervalo dalijimo pusiau metodas

Jis tinka ieškoti šaknies ir tada, kai intervale $[a, b]$ yra kelios šaknys, rasime bent vieną nors iš jų.

Minimizac. uždaviniui padarėme prielaidą, kad minimumo taškas vienintelis.

Toliau sprendžiame pavyzdį apie netiesinės vienos lygties šaknį

Skaičiuojame f-jos reikšmę

taške $y = \frac{a+b}{2}$: $f(y)$

Išsirenkame tą intervalą

kurio galuose galioja sąlyga

$$f(a)\,f(y) < 0 \Rightarrow [a, y]$$

$$f(y)\,f(b) < 0 \Rightarrow [y, b]$$

Sprendinio artinio fiksuoja

pagerėja du kartus (paklaida

sumažėja du kartus) atlikus

vieną iteraciją.

Turime geometrinę progresiją

konverguojančią greitai.

Iteracijų skaičius

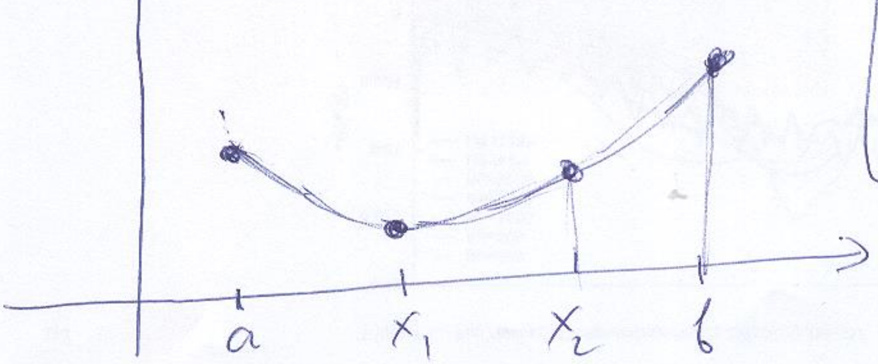
$$\boxed{\left(\frac{1}{2}\right)^N \le \varepsilon}$$

Dabar spręsime optimizavimo uždavinį auksinio pjūvio metodu

Intervale [a, b] pasirenkame du taškus $a < x_1 < x_2 < b$.

Kol kas neapibrėžiame tikslaus algoritmo, kaip randame šiuos taškus tokius, jog konvergavimo greitis į minimumo tašką x^* būtų didžiausias (greičiausiai mažėtų artinio paklaida).

Žinome tik f reikšmes taške $a < x_1 < x_2 < b$.



Kaip parinkti intervalą (intervalus), kuriuose tikrai negali egzistuoti minimumo taškas x^*?

<u>Garantuotas pasirinkimas</u> (universalus)

- atmesti tą tašką, kuris toliausiai nutolęs nuo <u>geriausio</u> (dabar žinomo) taško

$$x^{**} \doteq \arg\min_{x_j} f(x_j), \quad j = 0, \dots, 3$$

$$\underbrace{\underset{a}{\overset{\shortparallel}{x_0}} \leq x_1 < x_2 < \underset{b.}{\overset{\shortparallel}{x_3}}}$$

<u>Pvz</u> Mūsų pavyzdyje, tai bus taškas

$x^{**} = x_1 \Rightarrow$ $b = x_3$. yra krotimas taška

$\Downarrow$

pašaliname interv.

$(x_2, b]$.

Kaip parinkti taškus x_1 ir x_2, kad a) konvergavimo greitis būtų didžiausias.

ir

b) išnaudotume jau parinktus taškus ir kiekvienoje iteracijoje apskaičiuotume $f(x)$ reikšmę tik viename naujame taške.

Algoritmas

Iš anksto nežinome, kurį intervalą reikės atmesti, todėl

$$\boxed{x_1 - a = b - x_2}$$

Pirmoji lygtis

$$\boxed{\frac{x_2 - x_1}{x_2 - a} = \frac{x_1 - a}{b - a} = q}$$

Proporcingumas.

Naujas int.

Senas int.

Tada skaičiuojame

$$\frac{x_2 - x_1}{x_2 - a} = \frac{b-a-(x_1-a)-(b-x_2)}{b-a-(b-x_2)}$$

$$(b-x_2) = x-a$$

$$= \frac{(b-a) - 2(x_1-a)}{b-a-(x_1-a)} \quad (\text{dalijame iš } (b-a))$$

$$= \boxed{\frac{1-2q}{1-q} = q}$$

$$q^2 - 3q + 1 = 0$$

$$q = \frac{3-\sqrt{5}}{2} \approx 0{,}38 \quad (\text{auksinio pjūvio proporcija})$$

$$\boxed{x_1 = a + (b-a)q, \quad x_2 = b - (b-a)q.}$$

Gradientinis nusileidimas

Nagrinėkime vieno kintamojo uždavinį

$$\min_{x} f(x)$$

ir kelių kintamųjų atvejį:

$$\min_{(x_1,\dots,x_n)} f(x_1,\dots,x_n)$$

Tada iš matematinės analizės kurso ir analizinės geometrijos teorijos žinome, kad f-jos didžiausio didėjimo kryptis $\bar{x}$-taške yra apibrėžiama jos gradiento taške $\bar{x}$

$f'(\bar{x})$ $\quad n = 1.$

$(f_{x_1}(\bar{x}), \dots, f_{x_n}(\bar{x}))^T$ $\quad n > 1.$ vektorius stulpelis.

Kadangi mes domina minimizacijos uždavinys tai pokytį skaičiuojame antigradiento kryptimi

$$x^{m+1} = x^{m} - \alpha f'(x^{m})$$

$$\begin{pmatrix} x_1^{m+1} \\ \vdots \\ x_n^{m+1} \end{pmatrix} = \begin{pmatrix} x_1^{m} \\ \vdots \\ x_n^{m} \end{pmatrix} - \alpha \begin{pmatrix} f'_{x_1}(x_1^m,\dots,x_n^m) \\ \vdots \\ f'_{x_n}(x_1^m,\dots,x_n^m) \end{pmatrix}$$

Aišku, kad ~~t~~tokiu būdu artėsime prie kažkurio lokalaus minimumo taško.

Klausimai:

1) Koks yra gradientinio nusileidimo metodo konvergavimo greitis?

2) Kaip parinkti parametrą α?

Tarkime, kad nagrinėsime konvergavimą į elementą x^*.

Žinome, kad $f'(x^*) = 0$.

$z^m = x^m - x^*$ (pažymėkime).

$$z^{m+1} = z^m - \alpha \left(f'(x^m) - f'(x^*) \right)$$

$$= z^m - \alpha f''(\tilde{x}^m) z^m = q z^m$$

Konvergavimo greitis yra tiesinis.

Eksperimentinis parametro α parinkimas

Imkime keletą α reikšmių:

$$\alpha_1, \alpha_2, \ldots, \alpha_j$$

Apskaičiuojame x_{j+}^{m+1} reikšmes, juose randame f-jos $f(x)$ reikšmę

$$f_j^{m+1} = f(x_j^{m+1})$$

Pasirenkame tą $x_{j_0}^{m+1}$ $\left(x^{m+1} = x_{j_0}^{m+1}\right)$, kuriam:

$$f_{j_0}^{m+1} \leq f_j^{m+1} \quad \left(\arg\min_j f_j^{m+1}\right).$$

Niutono metodas.

Šiuo metodu sprendžiame lygtį

$$f'(x) = 0$$

$$x^* = x^m + (x^* - x^m)$$

$$0 = f'(x^*) = f'(x^m) + f''(x^m)(x^* - x^m) + O\left[(x^* - x^m)^2\right]$$

Naują artinį randame taip

$$\boxed{x^{m+1} = x^m - \frac{1}{f''(x^m)} f'(x^m)}$$

Taigi $\alpha = \frac{1}{f''(x^m)}$, o konvergavimo greitis yra kvadratinis (tiksliau sprendinio pakankamai mažoje aplinkoje)

Jeigu sprendžiame kelių kintamųjų f-jos minimizavimo uždavinį, tai

$$\vec{x}^{m+1} = \vec{x}^{m} - H^{-1}(\vec{x}^{m})(\operatorname{grad} f)(x^{m})$$

$H = \left(f''_{x_i x_j}(\vec{x}^{m})\right)$ yra antrųjų išvestinių matrica.

Modifikacija (reguliarizuotas Niutono metodas)

Pradžioje, kol artinys nėra pakankamai tikslus, naudojame Niutono ir gradientinio metodo kombinaciją:

$$\vec{x}^{m+1} = \vec{x}^{m} - \left[\gamma(1-\alpha)I + \alpha H^{-1}(x^{m})\right] \times$$
$$\times (\operatorname{grad} f)(x^{m})$$

γ - iterac. parametras

$\alpha = \alpha_1 < \alpha_2 < \dots < \alpha_g = 1.$