eman ta zabal zazu

Universidad del País Vasco Euskal Herriko Unibertsitatea

ZIENTZIA ETA TEKNOLOGIA FAKULTATEA

FACULTAD DE CIENCIA Y TECNOLOGÍA

2. Erroen kalkulua eta optimizazio problemak $\mathbb{R}$-n.

Jon Asier Bárcena-Petisco (Euskal Herriko Unibertsitatea) - Zenbakizko Metodoak I irakasgaia

Gai honetan ikusiko ditugun edukiak

- Funtzio baten erroak kalkulatzeko metodoak
- Funtzio baten puntu finkoak aurkitzeko metodoak
- Funtzio baten minimo eta maximo lokalak aurkitzeko metodoak

2.1. Sarrera eta oinarrizko metodoak.

Oinarrizko definizioak

Definizioak

Izan bedi $a, b \in \mathbb{R}$ zein $a < b$ eta $f \in C^0([a, b])$. Orduan,

- $x \in [a, b]$ f-ren **erroa** da baldin eta soilik baldin $f(x) = 0$.
- $x \in [a, b]$ f-ren **puntu finkoa** da baldin eta soilik baldin $f(x) = x$.

Oinarrizko definizioak

Definizioak

Izan bedi $a, b \in \mathbb{R}$ zein $a < b$ eta $f \in C^0([a, b])$. Orduan,

- $x \in [a, b]$ f-ren **erroa** da baldin eta soilik baldin $f(x) = 0$.
- $x \in [a, b]$ f-ren **puntu finkoa** da baldin eta soilik baldin $f(x) = x$.

Adibidea

Izan bedi $f(x) = x^2 - 1$. Orduan $x = \pm 1$ f-ren erroak dira. Gainera, $x = \frac{1 \pm \sqrt{5}}{2}$ balioak f-ren puntu finkoak dira.

Oinarrizko definizioak

Definizioak

Izan bedi $a, b \in \mathbb{R}$ zein $a < b$ eta $f \in C^0([a, b])$. Orduan,

- $x \in [a, b]$ f-ren **erroa** da baldin eta soilik baldin $f(x) = 0$.
- $x \in [a, b]$ f-ren **puntu finkoa** da baldin eta soilik baldin $f(x) = x$.

Adibidea

Izan bedi $f(x) = x^2 - 1$. Orduan $x = \pm 1$ f-ren erroak dira. Gainera, $x = \frac{1 \pm \sqrt{5}}{2}$ balioak f-ren puntu finkoak dira.

Oharra

x f-ren puntu finkoa da baldin eta soilik baldin $x \mapsto f(x) - x$ funtzioaren erroa bada.

Metodo grafikoa

Definizioa

Metodo grafikoa da programa informatiko baten bidez grafikoa egitea, eta grafiko horri begira erroaren hurbilpen bat ematea.

Metodo grafikoa

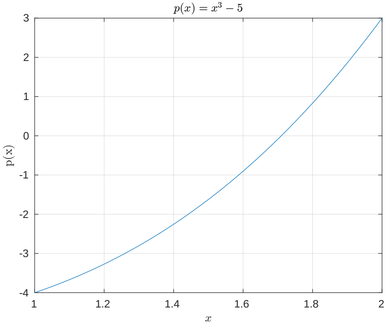
Definizioa

Metodo grafikoa da programa informatiko baten bidez grafikoa egitea, eta grafiko horri begira erroaren hurbilpen bat ematea.

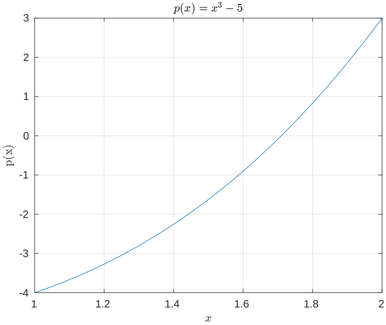
Oharra

Metodo grafikoak bi muga argi ditu: gizakion interbentzio zuzena eskatzen du, eta soilik aplika daiteke bi ekuazio edota aldagai ditugunean. Halere, ez gutxietsi metodo grafikoa. Oso erabilgarria da erro kopurua eta erroen hurbilpen bat lortzeko ekuazio batekin ari garenean.

Metodo grafikoaren adibidea



Metodo grafikoaren adibidea



Metodo grafikoa erabiliz, ikus daiteke 1.7 inguru $p(x) = x^3 - 5$ polinomioaren erroa dagoela.

Bisekzio metodoa

Bisekzio metodoa tarteko balioen teoreman oinarritzen den metodoa da. Idea hurrengoa da, $a, b \in \mathbb{R}$ bada zein $a < b$, $f \in C^0([a, b])$ bada zein $f(a)f(b) < 0$ eta $c = \frac{a+b}{2}$ bada, orduan hurrengo hiru kasu daude:

- $f(c) = 0$ bada, erro bat aurkitu dugu.
- $f(a)f(c) < 0$ bada, erro bat dugu $[a, c]$ tartean.
- $f(b)f(c) < 0$ bada, erro bat dugu $[c, b]$ tartean.

Bisekzio metodoaren pseudo-kodea

Bisekzio metodoaren pseudo-kodea

Sarrera: $a_0, b_0 \in \mathbb{R} : a_0 < b_0$;
Sarrera: $f \in C^0([a,b]) : f(a_0)f(b_0) < 0$
Sarrera: $n \in \mathbb{N}$

for $i := 0, ..., n-1$ **do**
  $c_i := \frac{a_i + b_i}{2}$
  if $f(a)f(c) \leq 0$ **then**
    $b_{i+1} := c_i, \quad a_{i+1} := a_i$
  else
    $a_{i+1} := c_i, \quad b_{i+1} := b_i$
  end if
end for

Irteera: a_n, b_n

Bisekzio metodoaren konbergentzia abiadura.

Teorema

Izan bedi $a_0, b_0 \in \mathbb{R} : a_0 < b_0$, $f \in C^0([a,b]) : f(a_0)f(b_0) < 0$ *eta* $n \in \mathbb{N}$. *Bisekzio metodoa aplikatu eta gero:*

$$b_n - a_n = \frac{b_0 - a_0}{2^n}$$

dugu. Gainera, existitzen da $x \in [a_n, b_n]$ *zein* $f(x) = 0$.

Teorema hau indukzioz froga daiteke i-n. Nahikoa da frogatzea edozein i baliorako

$$b_i - a_i = \frac{b_0 - a_0}{2^i}$$

dela, eta existitzen dela $x \in [a_i, b_i]$ zein $f(x) = 0$.

2.2. Definizio garrantzitsu batzuk.

Zenbakizko metodoetan erabiltzen den nomenklatura

Ez nahastu hurrengo nozioak:

- **Zenbakizko metodo** bat problema matematiko baten ebazpenaren hurbilpen bat, baldintza egokiak betetzen badira, ematen duten pausu garrantzitsuenak dira.
- **Algoritmo** bat problema bat ebazteko eman behar diren urratsak dira.
- **Pseudo-kodea** algoritmoaren edo metodoaren laburpen bat da, deskribatzen duena pausoz pausoz gizakion hizkuntzan, erraz uler dezan. Pseudo-kodeetan ez ohi da programatzeko pausu guztiak ematen, eta askatasun handia dago idazteko. Notazio matematikoa erabiltzen ohi da.
- **MATLAB-en inplementatzeko kodea**, bere izena dioenez, MATLAB-eko sintaxian idatzitako kodea da.

Nomenklaturaren adibidea

- Bisekzioaren metodoa errepikapen formulak dira.
- Bisekzioaren metodoaren pseudo-kodea irteera, errepikapen formula eta sarrera modu eskematikoan idatzita duen formula da.
- Bisekzioaren algoritmo bat litzateke pseudo-kodean idatzi dugun prozesua, hor adierazi baitugu noiz geldituko den.
- MATLAB-eko inplementazioa da algoritmo hori MATLAB-en sintaxian idatzita.

Hitz tekniko gehiago

Definizioa

Algoritmo batek behar duen informazio guztiari **sarrera** (*input*, ingelesez) deritzogu. Gainera, itzultzen dituen balioei **irteera** (*output*, ingelesez) deritzogu. Era berean, algoritmoak segida bat sortzen ari badu, sarrerako balioei **hasierako balioak** deritzogu, eta itzultzen dituen balioei **bukaerako balioak**.

Adibidea

Bisekzioaren metodorako definitu dugun algoritmoaren sarrera f funtzioa, n iterazio kopurua eta (a_0, b_0) balioak dira, (a_0, b_0) hasierako balioak, eta (a_n, b_n) bukaerako balioak eta irteera.

Algoritmoak gelditzeko irizpidea

Suposa dezagun algoritmo bat $\{x_n\}$ segida sortzen duela $\mathbb{R}^N$-n $F(x) = 0$ problema ebazteko. Hurrengo definizioak ditugu:

- Algoritmoaren **errorea** segidaren n. baliorako $\|F(x_n)\|$ da.
- Algoritmoaren **zehaztasun absolutua** segidaren n. baliorako $\|x_n - x_{n-1}\|$ da, eta **zehaztasun erlatiboa** segidaren n. baliorako $\frac{\|x_n - x_{n-1}\|}{\|x_{n-1}\|}$ da.
- $x_0, \ldots, x_{n-1}$ kalkulatu baditugu, x_n kalkulatzeari **iterazio** bat egitea deritzogu.

Errorea eta zehaztasuna algoritmoa gelditzeko irizpide ohikoak dira, iterazio kopuruarekin batera. Azken hau gomendagarria da begizta infinituetan ez sartzeko, hurrengo ataletan ikusiko dugun metodoekin gerta daitekeenez.

Metodoen sailkapena

- **Globalak:** hasierako balioak errotik urrun egon badaitezke.
- **Lokalak:** hasierako balioak errotik gertu egon behar badira.
- **Bakartze-metodoak:** erroa dagoen tartea asmatzen saiatzen badu.
- **Irekiak:** hurbiltze puntuak ematen baditu.

Adibidea

Bisekzioaren metodoa bakartze-metodo globala da. Hurrengo ataletan ikusiko ditugu metodoa lokalak eta irekiak.

Konbergentzia abiadura

Definizioak

Izen bedi x_n segida zein $x_n \to x^*$.

- Konbergentzia Q-**lineala** da existitzen bada $r \in [0, 1)$ eta $N \in \mathbb{N}$ non:

$$\frac{|x_{n+1} - x^*|}{|x_n - x^*|} \leq r, \quad \forall n \geq N.$$

  Gainera,

$$\frac{|x_{n+1} - x^*|}{|x_n - x^*|} \to 0$$

  bada, konbergentzia **Q-superlineala** da.

- Q-**konbergentzia ordena** $p > 1$ **da** existitzen bada $r \geq 0$ eta $N \in \mathbb{N}$ non:

$$\frac{|x_{n+1} - x^*|}{|x_n - x^*|^p} \leq r, \quad \forall n \geq N.$$

  Gainera, $p = 2$ kasuan **Q-koadratikoa** dela diogu.

Q letra "quotient" hitzetik dator.

Konbergentzia abiaduraren adibideak

Adibideak

- $x_n = 10^{-n}$ segidak linealki konbergitzen du 0-rantz.
- $y_n = 10^{-2^n}$ segidak koadratikoki konbergitzen du 0-rantz.

Konbergentzia abiaduraren adibideak

Adibideak

- $x_n = 10^{-n}$ segidak linealki konbergitzen du 0-rantz.
- $y_n = 10^{-2^n}$ segidak koadratikoki konbergitzen du 0-rantz.

Oharra

Konbergentzia koadratikoa lineala baina askoz azkarrago da. Hau barneratzeko kontuan izan dezakegu adibideetako segidek eginiko errorearen bornea adierazten badute, x_n esan nahiko luke iterazio bakoitzean zifra bat zehazten ari garela, eta y_n zehazki dakigun zifra kopurua bikoizten dela iterazio batekin.

2.3. Puntu finkoak lortzeko metodoak.

Puntu finkoen existentzia eta bakartasuna

Proposizioa

Izan bitez $a, b \in \mathbb{R}$ *zein* $a < b$ *eta* $f \in C^0([a,b];[a,b])$. *Orduan,*

1. f*-k puntu finko bat du* $[a,b]$ *tartean.*
2. *Gainera,* $f \in C^1([a,b])$ *bada eta* $|f'(x)| < 1$ *bada* $\forall x \in (a,b)$ *balioetarako, puntu finko hori bakarra da.*

Adibidea

Emaitza hau erabiliz froga daiteke kosinuak puntu finko bakarra duela $[0, \pi/2]$ tartean ($x = 0$ hain zuzen ere).

Puntu finkoen existentzia eta bakartasuna: frogapena

Frogapena.

1. 1. zatia frogatzeko nahikoa da $f(a) - a \geq 0$ eta $f(b) - b \leq 0$ dela konturatzea, eta tarteko balioen teorema erabiltzea.

Puntu finkoen existentzia eta bakartasuna: frogapena

Frogapena.

1 1. zatia frogatzeko nahikoa da $f(a) - a \geq 0$ eta $f(b) - b \leq 0$ dela konturatzea, eta tarteko balioen teorema erabiltzea.

2 2. zatia absurdora froga daiteke. Existitzen badira $x_1, x_2 \in [a, b]$ zein $x_1 < x_2$, $f(x_1) = x_1$ eta $f(x_2) = x_2$, existitzen da $c \in (x_1, x_2)$ zein

$$f'(c) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = 1,$$

absurdoa dena batez-besteko balioaren teoremagatik ("Kalkulu Diferentziala eta Integrala I" liburuaren 5.3.4 teorema).

□

Puntu-finkoak aurkitzeko metodo iteratiboa

Teorema (Puntu finkoaren teorema)

Izan bedi $a, b \in \mathbb{R}$ zein $a < b$ eta $f \in C^1([a,b];[a,b])$, eta $x^* \in [a,b]$ zein $f(x^*) = x^*$. Orduan,

1. $|f'(x^*)| < 1$ *bada, edozein* $\varepsilon \in (0, 1 - |f'(x^*)|)$ *baliorako existitzen da* $\delta > 0$ *zein* $x_0 \in (x^* - \delta, x^* + \delta)$ *bada, orduan:*
$$|f^n(x_0) - x^*| \leq (|f'(x^*)| + \varepsilon)^n |x_0 - x^*|.$$
2. $|f'(x^*)| \in (0,1)$ *bada, edozein* $\varepsilon \in (0, \min\{1 - |f'(x^*)|, |f'(x^*)|\})$ *baliorako existitzen da* $\delta > 0$ *zein* $x_0 \in (x^* - \delta, x^* + \delta)$ *bada, orduan:*
$$(|f'(x^*)| - \varepsilon)^n |x_0 - x^*| \leq |f^n(x_0) - x^*|.$$
3. $|f'(x^*)| > 1$ *bada, existitzen da* $\delta > 0$ *zein* $x_0 \in (x^* - \delta, x^* + \delta) \setminus \{x^*\}$ *bada, orduan existitzen da* $n \in \mathbb{N}$ *zein* $f^n(x_0) \notin (x^* - \delta, x^* + \delta)$.

Puntu-finkoak aurkitzeko metodo iteratiboa: oharrak

- Hemen $f^n(x_0)$ esan nahi du f funtzioa n aldiz konposatzen ari garela.

Puntu-finkoak aurkitzeko metodo iteratiboa: oharrak

- Hemen $f^n(x_0)$ esan nahi du f funtzioa n aldiz konposatzen ari garela.
- Lehenengo bi ataletan konbergitzearen arrazoi nagusia da $|f'(x^*)| < 1$ eta puntu finko bat egoteak inplikatzen duela **kontrakzio lokala** dugula; hots, irudien arteko distantzia murrizten dela aurreirudiekiko. Ohartu, teoremaren frogapenean kontrakzioa dela ikusiko dugu, eta kontrakzioa kuantifikatuko dugu.

Puntu-finkoak aurkitzeko metodo iteratiboa: oharrak

- Hemen $f^n(x_0)$ esan nahi du f funtzioa n aldiz konposatzen ari garela.
- Lehenengo bi ataletan konbergitzearen arrazoi nagusia da $|f'(x^*)| < 1$ eta puntu finko bat egoteak inplikatzen duela **kontrakzio lokala** dugula; hots, irudien arteko distantzia murrizten dela aurreirudiekiko. Ohartu, teoremaren frogapenean kontrakzioa dela ikusiko dugu, eta kontrakzioa kuantifikatuko dugu.
- Azkenengo atala ez konbergitzearen arrazoi nagusia $|f'(x^*)| > 1$ eta puntu finko bat egoteak inplikatzen duela **espantsio lokala** dugula; hots, irudien arteko distantzia handitzen dela aurreirudiekiko. Ohartu, teoremaren frogapenean handitzen dela ikusiko dugu, eta espantsioa kuantifikatuko dugu.

Puntu-finkoak aurkitzeko metodo iteratiboa: frogapena (i)

Frogapena. Froga dezagun 1. atala. Jarraitasunagatik, existitzen da $\delta > 0$ zein $|f'(x)| < |f'(x^*)| + \varepsilon$ dugun $I_\delta := (x^* - \delta, x^* + \delta)$ tartean. Orduan, indukzioz froga dezakegu $x_0 \in I_\delta$ bada edozein $n \in \mathbb{N}$-rako:

$$|f^n(x_0) - x| \leq (|f'(x^*)| + \varepsilon)^n |x - x_0|,$$

Horretarako, nahikoa da frogatzea edozein $x \in I_\delta$-baliorakorako:

$$|f(x) - x_0| \leq (|f'(x^*)| + \varepsilon)|x - x_0| \quad (1)$$

dugula, eta $x = f^{n-1}(x_0), f^{n-2}(x_0), ..., x_0$ balioekin aplikatzea.

Puntu-finkoak aurkitzeko metodo iteratiboa: frogapena (i)

Frogapena. Froga dezagun 1. atala. Jarraitasunagatik, existitzen da $\delta > 0$ zein $|f'(x)| < |f'(x^*)| + \varepsilon$ dugun $I_\delta := (x^* - \delta, x^* + \delta)$ tartean. Orduan, indukzioz froga dezakegu $x_0 \in I_\delta$ bada edozein $n \in \mathbb{N}$-rako:

$$|f^n(x_0) - x| \leq (|f'(x^*)| + \varepsilon)^n |x - x_0|,$$

Horretarako, nahikoa da frogatzea edozein $x \in I_\delta$-baliorakorako:

$$|f(x) - x_0| \leq (|f'(x^*)| + \varepsilon)|x - x_0| \tag{1}$$

dugula, eta $x = f^{n-1}(x_0), f^{n-2}(x_0), ..., x_0$ balioekin aplikatzea. (1) Barrow-ren erregela aplikatuz froga daiteke:

$$\frac{|f(x) - x^*|}{|x - x^*|} = \frac{\left|\int_{x^*}^{x} f'(s)ds\right|}{|x - x^*|} \leq |f'(x^*)| + \varepsilon.$$

Puntu-finkoak aurkitzeko metodo iteratiboa: frogapena (ii)

2. atala 1. atalaren analogoa da. Ezberdintasun bakarra da, I_δ definitzeko orduan behar dugula $(x^* - \delta, x^* + \delta)$ tartean: $|f'(x^*)| - \varepsilon < |f'(x)| < |f'(x^*)| + \varepsilon$ izatea.

Puntu-finkoak aurkitzeko metodo iteratiboa: frogapena (ii)

2. atala 1. atalaren analogoa da. Ezberdintasun bakarra da, I_δ definitzeko orduan behar dugula $(x^* - \delta, x^* + \delta)$ tartean: $|f'(x^*)| - \varepsilon < |f'(x)| < |f'(x^*)| + \varepsilon$ izatea.

Froga dezagun azkenik 3. atala. Izan bedi $K = (1 + |f'(x)|)/2$ eta $\delta > 0$ zein $|f'(y)| \geq K$ dugun $(x - \delta, x + \delta)$ tartean. Orduan, lehen bezala froga dezakegu $x_0, \ldots, f^i(x_0) \in (x - \delta, x + \delta)$ bada,

$$\left|f^j(x_0) - x\right| \geq K^j |x - x_0| \;\; \forall j = 1, ..., i.$$

Puntu-finkoak aurkitzeko metodo iteratiboa: frogapena (ii)

2. atala 1. atalaren analogoa da. Ezberdintasun bakarra da, I_δ definitzeko orduan behar dugula $(x^* - \delta, x^* + \delta)$ tartean: $|f'(x^*)| - \varepsilon < |f'(x)| < |f'(x^*)| + \varepsilon$ izatea.

Froga dezagun azkenik 3. atala. Izan bedi $K = (1 + |f'(x)|)/2$ eta $\delta > 0$ zein $|f'(y)| \geq K$ dugun $(x - \delta, x + \delta)$ tartean. Orduan, lehen bezala froga dezakegu $x_0, \ldots, f^i(x_0) \in (x - \delta, x + \delta)$ bada,

$$\left|f^j(x_0) - x\right| \geq K^j |x - x_0| \;\; \forall j = 1, ..., i.$$

Hortaz, kontuan izanik $K > 1$ dela, existitzen da $n \in \mathbb{N}$ zein $f^n(x_0) \notin (x - \delta, x + \delta)$. $\square$

Puntu finkoa aurkitzeko algoritmoa

Puntu finkoaren teoremari loturiko metodoaren pseudo-kodea

Sarrera: f funtzioa, x puntu finkoa, x_0 balioa, eta $\varepsilon > 0$.
1: $n := 0$
2: **while** $|f(x_n) - x_n| \geq \varepsilon$ **do**
3: $\quad x_{n+1} := f(x_n)$
4: $\quad n := n + 1$
5: **end while**
Irteera: x_n

Oharra

Algoritmoa inplementatzeko orduan, gomendagarria da gehitzea iterazio kopuruagatik eta zehaztasunagatik gelditzea.

2.4. Newton-Raphson-en metodoa.

Newton-Raphson-en metodoa motibatzen

Izan bedi f funtzioa bi aldiz deribagarria. Taylor-ren x_0 puntuan zentratutako 1. mailako polinomioa erabiliz:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + R_{1,x_0}[f](x).$$

Hortaz, $f(x) = 0$ berdin bada, hondarra kenduz:

$$0 \approx f(x_0) + f'(x_0)(x - x_0);$$

hau da,

$$x \approx x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Newton-Raphson-en metodoa

Teorema

Izan bedi $f \in C^2([a,b])$ *eta* $x^* \in (a,b)$ *zein* $f(x^*) = 0$ *eta* $\min_{[a,b]} |f'| > 0$. *Izan bedi:*

$$g(x) = x - \frac{f(x)}{f'(x)}, \quad K = \frac{|f''(x^*)|}{2|f'(x^*)|}.$$

Orduan $\varepsilon > 0$ *behar bezain txikia bada, existitzen dira* $\delta > 0$ *eta* $x_0 \in (x^* - \delta, x^* + \delta)$ *zein:*

$$\frac{|g^{n+1}(x_0) - x^*|}{|g^n(x_0) - x^*|^2} \in (K - \varepsilon, K + \varepsilon) \ \ \forall n \in \mathbb{N},$$

eta $g^n(x_0) \to x^*$.

Newton-Raphson-en inplementazioa eta adibideak

- Ordenagailu praktiketan inplementatuko dugu Newton-Raphson-en metodoa MATLAB-en bidez, eta abian ikusiko dugu.
- Ordenagailu praktiketan ikusiko dugu zer gertatzen den hipotesiak betetzen ez direnean.
- Ariketa orduan urrezko zenbakia Newton-Raphson-en metodoaren bidez kalkulatzean iterazio bakoitzean egiten den errorea kalkulatuko dugu. Gainera, $\sqrt{5}$ erro funtzioa Taylorren polinomioen bidez kalkulatzean iterazio bakoitzean egiten genuen errorearekin konparatuko dugu.

Newton-Raphson-en metodoaren frogapena (i)

1. urratsa: berdintza bat edozein $x \in [a, b]$ baliorako. Izan bedi $x \in [a, b]$. Orduan,

$$\begin{aligned} g(x) - x^* &= x - x^* - \frac{f(x)}{f'(x)} \\ &= \frac{-f(x) - f'(x)(x^* - x)}{f'(x)} \\ &= \frac{R_{1,x}[f](x^*)}{f'(x)} \\ &= \frac{\int_x^{x^*} f''(s)(x^* - s)ds}{f'(x)}. \end{aligned}$$

Newton-Raphson-en metodoaren frogapena (ii)

2. urratsa: ezberdintza lortzen. Jarraitasuna erabiliz, erraz froga daiteke existitzen dela $\delta_1 > 0$ zein edozein $x \in (x^* - \delta, x^* + \delta)$ baliorako:

$$(K - \varepsilon)|x - x^*|^2 \leq |g(x) - x^*| \leq (K + \varepsilon)|x - x^*|^2.$$

3. urratsa: konbergentzia. Azkenik, $\delta < \min\left\{\delta_1, \frac{1}{2(K+\varepsilon)}\right\}$ hautatuz,

$$|g(x) - x^*| \leq \frac{|x - x^*|}{2}$$

dugu edozein $x \in (x^* - \delta, x^* + \delta)$ baliorako. Ondorioz, edozein $x_0 \in (x^* - \delta, x^* + \delta)$ hasierako baliorako froga daiteke indukzioz:

$$|g^n(x_0) - x^*| \leq \frac{|x_0 - x^*|}{2^n};$$

hau da, $g^n(x_0) \to x^*$. $\square$

Newton-Raphson-en metodoaren pseudo-kodea:

Newton-Raphson-en metodoaren pseudo-kodea

Sarrera: f funtzioa, x puntu finkoa, x_0 balioa, eta $\varepsilon > 0$.

1: $n := 0$
2: **while** $|f(x_n)| \geq \varepsilon$ **do**
3: $\quad x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$
4: $\quad n := n + 1$
5: **end while**

Irteera: x_n

Oharra

Algoritmoa inplementatzeko orduan, iterazio kopuru maximoagatik eta zehaztasunagatik ere gelditzea gomendagarria da.

Newton-Raphson-en metodoa deribatua ezezaguna denean

Posible da Newton-Raphson-en metodoa egokitzea deribatua ezagutzen ez denean, metodo berri bat sortuz: **ebakitzailearen metodoa**. Kasu horretan

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$$

iterazioa dugu (eta x_0, x_1 aukeratu behar ditugu). Hots, kasu horretan erabili dugu:

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

2.5. Optimizazio problemak eta gradientearen jaitsieraren metodoa (*gradient descent method*).

Optimizazio problemak eta erroen bilaketa problemen arteko lotura.

Proposizioa

Izan bedi $a, b \in \mathbb{R}$ *zein* $a < b$, $f \in C^1([a, b])$, $x \in (a, b)$. *Orduan,* f *funtzioak maximo edo minimo lokala badu* x *puntuan:*

$$f'(x) = 0.$$

Oharra

Proposizio hori erabiliz, argi dago maximo edo minimo lokalak f'-ren erroak direla; hortaz, $f \in C^1$ bada, aurreko ataletako metodoak erabili ditzakegu.

Gradientearen jaitsieraren metodoa minimoak aurkitzeko

Gradientearen jaitsieraren metodoa (edo **gradientearen metodoa**, testuingurunetik argi ulertzen denean), metodo dinamiko bat da zeinetan pausoz pauso minimora hurbiltzen garen. Hurrengo printzipioetan oinarritzen da:

- Minimoa aurkitzeko deribatuaren kontrako norabidean mugitu behar gara.
- Pausu bakoitzean mugitzen garen kantitatea deribatuaren proportzionala izan behar da.

Iterazioa

$$y = x - \alpha f'(x)$$

izango da, non α parametro txiki bat den.

Gradientearen jaitsieraren metodoa: formalismo matematikoa

Teorema (Gradientearen jaitsieraren metodoa)

Izan bedi $f \in C^2([a,b])$ *eta* $x^* \in (a,b)$ *zein* f*-k minimo lokala duen* x^* *puntuan eta* $f''(x^*) > 0$. *Izan bedi* $\alpha \in \left(0, \frac{1}{f''(x^*)}\right)$, *eta*

$$g_\alpha(x) = x - \alpha f'(x).$$

Orduan, $\varepsilon > 0$ *balioa behar bezain txikia bada, existitzen da* $\delta > 0$ *zein* $x_0 \in (x^* - \delta, x^* + \delta)$ *bada:*

$$(1 - \varepsilon - \alpha f''(x^*))^n |x_0 - x^*| \leq |g_\alpha^n(x_0) - x^*| \leq (1 + \varepsilon - \alpha f''(x^*))^n |x_0 - x^*|.$$

Definizioa

Gradientearen metodoan α parametroari **pausuaren tamaina** edo **ikasketa-ratioa** (*learning rate*) deritzogu.

Gradientearen jaitsieraren metodoa: frogapena

Frogapena.

Kontuan izanik

$$g_\alpha(x) = x \Leftrightarrow f'(x) = 0,$$

eta x^*-ren ingurune batean f'-ren erro bakarra x^* dela, f minimo lokala du baldin eta soilik baldin g_α-ren puntu finkoa bada. Gainera, $g'_\alpha(x^*) = 1 - \alpha f''(x^*) \in (0,1)$. Hortaz, Puntu Finkoaren Teoremaren ondorio zuzena da. $\square$

Gradientearen jaitsieraren metodoaren pseudo-kodea

Gradientearen jaitsieraren metodoa

Sarrera: f funtzioa, x_0 balioa, $\alpha > 0$ eta $\varepsilon > 0$.

1: $n := 0$
2: **while** $|f'(x_n)| \geq \varepsilon$ **do**
3: $\quad x_{n+1} := x_n - \alpha f'(x_n)$
4: $\quad n := n + 1$
5: **end while**

Irteera: x_n

Oharra

Algoritmoa inplementatzeko orduan, iterazio kopuruagatik eta zehaztasunagatik gelditzea ere komeni da. Gainera, iterazio bakoitzean lortutako errorea $|f'(x)|$ da, ez $|f(x)|$.

Gradientearen igoeraren metodoa maximoak kalkulatzeko

Optimizazioaren inguruko teoria soilik egiten da minimoen kalkuluetarako f-ren maximo lokala kalkulatzeko nahikoa baita $-f$-ren minimo lokala kalkulatzea. Kasu honetan ez da salbuespena, aurreko emaitzak eta algoritmoak aplika baitezakegu $-f$-rako.

Gradientearen jaitsieraren metodoaren aplikazioak

- Gradientearen metodoa erraz orokor daiteke $\mathbb{R}^n$-n edo espazio funtzionaletan definitutako optimizazio problemak ebazteko. Gainera, edozein dimentsiotan hurbilpen ordena 1. dimentsioko ordenaren berdina da (beste kontu bat da gradientearen kalkulua).

Gradientearen jaitsieraren metodoaren aplikazioak

- Gradientearen metodoa erraz orokor daiteke $\mathbb{R}^n$-n edo espazio funtzionaletan definitutako optimizazio problemak ebazteko. Gainera, edozein dimentsiotan hurbilpen ordena 1. dimentsioko ordenaren berdina da (beste kontu bat da gradientearen kalkulua).
- Ikasketa automatikoko (machine learning) algoritmo asko gradientearen maximoan oinarritzen dira, askotan hipotesi asko dugu eta “koste funtzioa” minimizatzen duen hipotesia hautatzeko gradientearen metodoa erabiltzen da. Ordenagailu praktika bat dago honen inguruan. Ikasketa-ratio terminoa, adibidez, ikasketa automatikotik dator.