HÁSKÓLI ÍSLANDS
VERKFRÆÐI- OG NÁTTÚRUVÍSINDASVIÐ
RAUNVÍSINDADEILD

Töluleg greining (STÆ405G)

Útgáfa 0.9

Benedikt Steinar Magnússon

April 14, 2016

KAFLI 1

Inngangur

He was determined to discover the underlying logic behind the universe. Which was going to be hard, because there wasn't one.

- Terry Pratchett, Mort

1.1 Hvað er töluleg greining?

1.1.1 Tilraun að svari

- Fagið *töluleg greining* snýst um að búa til, greina og forrita aðferðir til þess að nálga á lausnum á stærðfræðilegum verkefnum.
- Aðferðirnar eru settar fram með reikniritum sem síðan eru forrituð og það þarf góðan skilning á eiginleikum lausnanna sem verið er að nálga til þess að geta greint hvernig forritin munu virka.
- Greining á reikniritum er aðallega fólgin í skekkjumati og mati á þeim aðgerðafjölda sem þarf til þess að ná að nálga lausn með fyrirfram gefinni nákvæmni, þ.e. hagkvæmni og nákvæmni reikniritsins.
- Líkanagerð í raunvísindum og verkfræði felur yfirleitt í sér eftirfarandi skref:
  1. *Greina* kerfið sem um ræðir
  2. *Smíða líkan* sem útskýrir hvernig kerfið hegðar sér, þó yfirleitt með töluverðum einföldunum.
  3. *Herma* kerfið í tölvu eins vel og hægt er. Hér þarf að ná ásættanlegri námkvæmni á þeim tíma sem útreikningar mega taka.
  4. *Túlka* niðurstöðurnar og bera saman við upphaflega kerfið.

  Töluleg greining kemur mikið við sögu í lið 3. og einnig í lið 4.

1.2 Dæmi: Eldflaug

Gerum ráð fyrir að við höfum eftirfarandi eldflaug undir höndum:

- Eldsneytið dugir í 18 sek., þ.e. $t \in [0, 18]$.
- Loftmótstaðan er $d = 0.1v^2$, þar sem $v(t)$ er hraðinn á tíma t.
- Krafturinn sem knýr flaugina er $T = 5000$ N.
- Massi eldsneytisins er $m = 180 - 10t$ kg.

- Massi flaugarinnar er $M = 120 + m = 300 - 10t$ kg.

Spurningin er: Í hvaða hæð er eldflaugin þegar eldsneytið klárast?

Úr öðru lögmáli Newtons fæst að $F = (Mv)'$. Kraftarnir sem verka á eldflaugina er T upp á við og loftmótstaðan og þyngdarkrafturinn niður á við. Þannig fæst

$$(Mv)' = F = T - Mg - d$$

það er

$$M'v + Mv' = T - Mg - d.$$

Þetta jafngildir því að

$$v' = \frac{T - Mg - d - M'v}{M} = \frac{5000 - (300 - 10t)g - 0,1v^2 + 10v}{300 - 10t},$$

og upphafsskilyrðin eru $v(0) = 0$.

Þar sem $h' = v$, þá er hæðin á tíma t gefin með $h(t) = \int_0^t v(s)\,ds$. Þegar eldsneytið klárast þá er hæðin $h(18) = \int_0^{18} v(s)\,ds$.

Verkefnið er því að finna v, og reikna svo heildið.

Diffurjafnan hér að ofan er ólínuleg og ekki aðgreinanleg þannig að við getum ekki vænst þess finna lausn með þeim aðferðum sem við höfum þegar lært. Eins er ekki víst að við getum auðveldlega fundið stofnfall h fyrir v til þess að reikna heildið, jafnvel þótt við hefðum v.

Hins vegar getum við leyst diffurjöfnuna tölulega með aðferðunum úr *kafla 6*, og heildið reiknum við svo tölulega með aðferðunum úr *kafla 5*.

1.3 Samleitni runa

1.3.1 Nokkur atriði um samleitni runa

Mörg reiknirit til nálgunar á einhverri rauntölu eru hönnuð þannig að reiknuð er runa $x_0, x_1, x_2, \dots$ sem á að nálgast lausnina okkar.

1.3.2 Skilgreining: Samleitni

Rauntalnaruna (x_n) er sögð vera *samleitin* (e. convergent) að *markgildinu* r ef um sérhvert $\varepsilon > 0$ gildir að til er $N > 0$ þannig að

$$|x_n - r| < \varepsilon, \qquad \text{ef} \quad n \geq N.$$

Þetta er táknað annað hvort með

$$\lim_{n\to\infty} x_n = r \qquad \text{eða} \qquad x_n \to r \text{ ef } n \to \infty.$$

Ef runan (x_n) er samleitin að markgildinu r þá segjum við einnig að hún *stefni á* r.

Hugsum okkur nú að (x_n) sé gefin runa sem stefnir á r og táknum skekkjuna með $e_n = r - x_n$.

Runan er sögð vera *línulega samleitin* (e. linear convergence) ef til er $\lambda \in]0,1[$ þannig að

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = \lambda,$$

ofurlínulega samleitin (e. superlinear convergence), ef

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = 0,$$

ferningssamleitin (e. quadratic convergence) ef til er $\lambda > 0$ þannig að

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} = \lambda,$$

og *samleitin af stigi* α (e. convergence of order α), þar sem $\alpha > 1$, ef til er $\lambda > 0$ þannig að

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^\alpha} = \lambda.$$

Athugasemd: Runa er ofurlínulega samleitin ef hún er samleitin af stigi $\alpha > 1$.

Ferningssamleitin runa er samleitin af stigi 2 þannig að hún er einnig ofurlínulega samleitin.

1.3.3 Skilgreining

Oft eru notuð veikari hugtök til þess að lýsa samleitni runa (t.d. ef við getum ekki fundið λ og α nákvæmlega).

Þannig segjum við að runan (x_n) sé *að minnsta kosti línulega samleitin* ef til er $\lambda \in]0,1[$ og $N > 0$ þannig að

$$|e_{n+1}| \le \lambda |e_n|, \qquad n \ge N,$$

ef til er $\lambda > 0$ og $N > 0$ þannig að

$$|e_{n+1}| \le \lambda |e_n|^2, \qquad n \ge N,$$

og *að minnsta kosti samleitin af stigi* α, þar sem $\alpha > 1$, ef til eru $\lambda > 0$ og $N > 0$ þannig að

$$|e_{n+1}| \le \lambda |e_n|^\alpha, \qquad n \ge N.$$

1.4 Setning Taylors

Sometimes it's better to light a flamethrower than curse the darkness. - Terry Pratchett, Men at Arms: The Play

1.4.1 Ritháttur fyrir diffranleg föll

Látum nú $f : I \to \mathbb{C}$ vera fall á bili I sem tekur gildi í tvinntölunum. Ef f er deildanlegt í sérhverjum punkti í I, þá táknum við afleiðuna með f'. Ef f' er deildanlegt í sérhverjum punkti í I, þá táknum við *aðra afleiðu* f með f'', og svo framvegis.

Við skilgreinum með þrepun $f^{(k)}$ fyrir $k = 0, 1, 2, \ldots$ þannig að $f^{(0)} = f$ og ef $f^{(k-1)}$ er deildanlegt í sérhverjum punkti í I, þá er $f^{(k)} = (f^{(k-1)})'$.

Við látum $C^k(I)$ tákna línulega rúmið sem samanstendur af öllum föllum $f : I \to \mathbb{C}$ þannig að $f', \ldots, f^{(k)}$ eru til í sérhverjum punkti í I og $f^{(k)}$ er samfellt fall á I.

1.4.2 Nálgun með Taylor-margliðu

Ef $a \in I$, m er jákvæð heiltala og $f \in C^m(I)$, þá nefnist margliðan

$$p(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(m)}(a)}{m!}(x-a)^m$$

Taylor-margliða fallsins f í punktinum a af stigi m, og er stundum táknuð með $T_m f(x;a)$.

Athugið að stig margliðunnar p er minna eða jafnt og m.

1.4.3 Setning Taylors

Látum $I \subseteq \mathbb{R}$ vera bil, $f : I \to \mathbb{C}$ vera fall, $m \geq 0$ vera heiltölu og gerum ráð fyrir að $f \in C^m(I)$ og að $f^{(m+1)}(x)$ sé til í sérhverjum innri punkti bilsins I. Þá er til punktur ξ á milli a og x þannig að

$$f(x) - T_m f(x;a) = \frac{f^{(m+1)}(\xi)}{(m+1)!}(x-a)^{m+1}.$$

Hægri hliðin er oft táknuð $R_m(x)$.

Athugasemd: Þetta þýðir að skekkjan í því að nálga fallið $f(x)$ með Taylor-margliðu af stigi m hagar sér eins og $(x-a)^{m+1}$.

1.4.4 Viðbót

Ef $f^{(m+1)}$ er samfellt á lokaða bilinu með endapunkta a og x, þá er

$$\begin{aligned} R_m(x) &= f(x) - T_m f(x;a) \\ &= \int_a^x \frac{(x-t)^m}{m!} f^{(m+1)}(t)dt \\ &= (x-a)^{m+1} \int_0^1 \frac{(1-s)^m}{m!} f^{(m+1)}(a + s(x-a))ds. \end{aligned}$$

1.4.5 Sýnidæmi: Nálgun á fallgildum $x - \sin x$

Vitum að $x \approx \sin x$ ef x er lítið. Tökum $x = 0.1$ og hugsum okkur að við séum að reikna á vél með 8 stafa nákvæmni. Hún gefur

$$\sin 0.1 = 0.099833417$$

Af því leiðir

$$0.1 - \sin 0.1 = 1.66583 \cdot 10^{-4}$$

Við höfum tapað tveimur markverðum stöfum í nákvæmni.

Ef við notum Taylor-nálgunina fyrir $\sin(x)$,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \cdots$$

og tökum fyrstu þrjá liðina, þ.e. skoðum 6. stigs Taylor-margliðu fallsins.

$x - \sin(x)$ er þá u.þ.b.

$$x - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!}\right) = \frac{x^3}{3!} - \frac{x^5}{5!}.$$

Fallgildið er þá

$$\frac{0.1^3}{3!} - \frac{0.1^5}{5!} = 1.6658334 \cdot 10^{-4}.$$

Skekkjan er gefin með

$$|R_6(0.1)| = \left|\frac{\sin^{(7)}(\xi)}{7!}0.1^7\right| = \left|\frac{-\cos(\xi)}{7!}0.1^7\right| \leq \frac{1}{7!}0.1^7 < 0.2 \cdot 10^{-10}.$$

Sem þýðir að allir 8 stafir reiknivélarinnar eru markverðir, þ.e. allir stafir $1.6658334 \cdot 10^{-4}$ eru réttir.

$\sin^{(7)}$ hér að ofan táknar 7. afleiðu $\sin$, sem er $-\cos$.

Ef við tökum $x = 0.01$ er þetta enn greinilegra. Reiknivélin gefur

$$\sin(0.01) = 0.0099998333$$

Þannig að

$$0.01 - \sin 0.01 = 0.1667 \cdot 10^{-7}$$

og við erum bara með 4 markverða stafi.

Hér dugir að taka aðeins þriðja stigs liðinn í Taylor-formúlunni

$$0.01 - \sin(0.01) \approx \frac{0.01^3}{3!} = 0.16666667 \cdot 10^{-7},$$

því skekkjan er

$$R_4(0.01) \leq \frac{0.01^5}{5!} < 10^{-12}$$

1.5 Skekkjur

Við allar úrlausnir á verkefnum í tölulegri greininingu þarf að fást við skekkjur. Þær eru af ýmsum toga:

- Gögn eru oft niðurstöður mælinga og þá fylgja þeim *mæliskekkjur*. Eins getum við þurft að notast við nálganir á föstum sem koma fyrir (t.d. π, Avogadrosar talan, ...).
- Við nálganir á lausnum á stærðfræðilegum verkefnum verða til *aðferðarskekkjur*. Þær verða til þegar reikniritin eru hönnuð og greining á reikniritum snýst fyrst og fremst um mat á aðferðarskekkjum.
- *Reikningsskekkjur* verða til í tölvum á öllum stigum, jafnvel þegar tölur eru lesnar inn í tugakerfi og þeim snúið yfir í tvíundarkerfi. Þær verða líka til vegna þess að tölvur geta einungis unnið með endanlegt mengi af tölum og allar útkomur þarf að nálga innan þess mengis. Þessar skekkjur nefnast oft *afrúningsskekkjur*.
- *Mannlegar villur* eru óumflýjanlegar. Það sem við getum gert er temja okkur vinnubrögð sem lágmarka líkur á þeim og auðvelda okkur að finna villur sem við gerum.

  Real stupidity beats artificial intelligence every time. – Terry Pratchett

1.5.1 Skekkja í nálgun á rauntölu r

Við getum stillt upp jöfnunum svona

$$r \text{ (rétt gildi) } = x \text{ (nálgunargildi)} + e \text{ (skekkja)}$$

þar sem talan x er nálgun á tölunni r, og þá nefnist

$$e = r - x$$

skekkjan (e. error) í nálgun á r með x eða bara *skekkja*.

Algildi skekkju (e. absolute error) er tölugildið $|e| = |r - x|$

Ef vitað er að $r \neq 0$, þá nefnist

$$\frac{|e|}{|r|} = \frac{|r-x|}{|r|}$$

hlutfallsleg skekkja (e. relative error) í nálgun á r með x.

Aðvörun: Auðvitað er talan r sem við leitum að óþekkt (annars þyrftum við ekki að framkvæma alla þessa reikninga), sem þýðir að við getum hvergi notað hana í reikningum.

1.5.2 Fyrirframmat á skekkju

Metið er áður en reikningar hefjast hversu umfangsmikla reikninga þarf að framkvæma til þess að nálgunin náist innan fyrirfram gefinna skekkjumarka.

Ef lausnin er fundin með ítrekunaraðferð er yfirleitt metið hversu margar ítrekarnir þarf til þess að nálgun verði innan skekkjumarka.

1.5.3 Eftirámat á skekkju

Um leið og reikningar eru framkvæmdir er lagt mat á skekkju og reikningum er hætt þegar matið segir að nálgun sé innan skekkjumarka. Það gerist yfirleitt þegar gildið sem við reiknum út breytist orðið lítið í hverju skrefi.

Hér þarf að skipta í tvö tilvik, fyrst skoðum við tilvikið þegar runan er ofurlínulega samleitin og seinna tilvikið er þegar við vitum aðeins að runan er línulega samleitin, en þá er matið aðeins flóknara.

1.5.4 Ofurlínuleg samleitni – Eftirámat á skekkju

Hugsum okkur að við séum að nálga töluna r með gildum rununnar x_n, að við höfum reiknað út $x_0, \ldots, x_n$ og viljum fá mat á skekkjunni $e_n = r - x_n$ í n-ta skrefi.

Við reiknum næst út x_{n+1} og skrifum $e_{n+1} = \lambda_n e_n$. Þá er

$$x_{n+1} - x_n = (r - x_n) - (r - x_{n+1}) = e_n - e_{n+1} = (1 - \lambda_n)e_n$$

og við fáum

$$e_n = \frac{x_{n+1} - x_n}{1 - \lambda_n}.$$

Ef við vitum að runan er *ofurlínulega samleitin*, þá stefnir λ_n á 0 og þar með er

$$e_n \approx x_{n+1} - x_n.$$

Við hættum því útreikningi þegar $|x_{n+1} - x_n| < \varepsilon$ þar sem ε er fyrirfram gefin tala, sem lýsir þeirri nákvæmni sem við viljum ná.

1.5.5 Línuleg samleitni – Eftirámat á skekkju

Skoðum nú tilvikið ef einu upplýsingarnar sem við höfum er að runan x_n sé *að minnsta kosti línulega samleitin*, þ.e. $c \in [0, 1)$ og $N \in \mathbb{N}$ þannig að

$$|e_{n+1}| \leq c|e_n|, \qquad \text{fyrir } n \geq N.$$

Þá stefnir $\lambda_n = e_{n+1}/e_n$ á fasta $\lambda \leq c$ og við höfum

$$\lambda_n = \frac{e_{n+1}}{e_n} = \frac{1-\lambda_n}{1-\lambda_{n+1}} \cdot \frac{x_{n+2}-x_{n+1}}{x_{n+1}-x_n} \approx \frac{x_{n+2}-x_{n+1}}{x_{n+1}-x_n}$$

Nú þurfum við að átta okkur á því hvernig þetta er nýtt í útreikningum.

Hugsum okkur að við höfum reiknað út $x_0, \ldots, x_n$ og viljum fá mat á e_n. Við reiknum þá út x_{n+1} og x_{n+2} og síðan hlutfallið $\kappa_n = (x_{n+2} - x_{n+1})/(x_{n+1} - x_n)$ sem við notum sem mat á λ_n. Eftirámatið á skekkjunni í ítrekunarskrefi númer n verður síðan

$$e_n \approx \frac{x_{n+1}-x_n}{1-\kappa_n}.$$

Ef stærðin í hægri hliðinni er komin niður fyrir fyrirfram gefin skekkjumörk ε, þá stöðvum við útreikningana.

1.5.6 Sýnidæmi

Okkur er gefin runa af nálgunum á lausn jöfnunnar

$$f(x) = e^x \sin x - x^2 = 0$$

og eigum að staðfesta hvort nálgunaraðferðin er ferningssamleitin:

n	x_n	$\|x_{n+1} - x_n\|$	$\frac{\|x_{n+1}-x_n\|}{\|x_n-x_{n-1}\|^2}$
0	3.00000000000000		
1	2.73251570951922	0.10052257507862	1.404
2	2.63199313444060	0.01373904283351	1.359
3	2.61825409160709	0.00024006192208	1.273
4	2.61801402968501	0.00000007236005	1.256
5	2.61801395732496	0.00000000000001	1.272

Við metum $e_n \approx |x_{n+1} - x_n|$ og þar af leiðandi er

$$|e_n|/|e_{n-1}|^2 \approx |x_{n+1} - x_n|/|x_n - x_{n-1}|^2.$$

Við sjáum að hlutfallið $|x_{n+1} - x_n|/|x_n - x_{n-1}|^2$ helst stöðugt og því ályktum við að aðferðin sé ferningssamleitin.

1.5.7 Útreikningur á samleitnistigi

Skoðum lítið dæmi um útreikninga á samleitnistigi.

Eftirfarandi runa stefnir á $\sqrt{3}$.

n	x_n
0	2.000000000000000
1	1.666666666666667
2	1.727272727272727
3	1.732142857142857
4	1.732050680431722
5	1.732050807565499

Er samleitnistigið 1.618?

Ef ekki, hvert er þá samleitnistigið?

Ef miðað er við að runan (x_n) sé ofurlínulega samleitin, þá er eðlilegt að taka $e_n \approx x_{n+1} - x_n$ sem mat á skekkjunni $e_n = \sqrt{3} - x_n$ í n-ta ítrekunarskrefinu.

Við byrjum á því að kanna hvernig tilgátan um að samleitnistigið kemur út á þessum tölum með $e_n = x_{n+1} - x_n$:

n	x_n	$\lvert e_n\rvert$	$\lvert e_n\rvert/\lvert e_{n-1}\rvert^{1.618}$
0	2.000000000000000	$3.3333 \cdot 10^{-1}$	
1	1.666666666666667	$6.0606 \cdot 10^{-2}$	$3.5851 \cdot 10^{-1}$
2	1.727272727272727	$4.8701 \cdot 10^{-3}$	$4.5439 \cdot 10^{-1}$
3	1.732142857142857	$9.2177 \cdot 10^{-5}$	$5.0837 \cdot 10^{-1}$
4	1.732050680431722	$1.2713 \cdot 10^{-7}$	$4.3004 \cdot 10^{-1}$
5	1.732050807565499		

Tveimur síðustu tölunum í aftasta dálki ber ekki nógu vel saman, svo það er vafasamt hvort talan 1.618 er rétta samleitnistigið.

Ef (x_n) er samleitin af stigi α, þá gildir $\lim_{n\to\infty} |e_{n+1}|/|e_n|^\alpha = \lambda$, þar sem $\lambda > 0$. Þar með höfum við nálgunarjöfnu ef n er nógu stórt,

$$\frac{|e_{n+1}|}{|e_n|^\alpha} \approx \frac{|e_{n+2}|}{|e_{n+1}|^\alpha} \qquad \text{þá og því aðeins að} \qquad \frac{|e_{n+1}|}{|e_{n+2}|} \approx \left|\frac{e_n}{e_{n+1}}\right|^\alpha.$$

Ef við lítum á þetta sem jöfnu og leysum út α, þá fáum við

$$\alpha_n = \frac{\ln(|e_{n+1}|/|e_{n+2}|)}{\ln(|e_n|/|e_{n+1}|)}.$$

Við getum reiknað út þrjú gildi á α úr þeim gögnum sem við höfum, $\alpha_0 = 1.479$, $\alpha_1 = 1.573$ og $\alpha_2 = 1.660$.

Ef við endurtökum útreikninga okkar hér að framan með 1.660 í stað 1.618, þá fæst

n	p_n	$\lvert e_n\rvert$	$\lvert e_n\rvert/\lvert e_{n-1}\rvert^{1.660}$
0	2.000000000000000	$3.3333 \cdot 10^{-1}$	
1	1.666666666666667	$6.0606 \cdot 10^{-2}$	$3.7551 \cdot 10^{-1}$
2	1.727272727272727	$4.8701 \cdot 10^{-3}$	$5.1143 \cdot 10^{-1}$
3	1.732142857142857	$9.2177 \cdot 10^{-5}$	$6.3639 \cdot 10^{-1}$
4	1.732050680431722	$1.2713 \cdot 10^{-7}$	$6.3639 \cdot 10^{-1}$
5	1.732050807565499		

Tölunum neðst í aftasta dálki ber saman með fimm réttum stöfum og því ályktum við að 1.660 sé nær því að vera rétta samleitnistigið.

1.6 Meira um skekkjur

1.6.1 Skilgreining: Markverðir stafir

Gerum ráð fyrir að $r \neq 0$, þá segjum við að x sé *nálgun á r með t markverðum stöfum (e. significant digits)* ef

$$\frac{|r-x|}{|r|} \leq 10^{-t}.$$

Getum útfært þetta aðeins ítarlegra. Ef

$$10^{-(t+1)} < \frac{|r-x|}{|r|} \leq 10^{-t}.$$

þá segjum við að nálgunin á r með x sé rétt með að minnsta kosti t markverðum stöfum og að hámarki með $t+1$ markverðum stöfum.

Athugið að ef e er minnsta heila talan þannig að $|r| < 10^e$, þá gefur seinni ójafnan matið

$$|r - x| = 0.0 \dots 0 a_t a_{t+1} \dots \cdot 10^e,$$

þar sem núllin aftan við punkt eru t talsins.

Einnig er hægt að útfæra þetta fyrir aðrar grunntölur en 10.

1.6.2 Úrlausn annars stigs jöfnu

Þegar núllstöðvar annars stigs jöfnunnar $ax^2 + bx + c = 0$ eru reiknaðar út úr formúlunni

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

verður til styttingarskekkja ef b^2 er miklu stærra heldur en $4ac$ vegna $|b| \approx \sqrt{b^2 - 4ac}$. Við komumst hjá þessum vandræðum með því að líta á margliðuna fullþáttaða $a(x - x_1)(x - x_2)$ og notfæra okkur að núllstöðvarnar x_1 og x_2 uppfylla $x_1 x_2 = c/a$.

Ef $b > 0$, þá reiknum við x_1 fyrst út úr formúlunni

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad \text{og síðan} \quad x_2 = \frac{c/a}{x_1}.$$

Ef aftur á móti $b < 0$, þá reiknum við fyrst x_1 út úr formúlunni

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad \text{og síðan} \qquad x_2 = \frac{c/a}{x_1}.$$

Ef $b^2 \approx 4ac$ þá lendum við í styttingarskekkjum, en við neyðumst til þess að lifa með þeim.

1.6.3 Áhrif gagnaskekkju

Hugsum okkur að við séum að finna nálgun á núllstöð falls $x \mapsto f(x, \alpha)$. Við viljum finna nálgun x á lausninni $r = r(\alpha)$ sem uppfyllir

$$f(r, \alpha) = 0$$

og við lítum á α sem stika (t.d. náttúrulegur fasti).

Gerum ráð fyrir að α_0 sé nálgun á α og að við þekkjum nálgun á $r(\alpha_0)$ sem er lausn á jöfnunni $f(x, \alpha_0) = 0$.

Við viljum athuga hversu mikil áhrif nálgun á α með α_0 hefur á lausnina okkar, þ.e. við þurfum að meta skekkjuna $r(\alpha) - r(\alpha_0)$.

Ef við gefum okkur að f sé samfellt deildanlegt í grennd um punktinn (x_0, α_0), þar sem $x_0 = r(\alpha_0)$ og $\partial_x f(x_0, \alpha_0) \neq 0$, þá segir setningin um fólgin föll að til sé grennd I um punktinn α_0 í $\mathbb{R}$ og samfellt deildanlegt fall $r : I \to \mathbb{R}$, þannig að $r(\alpha_0) = x_0$ og $f(r(\alpha), \alpha) = 0$ fyrir öll $\alpha \in I$.

Með öðrum orðum má segja að við getum alltaf leyst jöfnuna $f(x, \alpha) = 0$ með tilliti til x þannig að út komi lausn $x = r(\alpha)$ sem er samfellt diffranlegt fall af α.

Keðjureglan gefur okkur nú gildi afleiðunnar, því af jöfnunni $f(r(\alpha), \alpha) = 0$ leiðir að fallið $I \ni \alpha \mapsto f(r(\alpha), \alpha)$ er fast, þannig að

$$0 = \frac{\partial}{\partial \alpha} f(r(\alpha), \alpha) = f'_x(r(\alpha), \alpha) \cdot r'(\alpha) + f'_\alpha(r(\alpha), \alpha).$$

Þetta gefur

$$r'(\alpha) = \frac{-f'_\alpha(r(\alpha), \alpha)}{f'_x(r(\alpha), \alpha)}.$$

Nú látum við e tákna skekkjuna í nálguninni á α með α_0, $e = \alpha - \alpha_0$. Þá fáum við skekkjumatið

$$r(\alpha) - r(\alpha_0) \approx r'(\alpha_0) \cdot e = \frac{-f'_\alpha(r(\alpha_0), \alpha_0)}{f'_x(r(\alpha_0), \alpha_0)} \cdot e$$

og jafnframt mat á hlutfallslegri skekkju

$$\frac{|r(\alpha) - r(\alpha_0)|}{|r(\alpha)|} \approx \frac{|f'_\alpha(r(\alpha_0), \alpha_0)|}{|r(\alpha_0) f'_x(r(\alpha_0), \alpha_0)|} \cdot |e|.$$

1.6.4 Sýnidæmi

Við skulum nú líta á það verkefni að finna nálgun á minnstu jákvæðu lausn jöfnunnar $\sin(\pi x) = 1 - e^{-x}$, þar sem við gerum ráð fyrir því að þurfa að nálga π með 3.14.

Okkur eru gefnar niðurstöður úr nálguninni með einhverri aðferð. Við setjum $f(x, \alpha) = 1 - e^{-x} - \sin(\alpha x)$ og fáum

n	x_n	$\lvert x_{n+1} - x_n\rvert$	$\frac{\lvert x_{n+1}-x_n\rvert}{\lvert x_n - x_{n-1}\rvert^2}$
0			0.8
1	0.81276894538752	0.00014017936338	0.8597
2	0.81262876602414	0.00000001621651	0.8253
3	0.81262874980763	0.00000000000000	0.8444

Hér er $\alpha = \pi$ og $\alpha_0 = 3.14$ og þar með $|e| < 0.0016$.

Hlutafleiðurnar eru $f'_x(x, \alpha) = e^{-x} - \alpha \cos(\alpha x)$ og $f'_\alpha(x, \alpha) = -x \cos(\alpha x)$.

Við stingum tölunum okkar inn í matið og notum punktinn $(x_3, \alpha_0) = (0.8126, 3.14)$. Það gefur

$$\begin{aligned} r(\pi) - r(3.14) &\approx r'(3.14) \cdot e \\ &\approx \frac{|0.8126 \cdot \cos(0.8126 \cdot 3.14)|}{|e^{-0.8126} - 3.14 \cdot \cos(0.8126 \cdot 3.14)|}\, 0.0016 \\ &\approx 0.4 \cdot 10^{-3} \end{aligned}$$

Þetta mat segir okkur að við eigum að gera ráð fyrir að áhrif gagnaskekkjunnar séu þau að við fáum lausn með þremur réttum stöfum, $r(\pi) \approx 0.813$. Nálgun okkar á minnstu jákvæðu lausn jöfnunnar $\sin(\pi x) = 1 - e^{-x}$ er því 0.813.

1.6.5 O-ritháttur

Látum f og g vera tvö föll sem skilgreind eru á bili $I \subset \mathbb{R}$ og látum c vera tölu á I eða annan hvorn endapunkt I.

Við segjum að $f(t)$ *sé stórt O af* $g(t)$ og skrifum

$$f(t) = O(g(t)), \qquad t \to c,$$

ef til er fasti $C > 0$ þannig að ójafnan

$$|f(t)| \le C|g(t)|$$

gildi fyrir öll t í einhverri grennd um c.

Athugið að grennd um $c = +\infty$ er bil af gerðinni $]\alpha, +\infty[$ og grennd um $c = -\infty$ er bil af gerðinni $]-\infty, \alpha[$.

1.6.6 O-ritháttur og skekkja í Taylor-nálgnum

Oft er O-ritháttur notaður þegar fjallað er um skekkjur í Taylor-nálgunum,

$$\begin{aligned} f(x) - T_n f(x;c) &= f(x) - f(c) - f'(x-c) - \cdots - \frac{f^{(n)}(c)}{n!}(x-c)^n \\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-c)^{n+1} = O\big((x-c)^{n+1}\big), \quad x \to c \end{aligned}$$

1.6.7 Sýnidæmi

Það eru til haugar af dæmum, sem við þekkjum vel.

Setning Taylors gefur okkur:

$$\begin{aligned} x - \sin x &= O(x^3), \quad x \to 0 \\ x - \frac{x^3}{3!} - \sin x &= O(x^5), \quad x \to 0 \end{aligned}$$

1.6.8 O-ritháttur fyrir runur

Látum nú (a_n) og (b_n) vera tvær talnarunur. Við segjum að a_n *sé stórt O af* b_n og skrifum

$$a_n = O(b_n),$$

ef til er fasti $C > 0$ þannig að ójafnan

$$|a_n| \leq C|b_n|$$

gildi fyrir öll $n = 0, 1, 2, 3, \ldots$.

1.6.9 Tvö sýnidæmi

- Út frá Taylor-röðinni fyrir $\cos x$ fáum við að

$$\cos(1/n) - 1 + 1/(2n^2) = O(1/n^4)$$

- Út frá

$$\sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \leq \frac{1}{2\sqrt{n}}$$

  sjáum við að

$$\sqrt{n+1} - \sqrt{n} = O(\frac{1}{\sqrt{n}})$$

1.7 Fleytitalnakerfið

1.7.1 Framsetning á tölum

Ef r er rauntala frábrugðin 0 og β er náttúrleg tala, 2 eða stærri, þá er til einhlýtt ákvörðuð framsetning á r af gerðinni

$$r = \pm(0.d_1 d_2 \dots d_k d_{k+1} \dots)_\beta \times \beta^e$$

þar sem e er heiltala og d_j eru heiltölur

- $1 \le d_1 < \beta$,
- $0 \le d_j < \beta$, $j = 2, 3, 4, \dots$.

Tölvur reikna ýmist í *tvíundarkerfi* með $\beta = 2$ eða í *sextánundarkerfi* með $\beta = 16$, en við mannfólkið með okkar tíu fingur reiknum í *tugakerfi* með $\beta = 10$.

1.7.2 Mantissa

Formerkið og runan

$$\pm(0.d_1 d_2 \dots d_k d_{k+1} \dots)_\beta = \pm \sum_{j=1}^{\infty} \frac{d_j}{\beta^j}$$

nefnist *mantissa* tölunnar r.

Við skrifum

$$(0.d_1 d_2 \dots d_k)_\beta = \sum_{j=1}^{k} \frac{d_j}{\beta^j}$$

ef $d_{k+1} = d_{k+2} = \dots = 0$ og segjum þá að talan r hafi k-stafa mantissu.

1.7.3 Markverðir β-stafir

Ef rauntalan x er nálgun á r, þá segjum við að x sé nálgun á r með *að minnsta kosti* t *markverðum* β *-stöfum* ef

$$\frac{|r - x|}{|r|} \le \beta^{-t}.$$

Ef við höfum að auki að

$$\beta^{-t-1} < \frac{|r - x|}{|r|} \le \beta^{-t}.$$

þá segjum við að x sé nálgun á r með t *markverðum* β *-stöfum*.

Athugið að ef e er minnsta heila talan þannig að $|r| < \beta^e$, þá gefur seinni ójafnan matið

$$|r - x| = (0.0 \dots 0 a_t a_{t+1} \dots)_\beta \times \beta^e,$$

þar sem núllin aftan við punkt eru t talsins.

1.7.4 Afrúningur talna

Ef r er sett fram á stöðluðu β-fleytitöluformi, þá nefnist talan

$$x = (\pm 0.d_1 d_2 \dots d_k)_\beta \times \beta^e$$

afskurður tölunnar r *við* k *-ta aukastaf* r, en talan

$$x = \begin{cases} \pm(0.d_1 d_2 \dots d_k)_\beta \times \beta^e, & d_{k+1} < \beta/2, \\ \pm((0.d_1 d_2 \dots d_k)_\beta + \beta^{-k}) \times \beta^e, & d_{k+1} \geq \beta/2. \end{cases}$$

nefnist *afrúningur tölunnar* r *við* k *-ta aukastaf*.

Við köllum þessar aðgerðir *afskurð* (e. chopping) og *afrúning* (e. rounding).

1.7.5 Fleytitölukerfi

Fleytitölukerfi er endanlegt hlutmengi í $\mathbb{R}$, sem samanstendur af öllum tölum

$$\pm(0.d_1 d_2 \dots d_k)_\beta \times \beta^e$$

þar sem d_j eru heiltölur eins og áður var lýst, k er föst tala og við höfum mörk á veldisvísinum $m \leq e \leq M$.

Allar tölvur vinna með eitthvert fleytitölukerfi, oftast með grunntölu $\beta = 2$ eða $\beta = 16$ eins og áður sagði.

Eftir hverja aðgerð í tölvunni þarf að nálga útkomuna með *afskurði* eða *afrúningu*.

Ef við förum ekki varlega þá getur þetta magnað upp skekkju.

Sjá Úrlausn annars stigs jöfnu.

1.7.6 IEEE staðlar

- Single: $\beta = 2, k = 24, m = -125$ og $M = 128$,
- Double: $\beta = 2, k = 53, m = -1021$ og $M = 1024$.

1.7.7 Útreikningur í tugakerfi

Þegar reiknað er í tugakerfi er tölurnar afrúnaðar við k-ta aukastaf ef skekkjan í nálgun á þeim er minni en $\frac{1}{2} \times 10^{-k}$. Ef

$$\frac{|r - x|}{|r|} < 10^{-k-1}$$

þá treystum við öllum k stöfum mantissunnar, en ef

$$\frac{|r - x|}{|r|} > 10^{-k+q},$$

þá eru síðustu q stafir mantissunnar marklausir auk þess sem vænta má nokkurs fráviks í d_{k-q}.

KAFLI 2

Núllstöðvar

Build a man a fire, and he'll be warm for a day. Set a man on fire, and he'll be warm for the rest of his life.

- Terry Pratchett, Jingo

2.1 Nálgun á núllstöð

2.1.1 Skilgreining

Munum að talan $p \in I$ sögð vera *núllstöð* fallsins $f : I \to \mathbb{R}$ ef

$$f(p) = 0.$$

2.1.2 Dæmi

Það er auðvelt að finna núllstöðvar (**rót**) annars stigs margliðu $ax^2 + bx + c$, því

$$ax^2 + bx + c = 0$$

ef

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Svipaðar formúlur eru til fyrir núllstöðvar þriðja og fjórða stigs margliða. Einnig þekkjum við núllstöðvar hornafalla.

2.1.3 Athugasemd

Almennt er hins vegar erfitt að finna núllstöðvar falla. Til dæmis er ekki til almenn formúla fyrir núllstöðvar margliða af stigi 5 og hærra (sjá Abel-Ruffini setningin).

Eins er ekki hægt treysta á það að geta fundið nákvæmlega núllstöðvar almennra falla með því að nota þekkingu okkar á algebru og stærðfræðigreininu. Hverjar (og hversu margar) eru t.d. núllstöðvar

$$e^x + x^3?$$

Aðferðirnar í þessum kafla ganga út á að finna nálgun á núllstöðvum falla og í sumum tilvikum hjálpa þær okkur einnig að sýna fram á tilvist núllstöðva (sem er ekki alltaf sjálfgefin).

2.2 Helmingunaraðferð

Fyrsta aðferðin til að finna núllstöðvar sem við skoðum kallast helmingunaraðferð (e. bisection method).

2.2.1 Milligildissetningin

Ef f er samfellt á $[a, b]$ og y er einhver tala á milli $f(a)$ og $f(b)$, þá er til c þannig að $a < c < b$ og $f(c) = y$.

2.2.2 Afleiðing

Svo ef við höfum a og b þannig að $a < b$ og þannig að $f(a)$ og $f(b)$ hafi ólík formerki, þá hefur f núllstöð p á bilinu $[a, b]$.

Notum okkur þetta til þess að finna rætur.

1. Látum $x = \frac{1}{2}(a + b)$ vera miðpunkt $[a, b]$.
2. Reiknum $f(x)$, þá geta þrjú tilvik komið upp:
   (a) $f(x) = 0$ og leitinni að rót er lokið.
   (b) $f(a)$ og $f(x)$ hafa sama formerki, þannig að við leitum að rót á bilinu $[x, b]$.
   (c) $f(x)$ og $f(b)$ hafa sama formerki, þannig að við leitum að rót á bilinu $[a, x]$.

Í tilviki (ii) segir milligildissetningin að f hafi rót á bilinu $[x, b]$, og í tilviki (iii) er rótin á bilinu $[a, x]$. Þá getum við farið aftur í skref 1, nema með helmingi minna bil en áður.

Með því að ítreka þetta ferli n sinnum fáum við minnkandi runu af bilum

$$[a, b] = [a_1, b_1] \supset [a_2, b_2] \supset \cdots \supset [a_n, b_n].$$

Billengdin helmingast í hverju skrefi og milligildissetningin segir okkur að það sé núllstöð á öllum bilunum.

Rununa af bilunum

$$[a, b] = [a_1, b_1] \supset \cdots \supset [a_n, b_n] \supset \cdots$$

skilgreinum við með ítrun og notum til þess rununa $x_n = \frac{1}{2}(a_n + b_n)$.

Setjum $a_0 = a$, $b_0 = b$, og $x_0 = \frac{1}{2}(a + b)$.

Gefið er $x_0, \ldots, x_n$. Reiknum $f(x_n)$.

1. Ef $f(x_n) = 0$, þá er núllstöð fundin og við hættum.
2. Ef $f(x_n)$ og $f(a_n)$ hafa sama formerki, þá setjum við $a_{n+1} = x_n$, $b_{n+1} = b_n$, og $x_{n+1} = \frac{1}{2}(a_{n+1} + b_{n+1})$
3. annars setjum við $a_{n+1} = a_n$, $b_{n+1} = x_n$ og $x_{n+1} = \frac{1}{2}(a_{n+1} + b_{n+1})$.

2.2.3 Skekkjumat í helmingunaraðferð

Ef við látum miðpunktinn $p_n = \frac{1}{2}(a_n + b_n)$ vera nálgunargildi okkar fyrir núllstöð fallsins f í bilinu $[a_n, b_n]$, þá er skekkjan í nálguninni

$$e_n = p - p_n$$

og við höfum skekkjumatið

$$|e_n| \leq \frac{b_n - a_n}{2} = \frac{b_{n-1} - a_{n-1}}{2^2} = \ldots = \frac{b_1 - a_1}{2^n},$$

það er

$$|e_n| < \frac{b - a}{2^n}.$$

2.2.4 Fyrirframmat á skekkju

Nú er auðvelt að meta hversu margar ítrekanir þarf að framkvæma til þess að nálgunin lendi innan gefinna skekkjumarka.

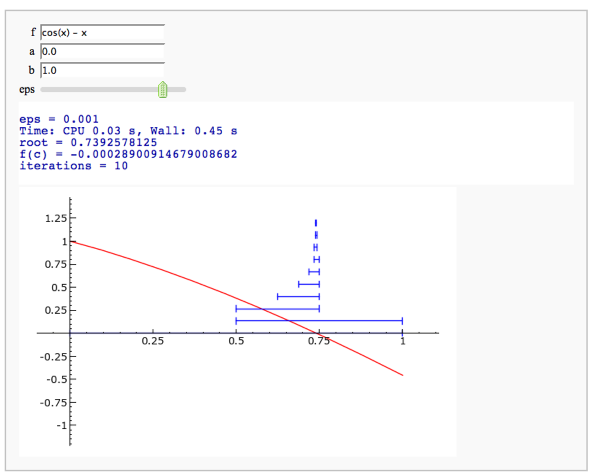
Ef $\varepsilon > 0$ er gefið og við viljum að $|e_n| < \varepsilon$, þá dugir að

$$|e_n| \leq \frac{b - a}{2^n} < \varepsilon.$$

Seinni ójafnan jafngildir því að

$$n > \frac{\ln\big((b - a)/\varepsilon\big)}{\ln 2}.$$

Double Precision Root Finding Using Bisection



2.3 Fastapunktsaðferð

Næsta aðferð sem við skoðum kallast fastapunktsaðferð (e. fixed point method) og er til að finna fastapunkta en ekki núllstöðvar. Það er hins vegar hægt að nota hana til þess að finna núllstöðvar, sjá athugasemd hér að *neðan*.

2.3.1 Skilgreining

Látum $f : [a, b] \to \mathbb{R}$ vera samfellt fall. Punktur $r \in [a, b]$ þannig að

$$f(r) = r$$

kallast *fastapunktur* fallsins f.

Athugasemd: Athugum að í fastapunktum skerast graf fallsins $y = f(x)$ og línan $y = x$. Verkefnið að ákvarða fastapunkta fallsins r er því jafngilt því að athuga hvar graf f sker línuna $y = x$.

2.3.2 Tenging við núllstöðvar

Verkefnið að finna fastapunkta fallsins $g(x)$ er jafngilt því að finna núllstöðvar fallsins $f(x) = g(x) - x$.

Þannig að ef við viljum t.d. finna núllstöð $f(x) = e^x + x^3$ þá er nóg að finna fastapunkt fallsins $g(x) = e^x + x^3 + x$.

2.3.3 Reiknirit

Byrjunarskref: Valin er tala $x_0 \in [a, b]$.

Ítrunarskref: Ef $x_0, \dots, x_n$ hafa verið valin, þá setjum við

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

Athugasemd: Til þess að þetta sé vel skilgreind runa, þá verðum við að gera ráð fyrir að $f(x) \in [a, b]$ fyrir öll $x \in [a, b]$. Þetta skilyrði er einnig skrifað

$$f([a, b]) \subset [a, b].$$

Athugasemd: Ef f er samfellt og runan er samleitin með markgildið r, þá er

$$r = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(r).$$

Þetta segir okkur að *ef* við getum séð til þess að runan verði samleitin, þá er markgildið fastapunktur.

2.3.4 Skilgreining: Herping

Fall $f : [a, b] \to \mathbb{R}$ er sagt vera *herping* ef til er fasti $\lambda \in [0, 1[$ þannig að

$$|f(x) - f(y)| \le \lambda |x - y| \qquad \text{fyrir öll } x, y \in [a, b].$$

Athugasemd: Sérhver herping er samfellt fall.

2.3.5 Setning

Ef f er deildanlegt fall á $]a, b[$, þá gefur meðalgildissetningin okkur til er ξ milli x og y þannig að

$$f(x) - f(y) = f'(\xi)(x - y).$$

Ef til er $\lambda \in [0, 1[$ þannig að $|f'(x)| \le \lambda$ fyrir öll $x \in [a, b]$, þá er greinilegt að f er herping.

2.3.6 Fastapunktssetningin

Látum $f : [a, b] \to [a, b]$ vera herpingu. Þá hefur f nákvæmlega einn fastapunkt r á bilinu $[a, b]$ og runan (x_n) þar sem

$$\begin{aligned} x_0 &\in [a, b] \quad \text{getur verið hvaða tala sem er og} \\ x_{n+1} &= f(x_n), \quad n \geq 0, \end{aligned}$$

stefnir á fastapunktinn.

Sönnunina brjótum við upp í nokkur skref.

1. skref, herping hefur í mesta lagi einn fastapunkt

Sönnum þetta með mótsögn.

Gerum ráð fyrir að r og s séu tveir ólíkir fastapunktar á $[a, b]$. Þá er

$$|r - s| = |f(r) - f(s)| \leq \lambda |r - s| < |r - s|$$

því $\lambda < 1$. Þetta fær ekki staðist, þannig að fjöldi fastapunkta er í mesta lagi einn

2. skref, fallið f hefur fastapunkt:

Látum $g(x) = f(x) - x$, þá eru núllstöðvar g nákvæmlega fastapunktar f.

Þar sem $a \leq f(x) \leq b$ fyrir öll $x \in [a, b]$ er

$$\begin{cases} g(a) = f(a) - a \geq 0 \\ g(b) = f(b) - b \leq 0 \end{cases}$$

Ef annað hvort $g(a) = 0$ eða $g(b) = 0$ höfum við fundið fastapunkt fallsins f og við getum hætt.

Ef hins vegar $g(a) > 0$ og $g(b) < 0$ þá hefur g ólík formerki í endapunktum bilsins $[a, b]$ og hefur því núllstöð r á bilinu skv. milligildissetningunni. Þá er r jafnframt fastapunktur f.

Skref 1 og 2 sýna því að fallið f hefur nákvæmlega einn fastapunkt á bilinu.

3. skref, runan (x_n) er samleitin

Látum r vera ótvírætt ákvarðaða fastapunktinn á $[a, b]$.

Við notfærum okkur að f er herping og að r er fastapunktur f, þá fæst að fyrir sérhvert $k \in \mathbb{N}$ þá er

$$|r - x_k| = |f(r) - f(x_{k-1})| \leq \lambda |r - x_{k-1}|$$

það er $|r - x_k| \leq \lambda |r - x_{k-1}|$.

Með því að nota þetta n-sinnum þá fæst að

$$\begin{aligned} |r - x_n| &\leq \lambda |r - x_{n-1}| & (k = n) \\ &\leq \lambda^2 |r - x_{n-2}| & (k = n - 1) \\ &\vdots & \vdots \\ &\leq \lambda^n |r - x_0| & (k = 1). \end{aligned}$$

Þar sem $\lambda < 1$ er því

$$\lim_{n \to +\infty} |r - x_n| \leq \lim_{n \to +\infty} \lambda^n |r - x_0| = 0,$$

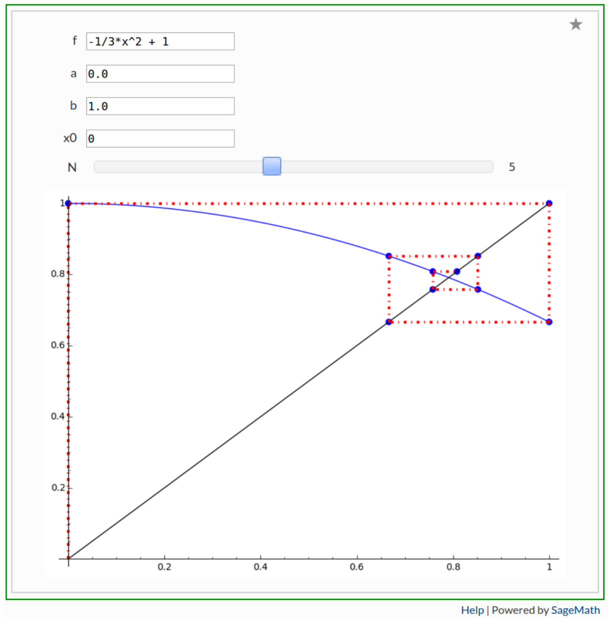
það er runan x_n stefnir á r.

2.3.7 Fastapunktsaðferð er að minnsta kosti línulega samleitin

Af skilgreiningunni á rununni x_n leiðir beint að

$$|e_{n+1}| = |r - x_{n+1}| = |f(r) - f(x_n)| \leq \lambda |r - x_n| = \lambda |e_n|$$

sem segir okkur að fastapunktsaðferð sé að minnsta kosti línulega samleitin ef f er herping.



2.4 Sniðilsaðferð

Næst er aðferð til að finna núllstöðvar sem kallast *sniðilsaðferð* (e. secant method)

Gefið er fallið $f : [a, b] \to \mathbb{R}$. Við ætlum að ákvarða núllstöð f, þ.e.a.s. $p \in [a, b]$ þannig að

$$f(p) = 0.$$

Rifjum upp að *sniðill* við graf f gegnum punktana $(\alpha, f(\alpha))$ og $(\beta, f(\beta))$ er gefinn með jöfnunni

$$y = f(\alpha) + f[\alpha, \beta](x - \alpha)$$

þar sem hallatalan er

$$f[\alpha, \beta] = \frac{f(\beta) - f(\alpha)}{\beta - \alpha} = \frac{f(\alpha) - f(\beta)}{\alpha - \beta}.$$

Sniðillinn sker x-ásinn í punkti s þar sem

$$0 = f(\alpha) + f[\alpha, \beta](s - \alpha) \quad \text{sem jafngildir því að} \quad s = \alpha - \frac{f(\alpha)}{f[\alpha, \beta]}.$$

2.4.1 Reiknirit

Byrjunarskref: Giskað er á tvö gildi x_0 og x_1.

Ítrunarskref: Fyrir $n > 1$ þá er punkturinn x_{n+1} skilgreindur sem skurðpunktur sniðilsins gegnum $(x_{n-1}, f(x_{n-1}))$ og $(x_n, f(x_n))$ við x-ás, þ.e.

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_n, x_{n-1}]}.$$

2.4.2 Samleitin runa stefnir á núllstöð f

Gefum okkur að runan (x_n) sé samleitin að markgildinu r. Meðalgildissetningin segir okkur þá að til sé punktur η_n á milli x_{n-1} og x_n þannig að

$$f[x_n, x_{n-1}] = f'(\eta_n),$$

og greinilegt er að $\eta_n \to r$.

Við fáum því

$$r = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty}\left(x_n - \frac{f(x_n)}{f'(\eta_n)}\right) = r - \frac{f(r)}{f'(r)}$$

Þessi jafna jafngildir því að $f(r) = 0$.

2.4.3 Skekkjumat í nálgun á $f(x)$ með $p_n(x)$

Sniðilinn sem við notum er graf 1. stigs margliðunnar

$$p_n(x) = f(x_n) + \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}(x - x_n) = f(x_n) + f[x_n, x_{n-1}](x - x_n)$$

Samkvæmt skilgreiningu er $p_n(x_{n+1}) = 0$ svo x_{n+1} uppfyllir jöfnuna

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_n, x_{n-1}]}.$$

Við þurfum að vita hver skekkjan er á því að nálga $f(x)$ með $p_n(x)$.

Niðurstaðan er að fyrir sérhvert $x \in [a, b]$ er til ξ_n sem liggur í minnsta bilinu sem inniheldur x, x_n og x_{n-1} þannig að

$$f(x) - p_n(x) = \frac{1}{2}f''(\xi_n)(x - x_n)(x - x_{n-1})$$

Ljóst er að matið gildir ef $x = x_{n-1}$ eða $x = x_n$.

Festum því punktinn x og gerum ráð fyrir að $x \neq x_1$ og $x \neq x_n$.

Skilgreinum fallið

$$g(t) = f(t) - p_n(t) - \lambda(t - x_n)(t - x_{n-1})$$

þar sem λ er valið þannig að $g(x) = 0$.

Látum nú $\alpha < \beta < \gamma$ vera uppröðun á punktunum x_{n-1}, x_n og x.

Fallið

$$g(t) = f(t) - p_n(t) - \lambda(t - x_n)(t - x_{n-1})$$

hefur núllstöð í öllum punktunum þremur.

Meðalgildissetningin gefur þá að $g'(t)$ hefur eina núllstöð í punkti á bilinu $]\alpha, \beta[$ og aðra í $]\beta, \gamma[$.

Af því leiðir aftur að $g''(t)$ hefur núllstöð, ξ_n, í $[\alpha, \gamma]$, sem er minnsta bilið sem inniheldur alla punktana x_{n-1}, x_n og x.

Af þessu leiðir

$$0 = g''(\xi_n) = f''(\xi_n) - 2\lambda \quad \text{þþaa} \quad \lambda = \tfrac{1}{2} f''(\xi_n).$$

Nú var λ upprunalega valið þannig að $g(x) = 0$. Þar með er

$$f(x) - p_n(x) = \frac{1}{2} f''(\xi_n)(x - x_n)(x - x_{n-1}).$$

2.4.4 Skekkjumat í sniðilsaðferð

Skoðum hvað af þessu leiðir:

Nú er $f(r) = 0$ og því

$$-p_n(r) = \frac{1}{2} f''(\xi_n) e_n \cdot e_{n-1}.$$

Eins er

$$-p_n(r) = -f[x_n, x_{n-1}] e_{n+1} = -f'(\eta_n) e_{n+1},$$

þar sem η_n fæst úr meðalgildissetningunni og liggur á milli x_n og x_{n+1}. Niðurstaðan verður því

$$e_{n+1} = \frac{-\frac{1}{2} f''(\xi_n)}{f[x_n, x_{n+1}]} e_n e_{n-1} = \frac{-\frac{1}{2} f''(\xi_n)}{f'(\eta_n)} e_n e_{n-1}$$

það er

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n e_{n-1}} = \lim_{n\to\infty} \frac{-\frac{1}{2} f''(\xi_n)}{f'(\eta_n)} = \frac{-\frac{1}{2} f''(r)}{f'(r)}.$$

2.4.5 Setning

Ef sniðilsaðferð er samleitin, $f \in C^2([a, b])$ (tvisvar diffranlegt) og $f'(r) \neq 0$, þá er sniðilsaðferðin ofurlínuleg.

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = \lim_{n\to\infty} \frac{|e_{n+1} e_{n-1}|}{|e_n e_{n-1}|} = \lim_{n\to\infty} \frac{|e_{n-1} \frac{1}{2} f''(r)|}{|f'(r)|} = 0$$

Raunar þá er sniðilsaðferðin samleitin af stigi $\alpha = (1 + \sqrt{5})/2 \approx 1,618$ og með $\lambda = \left(\frac{f''(r)}{2f'(r)}\right)^{\alpha - 1}$.

2.5 Aðferð Newtons

Í sniðilsaðferðinni létum við x_{n+1} vera skurðpunkt sniðils gegnum $(x_{n-1}, f(x_{n-1}))$ og $(x_n, f(x_n))$ við x-ás og fengum við rakningarformúluna

$$x_{n+1} = x_n - \frac{f(x_n)}{f[x_n, x_{n-1}]}.$$

Aðferð Newtons er nánast eins, nema í stað sniðils tökum við snertil í punktinum $(x_n, f(x_n))$.

Rakningarformúlan er eins, nema hallatalan verður $f'(x_n)$ í stað $f[x_n, x_{n-1}]$

2.5.1 Reiknirit

Byrjunarskref: Giskað er á eitt gildi x_0.

Ítrunarskref: Gefin eru $x_0, \dots, x_n$. Punkturinn x_{n+1} er skurðpunktur snertils gegnum $(x_n, f(x_n))$ við x-ás,

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

2.5.2 Upprifjun

Munum að snertill við graf f í punktinum x_n er

$$y = f(x_n) + f'(x_n)(x - x_n),$$

þessi lína sker x-ásinn ($y = 0$) þegar $x = x_n - \frac{f(x_n)}{f'(x_n)}$.

2.5.3 Samleitin runa stefnir á núllstöð f

Gefum okkur að runan (x_n) sé samleitin með markgildið r. Við fáum því

$$r = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty}\left(x_n - \frac{f(x_n)}{f'(x_n)}\right) = r - \frac{f(r)}{f'(r)}$$

Þessi jafna jafngildir því að $f(r) = 0$.

Þannig að ef runan er samleitin þá fáum við núllstöð.

2.5.4 Skekkjumat í nálgun á $f(x)$ með $p_n(x)$

Snertillinn við f í punktinum x_n er 1. stigs margliðan

$$p_n(x) = f(x_n) + f'(x_n)(x - x_n)$$

Samkvæmt skilgreiningu er $p_n(x_{n+1}) = 0$ svo x_{n+1} uppfyllir jöfnuna

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

Athugum að p_n er fyrsta Taylor nálgunin við fallið f kringum x_n. Setning Taylors gefur að til er ξ_n sem liggur á milli r og x_n þannig að

$$f(r) - p_n(r) = \frac{1}{2} f''(\xi_n)(r - x_n)^2.$$

2.5.5 Skekkjumat í aðferð Newtons

Nú er $f(r) = 0$ og því

$$-p_n(r) = \frac{1}{2} f''(\xi_n) e_n^2.$$

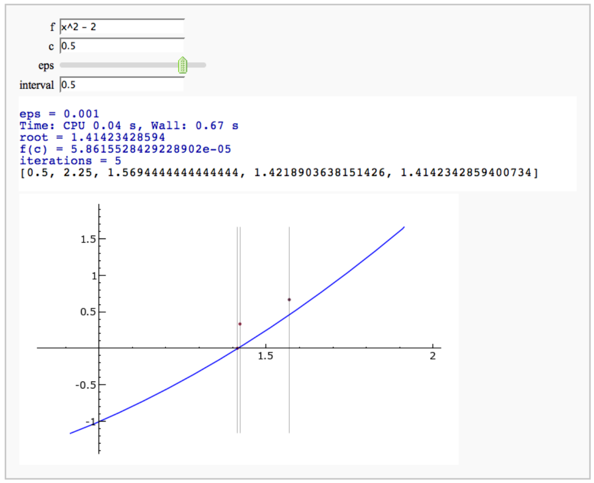
Eins er fæst af skilgreiningunni á p_n að

$$-p_n(r) = -f'(x_n) e_{n+1}$$

Niðurstaðan verður því

$$e_{n+1} = \frac{-\frac{1}{2} f''(\xi_n)}{f'(x_n)} e_n^2$$

Double Precision Root Finding Using Newton's Method



2.5.6 Setning

Ef aðferð Newtons fyrir fallið f er samleitin, $f \in C^2([a,b])$ og $f'(r) \neq 0$, þá fáum við:

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n^2} = \frac{-\frac{1}{2}f''(r)}{f'(r)}$$

Það þýðir að aðferð Newtons er ferningssamleitin.

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n^2} = \lim_{n\to\infty} \frac{-\frac{1}{2}f''(\xi_n)}{f'(x_n)} = \frac{-\frac{1}{2}f''(r)}{f'(r)}$$

Athugasemd: Athugið að það er ekki sjálfgefið að aðferð Newtons sé samleitin.

Auðvelt er að finna dæmi þar sem vond upphafságiskun x_0 skilar runu sem er ekki samleitin.

2.6 Samanburður á aðferðum

Aðferð	Samleitni	Stig samleitni
Helmingunaraðferð (e. bisection method)	Já, ef $f(a)f(b) < 0$	1, línuleg
Fastapunktsaðferð (e. fixed point iteration)	Ekki alltaf. En saml. ef f er herping	amk 1
Sniðilsaðferð (e. secant method)	Ekki alltaf	$\approx 1,618$, ef $f'(r) \neq 0$
Aðferð Newtons (e. Newtons method)	Ekki alltaf	2, ef $f'(r) \neq 0$

Aðvörun: Þó að aðferð Newtons sé samleitin af stigi 2, en sniðilsaðferðin af stigi u.þ.b. 1,618, þá er í vissum tilfellum hagkvæmara að nota sniðilsaðferðina ef það er erfitt að reikna gildin á afleiðunni f'.

KAFLI 3

Brúun

Over the centuries, mankind has tried many ways of combating the forces of evil... prayer, fasting, good works and so on. Up until Doom, no one seemed to have thought about the double-barrel shotgun. Eat leaden death, demon. – Terry Pratchett

3.1 Inngangur

3.1.1 Markmiðið

Viðfangsefni þessa kafla er að finna ferla sem ganga gegnum fyrirfram gefna *brúunarpunkta* $(x_0, y_0), \ldots, (x_m, y_m)$ í planinu eða liggja nálægt punktunum í einhverjum skilningi.

Fyrst viljum við finna graf margliðu p sem fer gegnum punktana. Þá þurfum við að gefa okkur að $x_i \neq x_j$ ef $i \neq j$.

Við sýnum fram á að það sé alltaf hægt að finna margliðu p af stigi $\leq m$ sem uppfyllir $p(x_i) = y_i$ í öllum punktum og að slík margliða sé ótvírætt ákvörðuð.

Hún nefnist *brúunarmargliða* fyrir punktana $(x_0, y_0), \ldots, (x_m, y_m)$.

Við alhæfum þetta verkefni með því að úthluta sérhverjum punkti jákvæðri heiltölu m_i og krefjast þess graf margliðunnar fari í gegnum alla punktana og til viðbótar að allar afleiður $p^{(j)}$ upp að stigi $m_i - 1$ taki einnig fyrirfram gefin gildi $y_i^{(j)}$.

3.1.2 Brúun

Við tilraunir þá fáum við oft aðeins strjálar mælingar, t.d. ef við mælum hljóðhraða við mismunandi hitastig. Hins vegar þá viljum við vita hvert sambandið er fyrir öll möguleg hitastig. Brúunin er margliða og hún skilgreind er fyrir allar rauntölur og ''brúar'' því gildin milli mælipunktanna.

3.1.3 Afhverju margliður?

- Einfalt að meta fallgildin fyrir margliður (Reiknirit Horners).
- Einfalt að diffra og heilda margliður.
- Margliður eru óendanlega oft diffranlegar.
- *Setning Weierstrass:* Látum f vera samfellt fall á bili $[a, b]$. Fyrir sérhvert $\varepsilon > 0$ þá er til margliða p þannig að

$$\|f - p\|_\infty := \max_{x \in [a,b]} |f(x) - p(x)| < \varepsilon.$$

Athugasemd: Setning Weierstrass segir að margliður nægja til að nálga samfelld föll. Það er sama hvað samfellda fall við skoðum það er alltaf til margliða sem nálgar það eins vel og við viljum á lokuðu bili.

3.1.4 Margliður

Fall p af gerðinni

$$p(x) = a_0 + a_1 x + \ldots + a_m x^m$$

þar sem m er heiltala og $a_0, \ldots, a_m$ eru tvinntölur nefnist margliða.

Stærsta talan j þannig að $a_j \neq 0$ nefnist *stig margliðunnar* p.

Ef allir stuðlarnir eru 0 þá nefnist p *núllmargliðan* og við segjum að stig hennar sé $-\infty$.

Munum að stuðullinn a_j við veldið x^j er gefinn með formúlunni

$$a_j = \frac{p^{(j)}(0)}{j!}, \quad j = 0, 1, 2, \ldots, m.$$

3.1.5 Mismunandi leiðir á framsetningu

Hægt er að setja sömu margliðuna fram á marga mismunandi vegu, en við nefnum framsetninguna hér að framan *staðalform margliðunnar* p.

Ef við veljum okkur einhvern punkt $x_0 \in \mathbb{R}$, þá getum við skrifað

$$p(x) = b_0 + b_1(x - x_0) + \ldots + b_m(x - x_0)^m$$

og stuðlarnir b_j eru gefnir með

$$b_j = \frac{p^{(j)}(x_0)}{j!}, \quad j = 0, 1, 2, \ldots, m.$$

Þessi formúla er jafngild þeirri staðreynd að ef p er margliða af stigi m. Þá er Taylor-röð p í sérhverjum punkti $x_0 \in \mathbb{R}$ bara margliðan p, og stuðlarnir í Taylor-röðinni eru gefnir með formúlunum fyrir b_j að ofan.

3.1.6 Newton-form margliðu

Ef við veljum okkur m punkta $x_0, \ldots, x_{m-1}$ þá nefnist framsetning af gerðinni

$$p(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \ldots + c_m(x - x_0)\cdots(x - x_{m-1})$$

Newton-form margliðunnar p miðað við punktana $x_0, \ldots, x_{m-1}$.

3.1.7 Reiknirit Horners

Við munum mikið fást við margliður á Newton-formi og því er nauðsynlegt að hafa hraðvirkt reiknirit til þess að reikna út fallgildi p út frá þessari framsetningu.

Eitt slíkt reiknirit er nefnt *reiknirit Horners*. Það byggir á því að nýta sér að þættirnir $(x - x_j)$ eru endurteknir í liðunum

$$(x - x_0), \quad (x - x_0)(x - x_1), \quad (x - x_0)(x - x_1)(x - x_2), \quad \dots$$

Þar sem við sleppum við að hefja í veldi þá komumst við af með fáar reikniaðgerðir hér.

Ef $m = 2$ má skrifa Newton-form p sem

$$p(x) = c_0 + (x - x_0)(c_1 + (x - x_1) \cdot c_2).$$

Ef $m = 3$ er það

$$p(x) = c_0 + (x - x_0)(c_1 + (x - x_1)(c_2 + (x - x_2)c_3))$$

og ef $m = 4$ er það

$$p(x) = c_0 + (x - x_0)(c_1 + (x - x_1)(c_2 + (x - x_2)(c_3 + c_4(x - x_3)))).$$

Reikniritið vinnur á þessari stæðu með því að margfalda upp úr svigunum frá hægri til vinstri.

Skilgreinum tölur $b_0, b_1, \dots$ á eftirfarandi hátt. Fyrst setjum við

$$b_n = c_n.$$

Fyrir hvert k frá $n - 1$ niður í 0 þá setjum við

$$b_k = c_k + (a - x_k)b_{k+1}.$$

Þá er $b_0 = p(a)$.

$$p(a) = \underbrace{c_0 + (a - x_0)(\underbrace{c_1 + (a - x_1)(\underbrace{c_2 + (a - x_2)(\underbrace{c_3 + (a - x_3)\underbrace{c_4}_{b_4}}_{b_3})}_{b_2})}_{b_1})}_{b_0}.$$

3.2 Margliðubrúun: Lagrange-form

3.2.1 Margliðubrúun

Látum nú $(x_0, y_0), \dots, (x_m, y_m)$ vera gefna punkta í plani. Við höfum áhuga á að finna margliðu p af lægsta mögulega stigi þannig að

$$p(x_k) = y_k, \quad k = 0, \dots, m.$$

Slík margliða nefnist *brúunarmargliða* fyrir punktana $(x_0, y_0), \dots, (x_m, y_m)$

eða *brúunarmargliða gegnum punktana* $(x_0, y_0), \dots, (x_m, y_m)$.

Augljóslega verðum við að gera ráð fyrir að x-hnitin séu ólík, það er $x_j \neq x_k$ ef $j \neq k$.

Verkefnið að finna margliðuna p nefnist *brúunarverkefni fyrir punktana* $(x_0, y_0), \dots, (x_m, y_m)$.

3.2.2 Setning: Brúunarmargliðan er ótvírætt ákvörðuð

Brúunarmargliðan af stigi $\leq m$ fyrir $(x_0, y_0), \ldots, (x_m, y_m)$ er ótvírætt ákvörðuð.

Ef $p(x)$ og $q(x)$ eru tvær brúunarmargliður af stigi $\leq m$ fyrir punktana $(x_0, y_0), \ldots, (x_m, y_m)$ þá er mismunurinn $r(x) = p(x) - q(x)$ margliða af stigi $\leq m$ með núllstöðvar $x_0, \ldots, x_m$. Þetta eru $m + 1$ ólíkir punktar og því er $r(x)$ núllmargliðan samkvæmt undirstöðusetningu algebrunnar. Þar með er $p(x) - q(x)$ núllmargliðan, þ.e. $p(x) = q(x)$.

3.2.3 Setning: Brúunarmargliðan er til

Til er margliða p af stigi $\leq m$ þannig að

$$p(x_0) = y_0, \quad \ldots \quad p(x_n) = y_n.$$

Við notum þrepun til að sýna fram á tilvistina.

Ef $m = 0$, þá erum við aðeins með eitt brúunarskilyrði, $p(x_0) = y_0$, og fastamargliðan $p(x) = y_0$ er lausn af stigi ≤ 0.

G.r.f. að við getum leyst öll brúunarverkefni þar sem fjöldi punkta er m og sýnum að við getum þá leyst verkefnið fyrir $m + 1$ punkt.

Látum q vera brúunarmargliðuna af stigi $\leq m - 1$ fyrir punktana $(x_0, y_0), \ldots, (x_{m-1}, y_{m-1})$ og r vera brúunarmargliðuna af stigi $\leq m - 1$ fyrir punktana $(x_1, y_1), \ldots, (x_m, y_m)$ og setjum síðan

$$p(x) = \frac{x - x_m}{x_0 - x_m} q(x) + \frac{x - x_0}{x_m - x_0} r(x),$$

$p(x)$ er greinilega margliða af stigi $\leq m$. Skoðum nú gildin á p

$$\begin{aligned} p(x_0) &= 1 \cdot q(x_0) + 0 \cdot r(x_0) = y_0, \\ p(x_k) &= \frac{x_k - x_m}{x_0 - x_m} y_k + \frac{x_k - x_0}{x_m - x_0} y_k = y_k, \qquad k = 1, \ldots, m-1, \\ p(x_m) &= 0 \cdot q(x_m) + 1 \cdot r(x_m) = y_m. \end{aligned}$$

Þar með er p brúunarmargliðan sem uppfyllir $p(x_j) = y_j$ fyrir $j = 0, \ldots, m$ og við höfum leyst brúunarverkefnið fyrir $m + 1$ punkt.

3.2.4 Lagrange-form brúunarmargliðunnar

Sönnunin á undan er í raun rakningarformúla til þess að reikna út gildi brúunarmargliðunnar p fyrir punktana $(x_0, y_0), \ldots, (x_m, y_m)$.

Hægt er að skrifa lausnina niður beint

$$p(x) = y_0 \ell_0(x) + y_1 \ell_1(x) + \cdots + y_m \ell_m(x),$$

þar sem $\ell_0, \ldots, \ell_m$ er ákveðinn grunnur fyrir rúm allra margliða $\mathcal{P}_m$ af stigi $\leq m$ og nefnast *Lagrange-margliður fyrir punktasafnið* $(x_0, y_0), \ldots, (x_m, y_m)$.

3.2.5 Lagrange-margliður, tilfellin $m = 0, 1, 2$

- $m = 0$ Ef $m = 0$ þá er $p(x) = y_0$ fastamargliða eins og við höfum séð.

- $m = 1$ Ef $m = 1$, þá blasir við að lausnin er

  $$p(x) = y_0 \frac{(x - x_1)}{(x_0 - x_1)} + y_1 \frac{(x - x_0)}{(x_1 - x_0)},$$

  sem er margliða af stigi ≤ 1 (þ.e. lína) sem leysir brúunarverkefnið.
- $m = 2$ Á hliðstæðan hátt fáum við fyrir $m = 2$ að

  $$p(x) = y_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + y_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + y_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}$$

  leysir brúunarverkefnið.

3.2.6 Lagrange-margliður almenna tilfellið

Almennt fæst lausnin

$$p(x) = y_0\, \ell_0(x) + y_1\, \ell_1(x) + \ldots + y_m\, \ell_m(x)$$

þar sem

$$\ell_k = \prod_{\substack{j=0 \\ j \neq k}}^{m} \frac{(x - x_j)}{(x_k - x_j)}$$

Athugasemd:

$$\ell_k(x_i) = \begin{cases} 1 & \text{ef } i = k \\ 0 & \text{ef } i \neq k \end{cases}$$

Allar margliðurnar ℓ_k eru af stigi m og því er p af stigi $\leq m$. Nú er augljóst útfrá ([p]) og ([l]) að p er lausn brúunarverkefnisins.

3.2.7 Sýnidæmi

Finnið brúunarmargliðuna gegnum punktana $(1, 1)$, $(2, 3)$ og $(3, 6)$ með því að nota Lagrange-margliður.

Reiknum fyrst margliðurnar ℓ_0, ℓ_1 og ℓ_2:

$$\begin{aligned} \ell_0 &= \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{(x-2)(x-3)}{2} \\ \ell_1 &= \frac{(x-1)(x-3)}{(2-1)(2-3)} = -(x-1)(x-3) \\ \ell_2 &= \frac{(x-1)(x-2)}{(3-1)(3-2)} = \frac{(x-1)(x-2)}{2} \end{aligned}$$

Þá fæst að brúunarmargliðan p er

$$p(x) = 1 \cdot \frac{(x-2)(x-3)}{2} - 3 \cdot (x-1)(x-3) + 6 \cdot \frac{(x-1)(x-2)}{2}$$

Þetta er greinilega annars stigs margliða og auðvelt er að sannfæra sig um að $p(1) = 1$, $p(2) = 3$ og $p(3) = 6$.

3.3 Margliðubrúun: Newton-form

3.3.1 Formúla fyrir $c_0, \dots, c_m$

Nú ætlum við að leiða út formúlu fyrir stuðlunum $c_0, \dots, c_m$ í Newton-formi brúunarmargliðunnar p miðað við röð brúunarpunktanna $x_0, \dots, x_{m-1}$.

Athugum að $c_m = a_m$, þar sem a_m er stuðullinn við veldið x^m í staðalframsetningunni á p.

Til þess að reikna út $c_0, \dots, c_m$ þurfum við að reikna út með skipulegum hætti stuðulinn við veldið x^j í brúunarmargliðunni gegnum punktana $(x_i, y_i), \dots, (x_{i+j}, y_{i+j})$, fyrir öll $i = 0, \dots, m$ og $j = 0, \dots, m-i$. Við táknum þennan stuðul með $y[x_i, \dots, x_{i+j}]$.

Aðvörun: Verkefnið er háð röð punktanna, þ.e. framsetningin (Newton-formið) á margliðunni breytist eftir röð punktanna. En auðvitað er margliðan og gildin á henni alltaf þau sömu
Dæmi: Skoðum margliðuna $p(x) = 2 - 7x + 5x^2$.
Ef $x_0 = 0$ og $x_1 = 2$ þá er Newton-form hennar

$$p(x) = 3 + 3(x - 0) + 5(x - 0)(x - 2).$$

En ef $x_0 = 2$ og $x_1 = 0$ þá er Newton-form hennar

$$p(x) = 8 + 3(x - 2) + 5(x - 2)(x - 0).$$

3.3.2 Mismunakvótar

Skilgreinum mismunakvóta $y[x_i, \dots, x_{i+j}]$ fyrir punktasafnið $(x_i, y_i), \dots, (x_{i+j}, y_{i+j})$ á eftirfarandi hátt:

- $j = 0$: $y[x_i] = y_i$.
- $j = 1$: $y[x_i, x_{i+1}] = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$
- $j = 2$: $y[x_i, x_{i+1}, x_{i+2}] = \frac{y[x_{i+1}, x_{i+2}] - y[x_i, x_{i+1}]}{x_{i+2} - x_i}$.
- $j > 2$: $y[x_i, \dots, x_{i+j}] = \frac{y[x_{i+1}, \dots, x_{i+j}] - y[x_i, \dots, x_{i+j-1}]}{x_{i+j} - x_i}$.

Athugasemd: Stærðin $y[x_{n-1}, x_n]$ hefur komið fyrir áður hjá okkur þegar við fjölluðum um *Sniðilsaðferð*, enda er sniðill ekkert annað en brúunarmargliða fyrir tvö punkta í planinu.

3.3.3 Upprifjun á tilvistarsönnuninni

Þrepunarskrefið í tilvistarsönnuninni fyrir brúunarmargliður gefur okkur nú hvernig mismunakvótarnir nýtast okkur.

Látum q vera brúunarmargliðuna af stigi $\leq m - 1$ fyrir punktana $(x_0, y_0), \dots, (x_{m-1}, y_{m-1})$ og r vera brúunarmargliðuna af stigi $\leq m - 1$ fyrir punktana $(x_1, y_1), \dots, (x_m, y_m)$ og setjum síðan

$$p(x) = \frac{x - x_m}{x_0 - x_m} q(x) + \frac{x - x_0}{x_m - x_0} r(x)$$

Gerum nú ráð fyrir að stuðullinn við veldið x^{m-1} í $q(x)$ sé $y[x_0, \dots, x_{m-1}]$ og stuðullinn við veldið x^{m-1} í $r(x)$ sé $y[x_1, \dots, x_m]$.

Við sjáum þá að stuðullinn við veldið x^m í $p(x)$ er

$$\frac{y[x_0,\dots,x_{m-1}]}{x_0-x_m}+\frac{y[x_1,\dots,x_m]}{x_m-x_0}=y[x_0,\dots,x_m]$$

Athugasemd: Fyrir $m=0$ gildir að $p(x)=y_0=y[x_0]$.

3.3.4 Mismunakvótatöflur fyrir $m = 0, 1, 2, 3$

Mismunakvótar eru venjulega reiknaðir út í svokölluðum *mismunakvótatöflum*.

Ef $m=0$ er mismunakvótataflan aðeins ein lína

i	x_i	$y[x_i]$
0	x_0	$y[x_0]=y_0$

Ef $m=1$ er taflan

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$
0	x_0	$y[x_0]=y_0$	$y[x_0,x_1]$
1	x_1	$y[x_1]=y_1$	

og margliðan er

$$p(x)=y[x_0]+y[x_0,x_1](x-x_0).$$

Ef $m=2$ verður taflan

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$	$y[x_i,x_{i+1},x_{i+2}]$
0	x_0	$y[x_0]=y_0$	$y[x_0,x_1]$	$y[x_0,x_1,x_2]$
1	x_1	$y[x_1]=y_1$	$y[x_1,x_2]$	
2	x_2	$y[x_2]=y_2$		

og margliðan er

$$p(x)=y[x_0]+y[x_0,x_1](x-x_0)+y[x_0,x_1,x_2](x-x_0)(x-x_1).$$

Skoðum loks tilfellið $m=3$

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$	$y[x_i,x_{i+1},x_{i+2}]$	$y[x_i,x_{i+1},x_{i+2},x_{i+3}]$
0	x_0	$y[x_0]=y_0$	$y[x_0,x_1]$	$y[x_0,x_1,x_2]$	$y[x_0,x_1,x_2,x_3]$
1	x_1	$y[x_1]=y_1$	$y[x_1,x_2]$	$y[x_1,x_2,x_3]$	
2	x_2	$y[x_2]=y_2$	$y[x_2,x_3]$		
3	x_3	$y[x_3]=y_3$			

Brúunarmargliðan fæst svo með því að nota stuðlana úr fyrstu línu töflunnar:

$$\begin{aligned}p(x)=&y[x_0]+y[x_0,x_1](x-x_0)+y[x_0,x_1,x_2](x-x_0)(x-x_1)\\&+y[x_0,x_1,x_2,x_3](x-x_0)(x-x_1)(x-x_2)\end{aligned}$$

3.3.5 Sýnidæmi

Við skulum reikna út aftur brúunarmargliðuna gegnum $(1,1)$, $(2,3)$ og $(3,6)$.

Stillum fyrst upp mismunakvótatöflu

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$	$y[x_i,x_{i+1},x_{i+2}]$
0	1	1		
1	2	3		
2	3	6		

Fyllum svo út í hana með að ganga á hvern dálk á fætur öðrum

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$	$y[x_i,x_{i+1},x_{i+2}]$
0	1	1	$\dfrac{3-1}{2-1}=2$	$\dfrac{3-2}{3-1}=1/2$
1	2	3	$\dfrac{6-3}{3-2}=3$	
2	3	6		

Lesum út brúunarmargliðuna p með að ganga á efstu línuna:

$$p(x)=1+2\cdot(x-1)+\frac{1}{2}\cdot(x-1)(x-2).$$

Reiknum út brúunarmargliðuna gegnum $(3,1)$, $(1,-3)$, $(5,2)$ og $(6,4)$. Stillum upp og fyllum út í mismunakvótatöflu:

i	x_i	$y[x_i]$,	$y[x_i,x_{i+1}]$,	$y[x_i,x_{i+1},x_{i+2}]$,	$y[x_i,\dots,x_{i+3}]$
1	3	1	$\frac{-3-1}{1-3}=2$	$\frac{5/4-2}{5-3}=-3/8$	$\frac{3/20-(-3/8)}{6-3}=7/40$
2	1	-3	$\frac{2-(-3)}{5-1}=5/4$	$\frac{2-5/4}{6-1}=3/20$	
3	5	2	$\frac{4-2}{6-5}=2$		
4	6	4			

Nú getum við lesið brúunarmargliðuna okkar úr töflunni með að ganga á efstu línuna, við fáum

$$p(x)=1+2(x-3)-\frac{3}{8}(x-3)(x-1)+\frac{7}{40}(x-3)(x-1)(x-5)$$

3.3.6 Samantekt

Ef gefnir eru punktar $(x_0,y_0),\dots,(x_m,y_m)$ í $\mathbb{R}^2$, þar sem $x_i\neq x_j$ ef $i\neq j$, þá er til nákvæmlega ein margliða p af stigi $\leq m$ þannig að

$$p(x_k)=y_k,\quad k=0,\dots,m$$

Newton-form margliðunnar p með tilliti til punktanna $x_0,\dots,x_{m-1}$ er

$$p(x)=y[x_0]+y[x_0,x_1](x-x_0)+\cdots+y[x_0,\dots,x_m](x-x_0)\cdots(x-x_m)$$

þar sem mismunakvótarnir eru reiknaðir með rakningarformúlunum $y[x_i]=y_i$ og

$$y[x_i,\dots,x_{i+j}]=\frac{y[x_{i+1},\dots,x_{i+j}]-y[x_i,\dots,x_{i+j-1}]}{x_{i+j}-x_i},\qquad i=0,\dots,m,\quad j=0,\dots,m-i.$$

3.3.7 Samantekt – Newton-form

Venja er að setja mismunakvótana upp í töflu og stuðlarnir í Newton-forminu raða sér í fyrstu línu töflunnar:

i	x_i	$y[x_i]$	$y[x_i,x_{i+1}]$	$y[x_i,x_{i+1},x_{i+2}]$	$y[x_i,\dots,x_{i+3}]$	$y[x_i,\dots,x_{i+4}]$	$\dots$
0	x_0	$y[x_0]=y_0$	$y[x_0,x_1]$	$y[x_0,x_1,x_2]$	$y[x_0,x_1,x_2,x_3]$	$y[x_0,x_1,x_2,x_3,x_4]$	$\dots$
1	x_1	$y[x_1]=y_1$	$y[x_1,x_2]$	$y[x_1,x_2,x_3]$	$y[x_1,x_2,x_3,x_4]$	$\dots$	
2	x_2	$y[x_2]=y_2$	$y[x_2,x_3]$	$y[x_2,x_3,x_4]$	$\dots$		
3	x_3	$y[x_3]=y_3$	$y[x_3,x_4]$	$\dots$			
4	x_4	$y[x_4]=y_4$	$\dots$				
$\vdots$	$\vdots$	$\vdots$					

3.3.8 Samantekt – Lagrange-margliður

Lagrange-form brúunarmargliðunnar er

$$p(x) = \sum_{k=0}^{m} y_k \ell_k(x)$$

þar sem ℓ_k eru Lagrange-margliðurnar með tilliti til punktanna $x_0, \ldots, x_m$,

$$\ell_k(x) = \prod_{\substack{j=0 \\ j\neq k}}^{m} \frac{(x - x_j)}{(x_k - x_j)} = \frac{(x - x_0)\cdots(x - x_{k-1})(x - x_{k+1})\cdots(x - x_m)}{(x_k - x_0)\cdots(x_k - x_{k-1})(x_k - x_{k+1})\cdots(x_k - x_m)}.$$

$$Enruppfylla$$

$$\ell_k(x_i) = \begin{cases} 1 & \text{ef } i = k \\ 0 & \text{ef } i \neq k \end{cases}$$

3.4 Samantekt

3.4.1 Lagrange-margliður

- Auðvelt að finna margliðuna
- Dýrara að reikna fallgildin

3.4.2 Newton-margliður

- Erfiðara að finna margliðuna
- Auðvelt að finna fallgildin (reiknirit Horners)

3.5 Margliðubrúun: Margfaldir punktar

Látum $a_1, \ldots, a_k$ vera ólíka punkta í $\mathbb{R}$, $m_1, \ldots, m_k$ vera jákvæðar heiltölur og hugsum okkur að gefnar séu rauntölur

$$y_i^{(j)}, \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, k.$$

Við viljum finna margliðu p af lægsta mögulega stigi þannig að margliðan $p = p^{(0)}$ og afleiður hennar $p^{(j)}$ uppfylli

$$p^{(j)}(a_i) = y_i^{(j)}, \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, k$$

Við nefnum verkefnið að finna slíka margliðu p *alhæft brúunarverkefni*, og margliða sem uppfyllir þessi skilyrði nefnist *brúunarmargliða fyrir brúunarverkefnið* sem lýst er með gefnu skilyrðunum.

3.5.1 Margfeldni punktanna

Við segjum að a_i sé *einfaldur brúunarpunktur* ef $m_i = 1$, *tvöfaldur brúunarpunktur* ef $m_i = 2$ o.s.frv.

Við skilgreinum nú töluna

$$m = m_1 + m_2 + \ldots + m_k - 1.$$

Brúunarmargliðan okkar p á að vera af stigi $\leq m$, og fjöldi skilyrða sem við setjum á hana eru $m+1$.

Athugasemd: Tilfellið $k = m+1$, $m_j = 1$ er það sem við skoðuðum hér á undan.

3.5.2 Sértilfelli

Tvö sértilfelli þekkjum við nú þegar.

1. *Allir punktar eins:* Ef allir punktarnir eru einfaldir, þá er alhæfða brúunarverkefnið sama verkefni og brúunarverkefnið sem við leystum í Margliðubrúun: Lagrange-form og Margliðubrúun: Newton-form.

$$p^{(0)}(a_i) = p(a_i) = y_i^{(0)},$$

   og lausnin var leidd út með

$$x_0 = a_1, \ldots, x_m = a_k \quad \text{og} \quad y_0 = y_1^{(0)}, \ldots, y_m = y_k^{(0)}.$$

2. *Einn punktur:* Ef aftur á móti $k = 1$, þá er lausn gefin með Taylor-margliðunni af röð m í punktinum a_1

$$p(x) = y_1^{(0)} + \frac{y'}{1!}(x-a_1) + \frac{y''}{2!}(x-a_1)^2 + \ldots + \frac{y_1^{(m)}}{m!}(x-a_1)^m.$$

3.5.3 Upprifjun

Munum að ef p er margliða og $p(a) = 0$ þá er p deilanleg með $(x-a)$. Það er, hægt er að skrifa

$$p(x) = (x-a)q(x),$$

þar sem q er margliða af stigi sem er einu lægra en stig p (sjá Undirstöðusetning algebrunnar).

3.5.4 Ótvíræðni lausnarinnar

Nú ætlum við að sýna fram á að til sé nákvæmlega ein margliða $p(x)$ af stigi $\leq m$ sem uppfyllir

$$p^{(j)}(a_i) = y_i^{(j)}, \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, k$$

Byrjum á að sýna að það er í mesta lagi ein margliða sem uppfyllir þetta.

Gerum ráð fyrir að $p(x)$ og $q(x)$ séu tvær margliður af stigi $\leq m$ sem uppfylla öll þessi skilyrði.

Þá uppfyllir margliðan $r(x) = p(x) - q(x)$ að

$$r^{(j)}(a_i) = 0, \quad j = 0, \ldots, m_i - 1, \quad i = 1, \ldots, k$$

Af þessu leiðir að $r(x)$ er deilanlegt með $(x-a_i)^{m_i}$ en samanlagt stig þessara þátta er $m_1 + \ldots + m_k = m+1$.

Nú er stig $r(x)$ minna eða jafnt m svo þetta getur aðeins gerst ef $r(x)$ er núllmargliðan.

Við höfum því að $p(x) = q(x)$ og ályktum að við höfum nákvæmlega eina lausn á brúunarverkefninu ef við getum sýnt fram á tilvist á lausn.

3.5.5 Tilvist á lausn

Nú beitum við sams konar röksemdafærslu og í byrjum kaflans til þess að sýna fram á tilvist á lausn, þ.e. við notum þrepun.

Smíðum margliðuna:

Ef $m = 0$, þá er lausnin fastamargliðan $p(x) = y_1^{(0)} = y_0$.

Gerum nú ráð fyrir að við getum fundið brúunarmargliðu af stigi $\leq m-1$ fyrir sérhvert alhæft brúunarverkefni þar sem samanlagður fjöldi skilyrðanna er m.

Lítum nú aftur á upprunalega brúunarverkefnið þar sem fjöldi skilyrðanna er $m+1$. Skilgreinum tvær runur af punktum

$$(x_0, x_1, \dots, x_m) = (\underbrace{a_1, \dots, a_1}_{m_1 \text{ sinnum}}, \underbrace{a_2, \dots, a_2}_{m_2 \text{ sinnum}}, \dots, \underbrace{a_k, \dots, a_k}_{m_k \text{ sinnum}})$$

og

$$(y_0, y_1, \dots, y_m) = (y_1^{(0)}, \dots, y_1^{(m_1-1)}, \dots, y_k^{(0)}, \dots, y_k^{(m_k-1)})$$

Við höfum séð að í því tilfelli að við höfum einn punkt, $k = 1$, $x_0 = x_1 = \dots = x_m = a_1$ er lausnin gefin með Taylor-margliðu í a_1.

Við megum því gera ráð fyrir punktarnir séu a.m.k. tveir, $k \geq 2$. Það gefur að $x_0 \neq x_m$.

Látum $q(x)$ vera margliðuna af stigi $\leq m-1$ sem uppfyllir sömu skilyrði og p, nema það síðasta um að $q^{(m_k-1)}(a_k)$ þurfi að vera $y_k^{(m_k-1)}$.

og látum $r(x)$ vera margliðuna sem uppfyllir öll brúunarskilyrðin, nema síðasta skilyrðið í fyrsta punkti um að $r^{(m_1-1)}(a_1)$ sé jafnt $y_1^{(m_1-1)}$.

Setjum síðan

$$p(x) = \frac{x - x_m}{x_0 - x_m} q(x) + \frac{x - x_0}{x_m - x_0} r(x) = \frac{x - a_k}{a_1 - a_k} q(x) + \frac{x - a_1}{a_k - a_1} r(x)$$

Sýnum að gefin fallgildi eru tekin

Nú þurfum við að staðfesta að öll skilyrðin séu uppfyllt.

Við byrjum á því að taka $j = 0$ sem svarar til þess að p taki fyrirfram gefin fallgildi,

$$\begin{aligned}
p(a_1) &= \frac{a_1 - a_k}{a_1 - a_k} q(a_1) + \frac{a_1 - a_1}{a_k - a_1} r(a_1) = q(a_1) = y_1^{(0)} \\
p(a_i) &= \frac{a_i - a_k}{a_1 - a_k} q(a_i) + \frac{a_i - a_1}{a_k - a_1} r(a_i) = \left(\frac{a_i - a_k}{a_1 - a_k} + \frac{a_i - a_1}{a_k - a_1} \right) y_i^{(0)} \\
&= y_i^{(0)}, \qquad \text{fyrir } i = 2, \dots, k-1, \\
p(a_k) &= \frac{a_k - a_k}{a_1 - a_k} q(a_k) + \frac{a_k - a_1}{a_k - a_1} r(a_k) = r(a_k) = y_k^{(0)}.
\end{aligned}$$

Sýnum svo að gildin á afleiðum p séu rétt.

Rifjum upp margliðuna p:

$$p(x) = \frac{x - x_m}{x_0 - x_m} q(x) + \frac{x - x_0}{x_m - x_0} r(x) = \frac{x - a_k}{a_1 - a_k} q(x) + \frac{x - a_1}{a_k - a_1} r(x)$$

Afleiður hennar eru

$$p^{(j)}(x) = \frac{(x - a_k)}{(a_1 - a_k)} q^{(j)}(x) + \frac{(x - a_1)}{(a_k - a_1)} r^{(j)}(x) + j \frac{\left(q^{(j-1)}(x) - r^{(j-1)}(x) \right)}{a_k - a_1}$$

Ef nú $m_i > 1$ þá er $q^{(j-1)}(a_i) = y^{(j-1)}(a_i) = r^{(j-1)}(a_i)$ fyrir $j = 1, \dots, m_i - 1$ og því kemur alltaf 0 út úr síðasta liðnum ef við setjum inn $x = a_i$, fyrir öll $i = 1, \dots, k$.

Af þessu sést að afleiður p uppfylla skilyrðin

$$p^{(j)}(a_i) = y_i^{(j)}, \qquad \text{fyrir } j = 0, \dots, m_i - 1, \quad i = 1, \dots, k.$$

3.5.6 Samantekt

Við höfum því sannað eftirfarandi:

Ef gefnar eru

- rauntölur $a_1, \dots, a_k$, með $a_j \neq a_k$ ef $j \neq k$,
- jákvæðar heiltölur $m_1, \dots, m_k$,
- rauntölur $y_i^{(j)}$, fyrir $j = 0, \dots, m_i - 1$, $i = 1, \dots, k$,

og talan m er skilgreind með $m = m_1 + \cdots + m_k - 1$, þá er til nákvæmlega ein margliða p af stigi $\leq m$ þannig að

$$p^{(j)}(a_i) = y_i^{(j)}, \qquad j = 0, \dots, m_i - 1, \quad i = 1, \dots, k.$$

3.5.7 Brúunarmargliðan fundin

Ef skilgreindar eru runurnar

$$(x_0, \dots, x_m) = (a_1, \dots, a_1, a_2, \dots, a_2, \dots, a_k, \dots, a_k)$$

þar sem a_1 kemur fyrir m_1 sinnum, a_2 kemur fyrir m_2 sinnum o.s.frv., og

$$(y_0, \dots, y_m) = (y_1^{(0)}, \cdot\cdot, y_1^{(m_1-1)}, y_2^{(0)}, \cdot\cdot, y_2^{(m_2-1)}, \cdots, y_k^{(0)}, \cdot\cdot, y_k^{(m_k-1)}),$$

þá er Newton-form margliðunnar p með tilliti til punktanna $x_0, \dots, x_{m-1}$ gefið með

$$p(x) = y[x_0] + y[x_0, x_1](x - x_0) + \cdots + y[x_0, \dots, x_m](x - x_0) \cdots (x - x_{m-1})$$

þar sem mismunakvótarnir $y[x_i, \dots, x_{i+j}]$ eru reiknaðir með rakningarformúlu þannig að $y[x_i] = y_i$ og

$$y[x_i, \dots, x_{i+j}] = \begin{cases} \dfrac{y[x_{i+1}, \dots, x_{i+j}] - y[x_i, \dots, x_{i+j-1}]}{x_{i+j} - x_i}, & \text{ef } x_i \neq x_{i+j}, \\ \dfrac{y_i^{(j)}}{j!}, & \text{ef } x_i = x_{i+j}. \end{cases}$$

3.6 Skynsamlegir skiptipunktar og Chebyshev margliður

3.6.1 Um val á brúunarpunktum

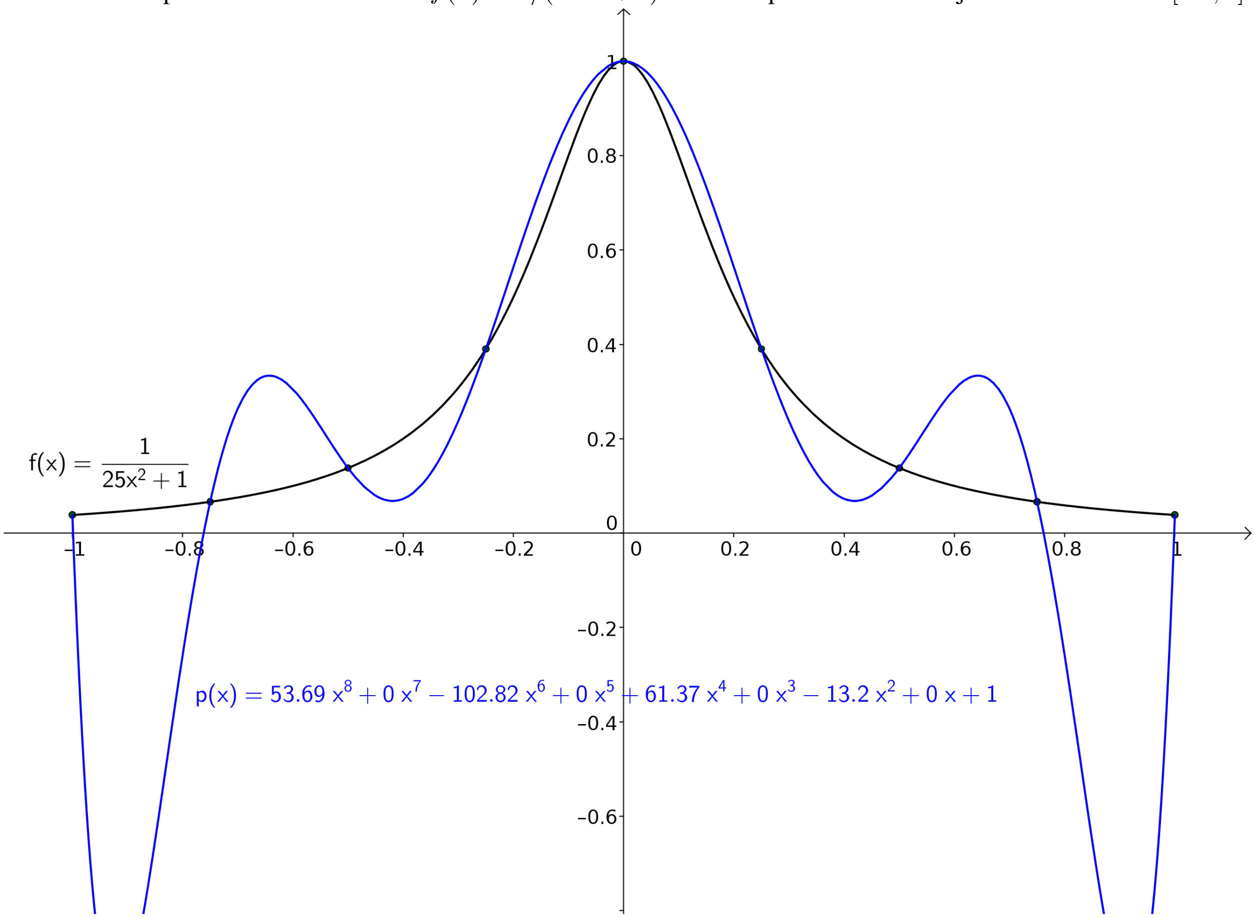
Látum $(t_0, y_0), \dots, (t_n, y_n)$ vera punkta í plani og gerum ráð fyrir að $a = t_0 < t_1 < \cdots < t_n = b$.

Við höfum nú lært að ákvarða margliðu p af stigi $\leq n$ sem tekur gildin y_i í punktunum t_i.

Ef punktarnir liggja á grafi fallsins f og nota á margliðuna til þess að nálga fallgildi f, þá getur það verið ýmsum erfiðleikum bundið þegar stig hennar stækkar. Til dæmis getur komið fram óstöðugleiki í útreikningum þannig að örlítil frávik í x geta leitt til mikilla frávika í $p(x)$, og þá hugsanlega í mikilli skekkju á $f(x) - p(x)$.

3.6.2 Dæmi um óheppilega skiptipunkta

Skoðum dæmi þar sem við brúum fallið $f(x) = 1/(25x^2 + 1)$ í 9 brúnarpunktum sem eru jafndreifðir á bilinu $[-1, 1]$.



Hér sjáum við „þægilegt” fall þar sem brúunarmargliðan gefur afskaplega vonda nálgun.

3.6.3 Val á brúunarpunktum

Það er ekki sjálfgefið að við getum valið í hvaða brúunarpunkta við notum, t.d. ef þeir ákvarðast af mælingum. Ef við hins vegar getum valið þá óhindrað, þá vaknar sú spurning hvernig er best að gera það?

3.6.4 Skilgreining

Fyrst þurfum við að útskýra betur hvað við eigum við með „best”. Við munum bara notast við tvær leiðir hér til að mæla skekkjuna, en það er ℓ_∞ og ℓ_2 staðlarnir, fyrir samfellt fall h á bilinu $[a, b]$ þá eru þeir skilgreindir með

$$\|h\|_\infty = \max_{x \in [a,b]} |h(x)|,$$

og

$$\|h\|_2 = \left(\int_a^b h(x)^2 \, dx \right)^{\frac{1}{2}}$$

Athugasemd: Það má líta þannig á þetta að ℓ_∞ staðallinn mæli hámarksskekkju og ℓ_2 mæli einhvers konar„meðaltalsskekkju", þar sem meðaltalið er reiknað með heildi.

3.6.5 Verkefnið

Verkefnið er því eftirfarandi: Fyrir gefið fall $f(x)$ á bili $[a, b]$ og fast n, þá viljum við finna $x_0, \dots, x_n$ sem lágmarka annað hvort

$$\|f - p\|_\infty \quad \text{eða} \quad \|f - p\|_2.$$

Þar sem p er brúunarmargliðan fyrir brúunarpunktana $(x_i, f(x_i))$.

Byrjum á að skoða ℓ_∞ tilvikið.

3.6.6 Skilgreining: Chebyshev margliður

Fyrir náttúrlega tölu n þá skilgreinum við *Chebyshev margliðuna* T_n á $[-1, 1]$ með

$$T_n(x) = \cos(n \arccos(x)).$$

Með því að setja inn $n = 0$ og $n = 1$ þá fæst að

$$T_0(x) = 1 \quad \text{og} \quad T_1(x) = x,$$

og með hornafallareglunum fæst að

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), n \geq 1.$$

Af jöfnunni hér á undan þá fæst með þrepun að

- $T_n(x)$ er margliða af stigi n.
- Forystustuðull T_n er 2^{n-1}.
- T_n er jafnstæð ef n er slétt og oddstæð ef n er oddatala.

3.6.7 Setning

Chebyshev margliðan T_n hefur n einfaldar núllstöðvar á bilinu $[-1, 1]$ og þær eru gefnar með

$$x_j = \cos\left(\frac{2j+1}{2n}\pi\right), \qquad j = 0, 1, 2, 3, \dots, n-1.$$

Auk þess eru útgildi T_n á $[-1, 1]$ staðsett í

$$z_j = \cos\left(\frac{j\pi}{n}\right), \qquad j = 0, 1, 2, \dots, n,$$

og fallgildin þar uppfylla $T_n(z_j) = (-1)^j$

3.6.8 Staðlaðar Chebyshev margliður

Margliða er kölluð *stöðluð* ef forystustuðull hennar er 1.

Stöðluðu Chebyshev margliðurnar $\tilde{T}$ eru skilgreindar á eftirfarandi hátt

$$\tilde{T}(x) = \begin{cases} T_0(x) & \text{ef } n = 0 \\ 2^{1-n} T_n(x) & \text{ef } n \geq 1 \end{cases}$$

Fyrir sérhverja staðlaða margliðu q af stigi n þá er

$$\frac{1}{2^{n-1}} = \max_{x \in [-1,1]} T_n(x) \leq \max_{x \in [-1,1]} |q(x)|.$$

Þ.e. af öllum stöðluðum margliðum þá eru stöðluðu Chebyshev margliðurnar „minnstar" á bilinu $[-1, 1]$.

3.6.9 Skynsamlegir skiptipunktar fyrir bilið $[-1, 1]$

Við vitum að skekkjan í því að nálga fallið f með brúunarmargliðu p með brúunarpunkta $x_0, \ldots, x_n$ er

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1) \cdots (x - x_n),$$

þar sem ξ er á minnsta bilinu sem inniheldur x og $x_0, x_1, \ldots, x_n$. Ef við skoðum jöfnuna að ofan þá sjáum við að þar sem n og f (og þar með $f^{(n+1)}$) er fast þá er stæðan $(x - x_0) \cdots (x - x_n)$ það eina sem við höfum einhverja stjórn á.

Með því að nota Chebyshev margliðurnar þá getum við lágmarkað þennan hluta skekkjunnar.

Athugið að $(x - x_0) \cdots (x - x_n)$ er stöðluð margliða af stigi $n + 1$. Þannig að samkvæmt því sem kom fram hér að ofan þá lágmörkum við framlag hennar til skekkjunnar með $(x - x_0) \cdots (x - x_n) = \tilde{T}_{n+1}$, þ.e. með því að velja

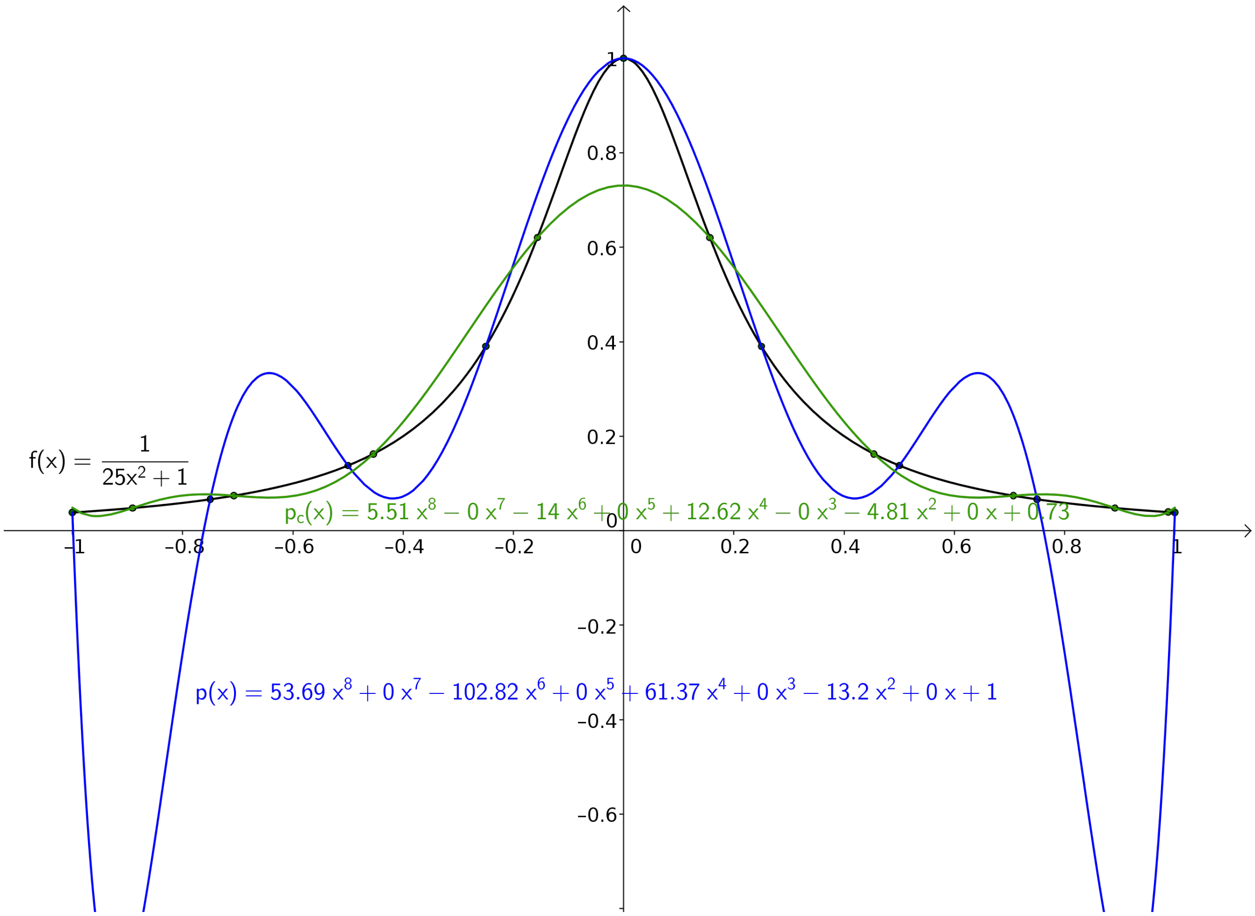
$$x_i = \cos\left(\frac{2i+1}{2(n+1)}\pi\right), \qquad i = 0, 1, \ldots, n.$$

Hæsta gildi $\tilde{T}_{n+1}$ er $\frac{1}{2^n}$, sem þýðir að við fáum skekkjumatið

$$\|f(x) - p(x)\|_\infty \leq \frac{\|f^{(n+1)}\|_\infty}{2^n (n+1)!}.$$

3.6.10 Dæmi um óheppilega skiptipunkta skoðað aftur

Skoðum aftur fallið $f(x) = 1/(25x^2 + 1)$, en í stað þess að taka 9 jafndreifaða brúunarpunkta á bilinu $[-1, 1]$, þá skulum við nota Chebyshev margliðurnar til að finna 9 punkta á bilinu.



3.6.11 Skynsamlegir skiptipunktar fyrir bil $[a, b]$

Hér á undan miðaðist allt við að finna brúunarmargliðu fyrir fallið f á bilinu $[-1, 1]$. Ef við viljum skoða almennt bil $[a, b]$ þá byrjum við á athuga að fallið $\eta : [-1, 1] \to [a, b]$,

$$\eta(t) = \frac{b-a}{2}t + \frac{b+a}{2}$$

skilgreinir línulega vörpun (hliðrun og stríkkun) frá $[-1, 1]$ yfir á $[a, b]$. Athugið að vörpunin sendir -1 í a og 1 í b.

Með því að taka rætur stöðluðu Chebyshev margliðunnar $\tilde{T}_{n+1}$ og varpa þeim með η yfir á bilið $[a, b]$ þá fáum við þá punkta $x_0, \dots, x_n \in [a, b]$ sem lágmarka $(x - x_0) \cdots (x - x_n)$ á bilinu $[a, b]$,

$$x_i = \eta\left(\cos\left(\frac{2i+1}{2(n+1)}\pi\right)\right) = \frac{b-a}{2}\cos\left(\frac{2i+1}{2(n+1)}\pi\right) + \frac{b+a}{2},$$

$i = 0, 1, 2, \dots, n.$

3.6.12 Lágmörkun á skekkju með tilliti til ℓ_2

Nú skulum við skipta um staðal, þannig að í stað þess að lágmarka $\|f - p\|_\infty$ þá skulum við reyna að lágmarka

$$\|f - p\|_2 = \left(\int_a^b (f - p)^2\, dx\right)^{1/2}$$

Við vitum að skekkjan í því að nálga fallið f með brúunarmargliðu p með brúunarpunkta $x_0, \dots, x_n$ er

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)(x - x_1)\cdots(x - x_n),$$

þar sem ξ er á minnsta bilinu sem inniheldur x og $x_0, x_1, \dots, x_n$.

Eins og áður þá sjáum við að stæðan $(x - x_0)\cdots(x - x_n)$ það eina sem við getum stjórnað með því að velja brúunarpunktana x_j.

3.6.13 Skilgreining: Legendre margliðurnar

Fyrir náttúrlega tölu n þá skilgreinum við *Legendre margliðurnar* svona

$$\begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_n(x) &= \frac{2n-1}{n} x P_{n-1}(x) - \frac{n-1}{n} P_{n-2}(x). \end{aligned}$$

Af skilgreiningunni hér á undan þá sjáum við að

- $P_n(x)$ er margliða af stigi n.
- Forystustuðull P_n er $\frac{2n-1}{n} \cdot \frac{2n-3}{n-2} \cdots \frac{3}{2} \cdot 1$.
- P_n er jafnstæð ef n er slétt og oddstæð ef n er oddatala.

3.6.14 Setning

$$\int_{-1}^{1} P_j(x) P_k(x)\, dx = \begin{cases} 0, & \text{ef } j \neq k \\ \frac{2}{2j+1}, & \text{ef } j = k. \end{cases}$$

Einnig gildir að ef q er margliða af stigi minna en n þá er

$$\int_{-1}^{1} q(x) P_n(x)\, dx = 0.$$

Þetta segir okkur að Legendre margliðurnar eru hornréttar (með tilliti til innfeldisins sem heildið skilgreinir).

3.6.15 Setning

P_n hefur n ólíkar núllstöðvar sem liggja allar á $[-1, 1]$.

3.6.16 Skilgreining: Staðlaðar Legendre margliður

Eins og þegar við fengumst við Chebyshev margliðurnar þá skilgreinum við *stöðluðu Legendre margliðurnar* $\tilde{P}_n$ með því að deila upp í P_n með forrystustuðlunum P_n.

Athugasemd: Setningarnar þrjár hér undan gilda um $\tilde{P}$ alveg eins og P.

3.6.17 Setning: Lágmörkun með Legendre margliðunum

Ef p er stöðluð margliða af stigi $n+1$ þá er $\|p\|_2 \geq \|\tilde{P}_{n+1}\|_2$.

Skilgreinum $q = p - \tilde{P}_{n+1}$, sem þýðir að q er margliða af stigi minna en $n+1$. Nú er

$$\begin{aligned}
\|p\|_2^2 &= \|\tilde{P}_{n+1} + q\|_2^2 \\
&= \int_{-1}^{1} (\tilde{P}_{n+1}(x) + q(x))^2 \, dx \\
&= \int_{-1}^{1} \tilde{P}_{n+1}(x)^2 + 2q(x)\tilde{P}_{n+1}(x) + q(x)^2 \, dx \\
&= \|\tilde{P}_{n+1}\|_2^2 + 2\int_{-1}^{1} q(x)\tilde{P}_{n+1}(x)\, dx + \|q\|_2^2 \\
&= \|\tilde{P}_{n+1}\|_2^2 + \|q\|_2^2 \geq \|\tilde{P}_{n+1}\|_2^2
\end{aligned}$$

því $\int_{-1}^{1} q(x)\tilde{P}_{n+1}(x)\, dx = 0$ og $\|q\|_2 \geq 0$.

Af síðustu setningu sjáum við að til þess að lágmarka

$$\|f(x) - p(x)\|_2 = \left\| \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)(x - x_1)\cdots(x - x_n) \right\|_2,$$

þá veljum við $x_1, \ldots, x_n$ þannig að $(x - x_0)(x - x_1)\cdots(x - x_n) = \tilde{P}_{n+1}$. Þ.e. x_j þurfa að vera rætur stöðluðu Legendre margliðunnar af stigi $n+1$.

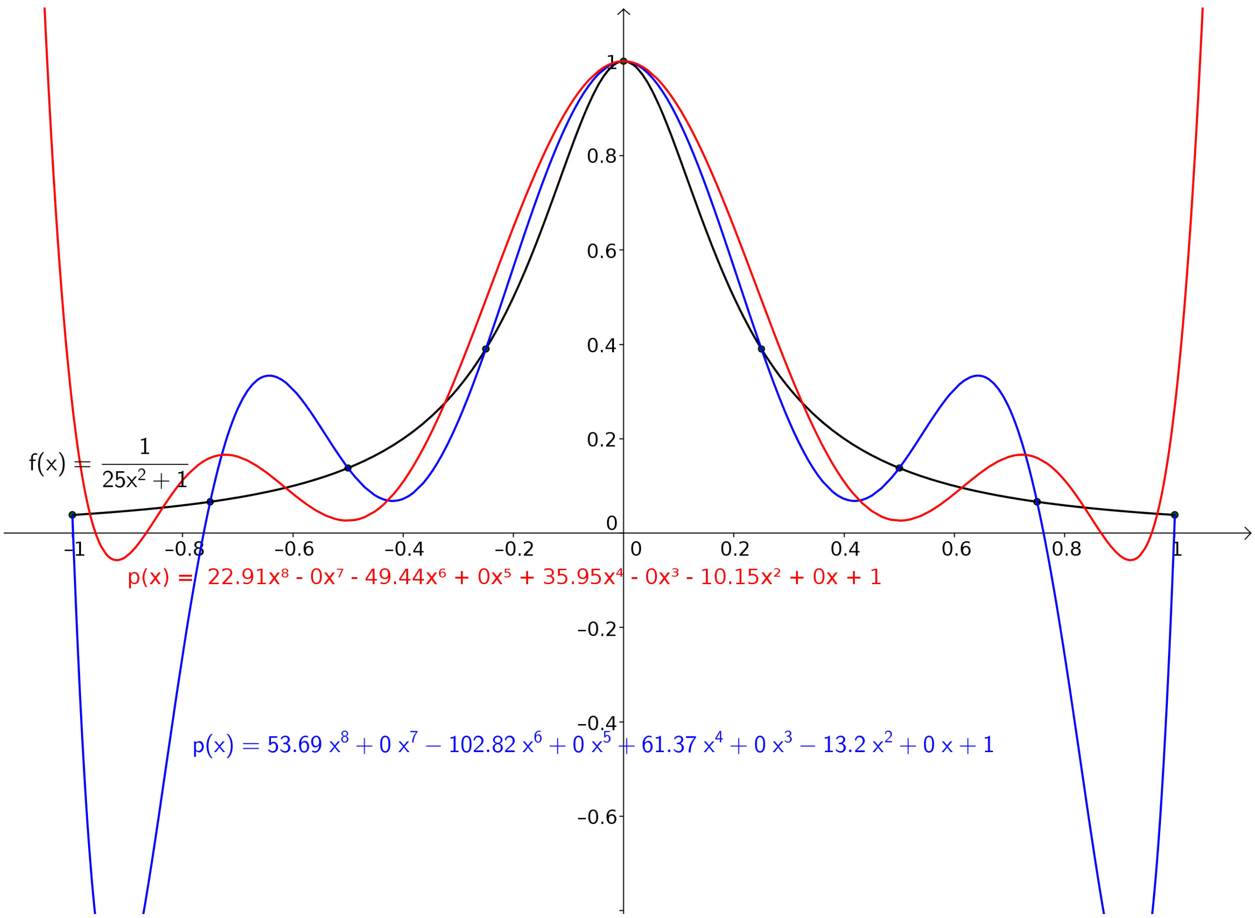
3.6.18 Núllstöðvar P_n, fyrir $n = 1, \ldots, 10$

Ólíkt Chebyshev margliðunum þá er ekki hlaupið að því að finna rætur $\tilde{P}_{n+1}$. Þannig að við þurfum að reikna þær tölulega og geyma í töflu.

k	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}
1	0.0000									
2	-0.5774	0.5774								
3	-0.7746	0.0000	0.7746							
4	-0.8611	-0.3400	0.3400	0.8611						
5	-0.9062	-0.5385	0.0000	0.5385	0.9062					
6	-0.9325	-0.6612	-0.2386	0.2386	0.6612	0.9325				
7	-0.9491	-0.7415	-0.4058	0.0000	0.4058	0.7415	0.9491			
8	-0.9603	-0.7967	-0.5255	-0.1834	0.1834	0.5255	0.7967	0.9603		
9	-0.9682	-0.8360	-0.6134	-0.3243	0.0000	0.3243	0.6134	0.8360	0.9682	
10	-0.9739	-0.8651	-0.6794	-0.4334	-0.1489	0.1489	0.4334	0.6794	0.8651	0.9739

3.6.19 Dæmi um óheppilega skiptipunkta skoðað aftur

Skoðum enn einu sinni fallið $f(x) = 1/(25x^2+1)$, en í stað þess að taka 9 jafndreifaða brúunarpunkta á bilinu $[-1, 1]$, þá skulum við nota Legendre margliðurnar til að finna 9 punkta á bilinu.



3.6.20 Athugasemd um ℓ_∞ og ℓ_2

Athugasemd: Það má líta þannig á þetta að ℓ_∞ staðallinn mæli hámarksskekkju og ℓ_2 mæli einhvers konar„heildarskekkju", þar sem skekkjan er reiknað með heildinu hér á undan og svarar því hér um bil til flatarmálsins á milli fallsins og brúunarmargliðunnar.

- ℓ_∞ staðallinn mælir hámarksskekkju, þannig að með því nota Chebyshev margliðurnar þá erum við að reyna að lágmarka mestu skekkju á bilinu.
- ℓ_2 mæli einhvers konar„heildarskekkju", þar sem skekkjan er reiknað með heildi. Þannig að með því að nota Legendre margliðurnar þá erum við í einhverjum skilningi að lágmarka flatarmál.

3.7 Skekkjumat

3.7.1 Nálgun á föllum með margliðum

Lítum nú aftur á almenna brúunarverkefnið og gefum okkur að tölurnar $y_i^{(j)}$ séu af gerðinni $f^{(j)}(a_i)$ þar sem $f : I \to \mathbb{R}$ er fall á bili I sem inniheldur alla punktana $a_1, \dots, a_k$.

Þá snýst brúunarverkefnið um að finna margliðu af stigi $\le m$ sem uppfyllir

$$p^{(j)}(a_i) = f^{(j)}(a_i), \quad j = 0, \dots, m_i - 1, \quad i = 1, \dots, k.$$

Við vitum að lausn þess er ótvírætt ákvörðuð. Ef við notum Newton form lausnarinnar, þá táknum við mismunakvótana með

$$f[x_i, \dots, x_{i+j}]$$

í stað

$$y[x_i, \dots, x_{i+j}]$$

3.7.2 Nálgun á fallgildum

Runurnar $(x_0, \dots, x_m)$ og $(y_0, \dots, y_m)$ eru skilgreindar með

$$(x_0, x_1, \dots, x_m) = (\underbrace{a_1, \dots, a_1}_{m_1 \text{ sinnum}}, \underbrace{a_2, \dots, a_2}_{m_2 \text{ sinnum}}, \dots, \underbrace{a_k, \dots, a_k}_{m_k \text{ sinnum}})$$

og

$$\begin{aligned}(y_0, y_1, \dots, y_m) = (f^{(0)}(a_1), \dots, f^{(m_1-1)}(a_1), f^{(0)}(a_2), \dots, f^{(m_2-1)}(a_2) \\ \dots, f^{(0)}(a_k), \dots, f^{(m_k-1)}(a_k))\end{aligned}$$

3.7.3 Skekkjumat

Nú tökum við punkt $x \in I$ og spyrjum um skekkjuna $f(x) - p(x)$ í nálgun á $f(x)$ með $p(x)$. Ef x er einn punktana $a_1, \dots, a_k$, þá er $p(x) = f(x)$ og skekkjan þar með 0, svo við skulum gera ráð fyrir að $x \neq a_i$, $i = 1, \dots, k$.

Við bætum nú $(x, f(x))$ sem einföldum brúunarpunkti við alhæfða brúunar verkefnið og fáum sem lausn $q(t)$ á þessu aukna verkefni. Margliðan q er af stigi $\leq m+1$. Við notum táknið t fyrir breytu, því x er frátekið.

Þá uppfyllir $q(t)$ að $q(x) = f(x)$ auk allra skilyrðanna

$$q^{(j)}(a_i) = p^{(j)}(a_i) = f^{(j)}(a_i)$$

í verkefninu sem við byrjuðum með.

Við getum þá skrifað (sjá *Newton-margliður* til hliðsjónar)

$$\begin{aligned}q(t) &= p(t) + f[x_0, \dots, x_m, x](t - x_0) \cdots (t - x_m) \\ &= p(t) + f[x_0, \dots, x_m, x](t - a_1)^{m_1} \cdots (t - a_k)^{m_k}.\end{aligned}$$

Þegar við gefum breytunni t gildið x, þá fáum við $q(x) = f(x)$ og því fæst formúla fyrir skekkjunni

$$f(x) - p(x) = f[x_0, \dots, x_m, x](x - a_1)^{m_1} \cdots (x - a_k)^{m_k}$$

Nú ætlum við að finna leið til þess að meta skekkjuliðinn. Til þess þurfum við að gefa okkur að f hafi að minnsta kosti $m+1$ afleiðu.

3.7.4 Tilfellið þegar við höfum aðeins einn punkt

Munum nú að í tilfellinu þegar við erum bara með einn punkt a_1, þá erum við með $m+1$ skilyrði

$$p^{(j)}(a_1) = f^{(j)}(a_1), \qquad j = 0, \dots, m$$

og við fáum að p er Taylor-margliða fallsins f í punktinum a_1. Þá er $x_0 = \cdots = x_m = a_1$ og við fáum

$$f(x) - p(x) = f[x_0, \dots, x_m, x](x - a_1)^{m+1}$$

Nú segir setning Taylors okkur að til sé punktur ξ milli a_1 og x þannig að

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a_1)^{m+1}$$

Við getum því dregið þá ályktun að í þessu sértilfelli er

$$f[x_0, \dots, x_m, x] = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

Það kemur í ljós að þetta er almenn regla sem gildir fyrir *öll* alhæfðu brúunarverkefnin.

Tilfellið $m = 1$ er meðalgildisreglan

Munum að tilfellið $m = 1$ er meðalgildisreglan

$$f[a_1, x] = \frac{f(x) - f(a_1)}{x - a_1} = f'(\xi).$$

3.7.5 Margfeldni núllstöðva

Samfellt fall φ á bili I er sagt hafa núllstöð *af stigi að minnsta kosti* $m > 0$ í punktinum $a \in I$, ef til er samfellt fall ψ á I þannig að

$$\varphi(x) = (x-a)^m \psi(x)$$

Við segjum að φ hafi núllstöð af *margfeldni* m ef $\psi(a) \neq 0$.

Athugið að ef φ er deildanlegt I með samfellda afleiðu, þá er ψ deildanlegt með samfellda afleiðu í $I \setminus \{a\}$ og við höfum

$$\begin{aligned}\varphi'(x) &= m(x-a)^{m-1}\psi(x) + (x-a)^m \psi'(x) \\ &= (x-a)^{m-1}\big(m\psi(x) + (x-a)\psi'(x)\big)\end{aligned}$$

Ef afleiðan ψ' er takmörkuð í grennd um a, þá sjáum við á þessari formúlu að φ' hefur núllstöð af stigi að minnsta kosti $m-1$ í a.

Hugsum okkur nú að við séum með $a_1, \dots, a_k$ ólíka punkta í bilinu I og að $m_1, \dots, m_k$ séu jákvæðar náttúrlegar tölur.

Ef fallið φ hefur núllstöðvar í öllum punktunum a_j og núllstöðin a_j er af stigi að minnsta kosti m_j. Við segjum að þá hafi φ *að minnsta kosti*

$$n = m_1 + \cdots + m_k$$

núllstöðvar taldar með margfeldni.

Eins þá segjum við að φ hafi n núllstöðvar í $\{a_1, \dots, a_k\}$ *taldar með margfeldni* ef φ hefur núllstöðvar í öllum punktum $a_1, \dots, a_k$ og samanlögð margfeldni þeirra er n

Hugsum okkur nú að fallið φ hafi núllstöð af stigi m_j í punktunum a_j fyrir öll $j = 1, \dots, k$ og að $n = m_1 + \cdots + m_k$.

Til einföldunar gerum við ráð fyrir að

$$a_1 < a_2 < \cdots < a_k.$$

Þá gefur meðalgildissetningin að φ' hefur að minnsta kosti eina núllstöð á sérhverju bilanna

$$]a_1, a_2[, \ ]a_2, a_3[, \ \dots]a_{k-1}, a_k[$$

Þau eru samanlagt $k-1$ talsins. Að auki vitum við að φ' hefur núllstöðvar af stigi að minnsta kosti $m_j - 1$ í punktinum a_j. Ef við leggjum þetta saman, þá fáum við að φ' hefur núllstöðvar af margfeldni að minnsta kosti

$$k - 1 + (m_1 - 1) + \cdots + (m_k - 1) = n - 1$$

í minnsta lokaða bilinu sem inniheldur alla punktana $a_1, \dots, a_k$.

3.7.6 Setning: Skekkjumat

Nú ætlum við að sýna fram á að fyrir föll f sem eru $(m+1)$ sinnum samfellt deildanleg að til sé ξ á minnsta bili sem inniheldur $a_1, \ldots, a_k$ og x þannig að

$$f[x_0, \ldots, x_m, x] = \frac{f^{(m+1)}(\xi)}{(m+1)!}$$

Við skilgreinum fallið

$$g(t) = f(t) - p(t) - \lambda w(t),$$

þar sem

$$w(t) = (t-a_1)^{m_1} \cdots (t-a_k)^{m_k}$$

og talan λ er valin þannig að $g(x) = 0$.

Nú er $p^{(j)}(a_i) = f^{(j)}(a_i)$ fyrir $j = 0, \ldots, m_i - 1$, þá gefur setning Taylors okkur að g hefur núllstöð af stigi m_i í sérhverjum punktanna a_i. Auk þess hefur g núllstöð í x. Samanlagt eru þetta að minnsta kosti $m+2$ núllstöðvar taldar með margfeldni.

Höfum:

- g hefur að minnsta kosti $m+2$ núllstöðvar taldar með margfeldni,
- g' hefur að minnsta kosti $m+1$ núllstöð talda með margfeldni,
- g'' hefur að minnsta kosti m núllstöðvar taldar með margfeldni
- og þannig áfram, þar til við ályktum að
- $g^{(m+1)}$ hefur að minnsta kosti eina núllstöð.

Tökum eina slíka og köllum hana ξ.

Munum að

$$g(t) = f(t) - p(t) - \lambda w(t),$$

þar sem

$$w(t) = (t-a_1)^{m_1} \cdots (t-a_k)^{m_k} = t^{m+1} + b_m t^m + \cdots + b_1 t + b_0$$

Margliðan p hefur stig $\leq m$ svo $p^{(m+1)}(x) = 0$ fyrir öll x

og margliðan w er af stigi $m+1$ með stuðul 1 við hæsta veldið, svo $w^{(m+1)}(t) = (m+1)!$. Við höfum því

$$0 = g^{(m+1)}(\xi) = f^{(m+1)}(\xi) - \lambda \cdot (m+1)!$$

sem jafngildir því að

$$\lambda = \frac{f^{(m+1)}(\xi)}{(m+1)!}$$

Við setjum nú inn $t = x$ sem gefur

$$0 = g(x) = f(x) - p(x) - \lambda w(x),$$

og við fáum þar með formúlu fyrir skekkjunni á nálgun á $f(x)$ með alhæfðu brúunarmargliðunni $p(x)$,

$$f(x) - p(x) = \lambda w(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!}(x-a_1)^{m_1} \cdots (x-a_k)^{m_k}$$

3.7.7 Samantekt

Ef gefið er fall $f : I \to \mathbb{R}$ á bili I, $a_1, \dots, a_k$ í I, með $a_j \neq a_k$ ef $j \neq k$, jákvæðar heiltölur $m_1, \dots, m_k$, talan m er skilgreind með $m = m_1 + \cdots + m_k - 1$, og gert er ráð fyrir að $f \in C^{m+1}(I)$, þá er til nákvæmlega ein margliða p af stigi $\leq m$ þannig að

$$p^{(j)}(a_i) = f^{(j)}(a_i), \qquad j = 0, \dots, m_i - 1, \quad i = 1, \dots, k.$$

Newton-form margliðunnar p er gefið með

$$p(x) = f[x_0] + f[x_0, x_1](x - x_0) + \cdots + f[x_0, \dots, x_m](x - x_0) \cdots (x - x_{m-1})$$

þar sem mismunakvótarnir $f[x_i, \dots, x_{i+j}]$ eru skilgreindir sem $y[x_i, \dots, x_{i+j}]$ út frá gögnunum $y_i^{(j)}$. Fyrir sérhvert x í I er skekkjan $f(x) - p(x)$ í nálgun á $f(x)$ með $p(x)$ gefin með

$$f(x) - p(x) = f[x_0, \dots, x_m, x](x - a_1)^{m_1} \cdots (x - a_k)^{m_k}.$$

Fyrir sérhvert $i = 1, \dots, k$ og $j = 0, \dots, m - i$ þá gildir að til er tala ξ á minnsta bilinu sem inniheldur $x_i \dots, x_{i+j}$ þannig að

$$f[x_i, \dots, x_{i+j}] = \frac{f^{(j)}(\xi)}{j!},$$

því gildir sérstaklega að til er tala ξ á minnsta bilinu sem inniheldur $a_1, \dots, a_k$ og x þannig að

$$f(x) - p(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!}(x - a_1)^{m_1} \cdots (x - a_k)^{m_k}.$$

3.7.8 Sýnidæmi

Látum $f(x) = x^2 \ln x$.

1. Setjið upp mismunakvótatöflu til þess að reikna út brúunarmargliðu p af stigi ≤ 3 fyrir fallið f, sem hefur tvo tvöfalda brúunarpunkta $a_1 = 1$ og $a_2 = 2$. Skrifið upp Newton-form margliðunnar p.
2. Reiknið út $p(1.3)$. Notið aðferðarskekkju fyrir margliðubrúun til þess að meta skekkjuna $f(1.3) - p(1.3)$ að ofan og neðan og fáið þannig bil þar sem rétta gildið liggur. Veljið miðpunkt bilsins sem nálgunargildi fyrir $f(1.3)$ og afrúnið gildið miðað við mörk bilsins.
3. Látum nú q vera brúnunarmargliðuna af stigi ≤ 4 sem uppfyllir sömu skilyrði og gefin eru í fyrsta lið að viðbættu því að $a_2 = 2$ á að vera þrefaldur brúunarpunktur. Sýnið hvernig hægt er að ákvarða mismunakvótatöfluna fyrir q með því að stækka töfluna í **a)**. Ákvarðið síðan q og reiknið út $q(1.3)$.

1. og 2. Til þess að spara pláss skulum leysa fyrsta og þriðja lið báða í einu með því að reikna strax út mismunakvótatöfluna fyrir fjórða stigs margliðuna í 3. lið. Punktarnir $x_0, \dots, x_4$ eru þá $1, 1, 2, 2, 2$ og við höfum gefin fallgildin

$$f(1) = f[x_0] = f[x_1] = 0 \qquad \text{og} \qquad f(2) = f[x_2] = f[x_3] = f[x_4].$$

Í 1. lið eru punktarnir tvöfaldir svo við höfum gefin gildi afleiðunnar $f'(x) = 2x \ln x + x$ í punktunum 1 og 2.

$$f'(1) = f[1, 1] = f[x_0, x_1] = 1 \ \text{ og } \ f'(2) = f[2, 2] = f[x_2, x_3] = 4 \ln 2 + 2.$$

Í 3. lið er gildið á 2. afleiðu $f''(x) = 2 \ln x + 3$ gefið í punktinum 2. Það gefur okkur

$$f''(2)/2! = f[2, 2, 2] = f[x_2, x_3, x_4] = \ln 2 + \tfrac{3}{2}.$$

Við setjum þessi gildi inn í mismunakvótatöfluna og fyllum hana út með því að taka mismunakvóta milli allra gilda

i	x_i	$f[x_i]$	$f[x_i,x_{i+1}]$	$f[x_i,x_{i+1},x_{i+2}]$	$f[x_i,\dots,x_{i+3}]$	$f[x_i,\dots,x_{i+4}]$
0	1	0	1	$4\ln 2-1$	$-4\ln 2+3$	$5\ln 2-\frac{7}{2}$
1	1	0	$4\ln 2$	2	$\ln 2-\frac{1}{2}$	
2	2	$4\ln 2$	$4\ln 2+2$	$\ln 2+\frac{3}{2}$		
3	2	$4\ln 2$	$4\ln 2+2$			
4	2	$4\ln 2$				

Margliðan í 1. lið er því

$$p(x)=(x-1)+(4\ln 2-1)(x-1)^2+(-4\ln 2+3)(x-1)^2(x-2).$$

en í 3. lið er hún

$$q(x)=p(x)+(5\ln 2-\tfrac{7}{2})(x-1)^2(x-2)^2$$

2. Við stingum gildinu $x=1.3$ inn í margliðuna og fáum $p(1.3)=0.445206074$. Skekkjan er

$$f(x)-p(x)=\frac{f^{(4)}(\xi)}{4!}(x-1)^2(x-2)^2$$

þar sem ξ er einhver punktur á bilinu $[1,2]$.

Við þurfum því að meta fjórðu afleiðuna,

$$f(x)=x^2\ln x,\quad f'(x)=2x\ln x+x,\quad f''(x)=2\ln x+3,$$
$$f'''(x)=2/x,\quad f^{(4)}(x)=-2/x^2.$$

Ef $x\in[1,2]$, þá höfum við matið $-2\le f^{(4)}(x)\le -\frac{1}{2}$.

Af ójöfnunum $-2\le f^{(4)}(x)\le -\frac{1}{2}$ leiðir síðan að

$$\alpha=\frac{-2\cdot(0.3)^2\cdot(-0.7)^2}{24}\le f(1.3)-p(1.3)\le\frac{-0.5\cdot(0.3)^2\cdot(-0.7)^2}{24}=\beta.$$

Við reiknum út úr báðum brotunum

$$\alpha=-0.003675\qquad\text{og}\qquad\beta=-0.00091875.$$

Þar með er $f(1.3)$ á bilinu milli $p(1.3)+\alpha=0.441531$ og $p(1.3)+\beta=0.444287$.

Nálgunargildi okkar á að vera miðpunktur þessa bils og algildi skekkjunnar verður þá hálf billengdin. Það færir okkur nálgunina $f(1.3)\approx 0.442909$ og skekkjuna ± 0.0014. Réttur afrúningur er $f(1.3)=0.44$.

Við eigum aðeins eftir að reikna út gildi margliðunnar q í punktinum 1.3. Út úr mismunakvótatöflunni fáum við

$$q(x)=p(x)+(5\ln 2-\tfrac{7}{2})(x-1)^2(x-2)^2$$

sem gefur okkur gildið

$$q(1.3)=0.445206074-0.001511046=0.4436950278$$

Til samanburðar höfum við rétt gildi

$$f(1.3)=0.443395606950060\dots.$$

3.8 Splæsibrúun

Látum $(t_0, y_0), \dots, (t_n, y_n)$ vera punkta í plani og gerum ráð fyrir að $a = t_0 < t_1 < \cdots < t_n = b$.

Við höfum nú lært að ákvarða margliðu p af stigi $\leq n$ sem tekur gildin y_i í punktunum t_i.

Ef punktarnir liggja á grafi fallsins f og nota á margliðuna til þess að nálga fallgildi f, þá getur það verið ýmsum erfiðleikum bundið þegar stig hennar stækkar eins og við sáum í byrjun kaflans Skynsamlegir skiptipunktar og Chebyshev margliður. Lausnin þar var að reyna að velja brúunarpunktana skynsamlega. Ef við hins vegar getum ekki valið brúunarpunktana eftir eigin höfði þá erum við í vandræðum og þurfum við að leita annarra leiða.

3.8.1 Almennt um splæsibrúun

Splæsibrúun er leið út úr þessum vandræðum.

Með henni er fundið samfellt fall S sem brúar gefnu punktana, $S(t_i) = y_i$, og er þannig að einskorðun þess við hlutbilin $[t_i, t_{i+1}]$ er gefið með margliðu af stigi $\leq m$, þar sem m er fyrirfram gefin tala.

Algengast er að nota $m = 3$.

3.8.2 Fyrsta stigs splæsibrúun:

Ef stigið m er 1, þá erum við einfaldlega að draga línustrik milli punktanna og sjáum í hendi okkar að lausnin er

$$S(x) = \begin{cases} S_0(x) = \dfrac{y_1 - y_0}{t_1 - t_0}(x - t_0) + y_0, & x \in [t_0, t_1], \\ S_1(x) = \dfrac{y_2 - y_1}{t_2 - t_1}(x - t_1) + y_1, & x \in [t_1, t_2], \\ \vdots & \\ S_{n-1}(x) = \dfrac{y_n - y_{n-1}}{t_n - t_{n-1}}(x - t_{n-1}) + y_{n-1}, & x \in [t_{n-1}, t_n]. \end{cases}$$

Þessi aðferð er ekki mikið notuð því hún er ósannfærandi fyrir deildanleg föll.

3.8.3 Þriðja stigs splæsibrúun

Algengast er að framkvæma splæsibrúun með þriðja stigs margliðum.

Við skulum tákna einskorðun S við hlutbilið $[t_i, t_{i+1}]$ með S_i og skrifa

$$S_i(x) = a_i + b_i(x - t_i) + c_i(x - t_i)^2 + d_i(x - t_i)^3, \qquad x \in [t_i, t_{i+1}).$$

Við ætlum að leiða út jöfnur fyrir stuðlunum a_i, b_i, c_i og d_i. Kröfurnar sem við setjum eru:

1. S er tvisvar sinnum samfellt deildanlegt á öllu bilinu $[a, b]$
2. S taki gildin y_i í punktunum t_i

Setjum til einföldunar $h_i = t_{i+1} - t_i$ fyrir $i = 0, \dots, n - 1$.

Skilyrðin tvö má því skrifa sem eftirfarandi jöfnuhneppi:

Á hverju hlutbili $[t_i, t_{i+1}]$ höfum við:

$$S_i(x) = a_i + b_i(x - t_i) + c_i(x - t_i)^2 + d_i(x - t_i)^3, \qquad x \in [t_i, t_{i+1}),$$

sem þýðir að skilyrðin tvö má skrifa sem

$$
\begin{aligned}
a_i = &\quad S_i(t_i) &= \quad y_i, \quad (1)\\
a_i + b_ih_i + c_ih_i^2 + d_ih_i^3 = &\quad S_i(t_{i+1}) = S_{i+1}(t_{i+1}) &= \quad a_{i+1}, \quad (2)\\
b_i + 2c_ih_i + 3d_ih_i^2 = &\quad S_i'(t_{i+1}) = S_{i+1}'(t_{i+1}) &= \quad b_{i+1}, \quad (3)\\
2c_i + 6d_ih_i = &\quad S_i''(t_{i+1}) = S_{i+1}''(t_{i+1}) &= \quad 2c_{i+1}, \quad (4)
\end{aligned}
$$

Í (1) höfum við $i = 0, \ldots, n$ og í (2)-(4) höfum við $i = 0, \ldots, n-2$.

Samtals: $(n+1) + 3(n-1) = 4n-2$ línulegar jöfnur til þess að ákvarða $4n$ óþekktar stærðir.

Það er því ljóst að okkur vantar tvö skilyrði til þess að geta fengið ótvírætt ákvarðaða lausn.

Fyrstu jöfnurnar gefa strax gildi a_i og (4) gefur að

$$
d_i = \frac{c_{i+1} - c_i}{3h_i}, \quad i = 0, \ldots, n-2
$$

Ef við setjum þetta inn í (2) og (3) fæst

$$
\begin{aligned}
a_{i+1} &= a_i + b_ih_i + \frac{c_{i+1} + c_i}{3}h_i^2, \quad i = 0, \ldots, n-2\\
b_{i+1} &= b_i + (c_{i+1} + c_i)h_i, \quad i = 0, \ldots, n-2
\end{aligned}
$$

Þegar við leysum fyrri jöfnuna fyrir b_i fæst

$$
b_i = \frac{a_{i+1} - a_i}{h_i} - \frac{c_i + c_{i+1}}{3}h_i, \quad i = 0, \ldots, n-2
$$

og ef við setjum þetta inn í seinni jöfnuna fæst á endanum að

$$
h_{i-1}c_{i-1} + 2(h_{i-1} + h_i)c_i + h_ic_{i+1} = \frac{3}{h_i}(a_{i+1} - a_i) - \frac{3}{h_{i-1}}(a_i - a_{i-1}), \quad i = 1, \ldots, n-1
$$

3.8.4 Jöfnuhneppið

$$
\begin{bmatrix}
.?. & .?. & & & & \\
h_0 & 2(h_0+h_1) & h_1 & & & \\
 & h_1 & 2(h_1+h_2) & h_2 & & \\
 & & \ddots & \ddots & \ddots & \\
 & & & h_{n-2} & 2(h_{n-2}+h_{n-1}) & h_{n-1}\\
 & & & & .?. & .?.
\end{bmatrix}
\begin{bmatrix}
c_0\\ c_1\\ c_2\\ \vdots\\ c_{n-1}\\ c_n
\end{bmatrix}
= 3
\begin{bmatrix}
.?.\\
\frac{a_2-a_1}{h_1} - \frac{a_1-a_0}{h_0}\\
\frac{a_3-a_2}{h_2} - \frac{a_2-a_1}{h_1}\\
\vdots\\
\frac{a_n-a_{n-1}}{h_{n-1}} - \frac{a_{n-1}-a_{n-2}}{h_{n-2}}
\end{bmatrix}
$$

En það vantar í þetta einhver skilyrði á c_0 og c_n.

Þegar þau hafa verið sett inn, þá getum við leyst þetta hneppi, reiknað svo gildi b_i og d_i og þá höfum við fundið splæsifallið okkar.

Það eru til margar leiðir til að ákvarða c_0 og c_n, en fjórar eru algengastar.

3.8.5 Tilfelli 1: Ekki-hnúts endaskilyrði

Ef við höfum engar upplýsingar um fallið f í t_1 og t_{n-1} liggur beint við að krefjast þess að S''' sé samfellt þar, sem þýðir að $d_0 = d_1$ og $d_{n-2} = d_{n-1}$. Með að nota jöfnurnar fyrir d_i má skrifa þetta sem

$$
\begin{aligned}
h_1c_0 - (h_0 + h_1)c_1 + h_0c_2 &= 0\\
h_{n-1}c_{n-2} - (h_{n-2} + h_{n-1})c_{n-1} + h_{n-2}c_n &= 0
\end{aligned}
$$

og þessar jöfnur, ásamt hinum, má leysa til að ákvarða c_i-in.

3.8.6 Tilfelli 2: Þvinguð endaskilyrði

Ef hallatala fallsins f er þekkt í endapunktum bilsins er eðlilegt að nota þær upplýsingar við ákvörðun splæsifallsins. Gerum því ráð fyrir að $f'(t_0) = A$ og $f'(t_n) = B$. Skilyrðið $S'(t_0) = A$ gefur þá að

$$A = \frac{a_1 - a_0}{h_0} - \frac{2c_0 + c_1}{3} h_0,$$

eða

$$2h_0 c_0 + h_0 c_1 = 3\left(\frac{a_1 - a_0}{h_0} - A\right)$$

og $S'(t_n) = B$ gefur

$$B = b_{n-1} + 2c_{n-1}h_{n-1} + 3d_{n-1}h_{n-1}^2$$

og með að nota formúlurnar fyrir b_{n-1} og d_{n-1} fæst

$$c_{n-1}h_{n-1} + 2c_n h_{n-1} = 3\left(B - \frac{a_n - a_{n-1}}{h_{n-1}}\right).$$

3.8.7 Tilfelli 3: Náttúrleg endaskilyrði

Einfaldasta lausnin er að setja $c_0 = c_n = 0$, en það jafngildir því að $S''(t_0) = S''(t_n) = 0$.

3.8.8 Tilfelli 4: Lotubundið endaskilyrði

Hugsum okkur að við viljum framlengja S í tvisvar samfellt deildanlegt $(b-a)$-lotubundið fall á $\mathbb{R}$. Það setur skilyrðin

$$y_0 = S(t_0) = S_(t_n) = y_n, \quad S'(t_0) = S'(t_n), \quad \text{og} \quad S''(t_0) = S''(t_n)$$

Fljótséð er að $S''(t_0) = S''(t_n)$ þýðir að $c_0 = c_n$, eða

$$c_0 - c_n = 0.$$

Þetta er fyrri jafnan sem við þurfum.

Nú gefur $S'(t_0) = S'(t_n)$ að

$$b_0 = b_{n-1} + 2c_{n-1}h_{n-1} + 3d_{n-1}h_{n-1}^2$$

og með að setja inn formúlurnar fyrir b_0, b_{n-1}, d_{n-1} og nota að $c_0 = c_n$ fæst jafnan

$$h_0 c_1 + 2h_{n-1}c_{n-1} + (2h_0 + 2h_{n-1})c_n = 3\left(\frac{a_1 - a_0}{h_0} - \frac{a_n - a_{n-1}}{h_{n-1}}\right).$$

3.8.9 Teikning á ferlum í planinu

Hægt er að nota brúun til þess að nálga ferla í $\mathbb{R}^n$. Skoðum tilvikið $n = 2$.

Gerum nú ráð fyrir að við höfum gefna punkta $(x_0, y_0), \ldots, (x_n, y_n) \in \mathbb{R}^2$ og að við viljum finna samfelldan splæsiferil í gegnum þá. Þetta er gert í nokkrum skrefum:

1. Ákveðið er stikabil $[a, b]$ og skiptingu á því

$$a = t_0 < t_1 < \cdots < t_n = b$$

   til dæmis $[0, n]$ og skiptinguna

$$0 = t_0 < t_1 = 1 < \cdots < t_n = n.$$

2. Ákveðið er hvaða endaskilyrði eiga við.
3. Búin eru til tvö splæsiföll $R(t)$ fyrir punktasafnið $x_0, \dots, x_n$ og $S(t)$ fyrir punktasafnið $y_0, \dots, y_n$.
4. Stikaferillinn $[a, b] \ni t \mapsto (R(t), S(t))$ er síðan teiknaður, en hann uppfyllir $(R(t_j), S(t_j)) = (x_j, y_j)$, $j = 0, \dots, n$.

Aðvörun: Athugið að hér er t breytan okkar en x og y eru gildin sem við viljum að ferillinn taki.
Þetta er frábrugðið því þegar við skoðum graf af einni breytu en þá er x venjulega breytan og y gildin sem viljum taka.

3.9 Aðferð minnstu fervika

Látum $(x_1, y_1), \dots, (x_m, y_m)$ vera safn punkta í plani með $x_j \in [a, b]$ fyrir öll j og látum $f_1, \dots, f_n$ vera raungild föll á $[a, b]$.

Við viljum finna það fall f af gerðinni

$$f(x) = c_1 f_1(x) + \cdots + c_n f_n(x)$$

með stuðla $c_1, \dots, c_n$ þannig að punktarnir $(x_j, f(x_j))$ nálgi gefna punktasafnið sem best og þá er átt við að fernings-summuna

$$\sum_{i=1}^{m} \big(y_i - f(x_i)\big)^2$$

verði eins lítil og mögulegt er.

3.9.1 Jafna bestu línu

Flestir hafa heyrt talað um bestu línu gegnum punktasafn, hún fæst með að taka hér $f_1(x) = 1$ og $f_2(x) = x$, en lítið mál er að finna einnig besta fleygboga, bestu margliðu af fyrirfram ákveðnu stigi eða einhverja aðra samantekt falla gegnum punktasafnið.

3.9.2 Smávegis línuleg algebra

Til þess að finna þessi gildi á stuðlunum c_i er heppilegt að notfæra sér nokkrar niðurstöður úr línulegri algebru. Fyrir gefin gildi á $c_1, \dots, c_n$ setjum við

$$b_i = f(x_i) = c_1 f(x_i) + \cdots + c_n f_n(x_i), \qquad i = 1, \dots, m,$$

og skilgreinum síðan dálkvigrana

$$b = [b_1, \dots, b_m]^T, \qquad y = [y_1, \dots, y_m]^T, \quad \text{og} \quad c = [c_1, \dots, c_n]^T,$$

Þá er $Ac = b$, þar sem A er $m \times n$ fylkið

$$A = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \dots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \dots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_m) & f_2(x_m) & \dots & f_n(x_m) \end{bmatrix}.$$

Verkefnið snýst nú um að finna þann vigur $c \in \mathbb{R}^n$ sem lágmarkar

$$\sum_{i=1}^{m} \big(y_i - b_i\big)^2 = \|y - b\|^2 = \|y - Ac\|^2$$

þar sem $\|\cdot\|$ táknar evklíðska fjarlægðina (staðalinn) á $\mathbb{R}^m$.

Vigrar af gerðinni $b = Ac$ spanna dálkrúm fylkisins A og þá má skrifa sem línulegar samantektir af gerðinni

$$b = c_1 A_1 + \cdots + c_n A_n$$

þar sem A_j er dálkur númer j.

Verkefnið snýst um að finna þann vigur í dálkrúminu sem næstur er y. Vigurinn b er næstur y ef og aðeins ef $y - b$ er hornréttur á alla vigra dálkrúmsins.

Þessi skilyrði má fá með innfeldi

$$A_j \cdot (y - b) = 0, \qquad j = 1, \dots, n$$

Með fylkjarithætti fæst ein jafna

$$A^T(y - b) = 0.$$

Setjum nú inn $b = Ac$. Þá ákvarðast c af hneppinu

$$A^T(y - Ac) = 0$$

sem jafngildir

$$(A^T A)c = A^T y$$

Við þurfum því aðeins að leysa þetta jöfnuhneppi

$$(A^T A)c = A^T y$$

fyrir c til að finna stuðlana okkar. Ef fylkið $A^T A$ hefur andhverfu, þá fæst alltaf ótvírætt ákvörðuð lausn c.

Ef fylkið $A^T A$ hefur ekki andhverfu eða að það hefur ákveðu sem er mjög nálægt 0, þá þurfum við að beita flóknari brögðum. Við komum að því síðar.

3.9.3 Jafna bestu línu

Algengt er að menn vilji finna beina línu sem best fellur að punktasafninu $(x_1, y_1,) \dots, (x_m, y_m)$. Þá er $n = 2$ og við tökum lausnagrunninn $f_1(x) = 1$ og $f_2(x) = x$.

Fylkið er þá

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}.$$

og þar með

$$A^T A = \begin{bmatrix} m & \sum_{j=1}^m x_j \\ \sum_{j=1}^m x_j & \sum_{j=1}^m x_j^2 \end{bmatrix}. \qquad \text{og} \qquad A^T y = \begin{bmatrix} \sum_{j=1}^m y_j \\ \sum_{j=1}^m x_j y_j \end{bmatrix}.$$

Athugasemd: Það er auðvelt að leysa þetta tilvik því við höfum einfalda formúlu fyrir andhverfum 2×2 fylkja (svo lengi sem ákveðan er ekki 0).

Sjá Wikipedia.

3.9.4 Jafna bestu annars stigs margliðu

Ef við viljum finna bestu annars stigs margliðu gegnum punktasafnið, þá er $n = 3$ og við tökum lausnagrunninn $f_1(x) = 1$, $f_2(x) = x$ og $f_3(x) = x^2$.

Þetta val gefur fylkið

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 \end{bmatrix}.$$

Fylkið $A^T A$ er þá 3×3 og vigurinn $A^T y$ er dálkvigur með 3 hnit.

3.9.5 Sýnidæmi: besta annars stigs margliða

Gefin eru mæligildin

x	0	1	2	3	4	5	6
y	2.7	-0.5	-1.7	-1.9	-1.5	0.2	2.3

Beitið aðferð minnstu fervika til þess að finna þá annars stigs margliðu sem best fellur að þessum gögnum Teiknið upp gögnin og graf marliðunnar.

Við leitum hér að þremur tölum c_1, c_2 og c_3 þannig að annars stigs margliðan $f(x) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x)$ falli sem best að gögnunum. Grunnföllin þrjú eru $f_1(x) = 1$, $f_2(x) = x$ og $f_3(x) = x^2$.

Í þessu dæmi er fylkið A gefið með

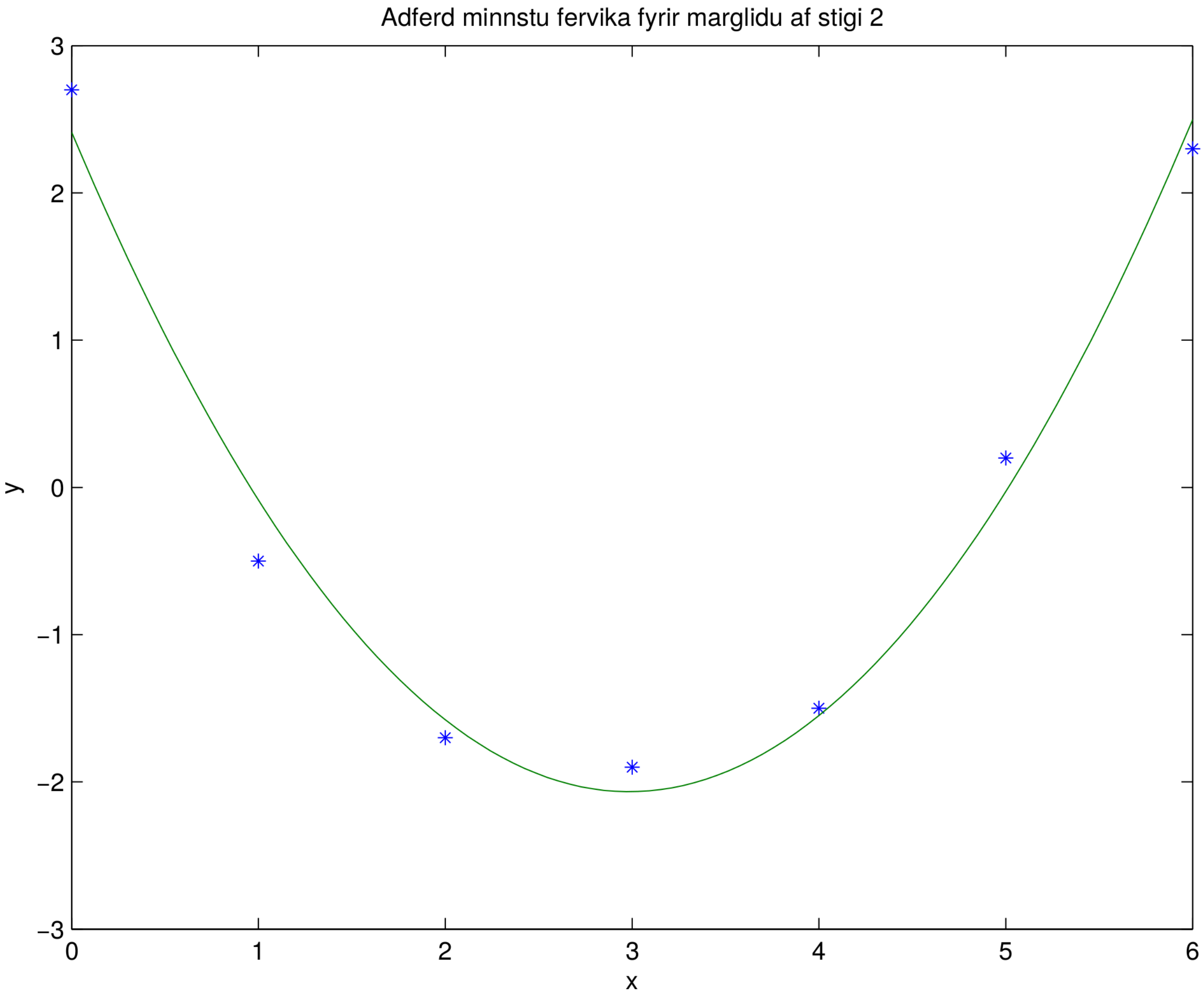
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 6 & 36 \end{bmatrix},$$

því stak númer (i, j) í A er gefið með $A_{ij} = f_j(x_i)$.

Nú látum við matlab um afganginn

```
%  Matlab forrit sem teiknar upp bestu margliðunálgun á gefnum gögnum
x=[0; 1; 2; 3; 4; 5; 6]
y=[2.7; -0.5; -1.7; -1.9; -1.5; 0.2; 2.3 ]
m=length(x);

% Við leitum að bestu margliðu af stigi 2 eða lægri
% og því eru  grunnföllin eru 3 talsins.
n=3;

% Stuðlafylkið er A=(a_{ij}), a_{ij}=x_i^{j-1}
A(1:m,1)=ones(m,1);
A(1:m,2)=x;
for j=3:n
    A(1:m,j)=A(1:m,j-1).*x;
end
% Reiknum úr úr normaljöfnuhneppinu A^TAc=A^Ty:
c=(A'*A)\(A'*y);

% Teikning undirbúin
N=100;
X=linspace(min(x),max(x),N);

% Hliðrun í reikniriti horners er 0
%
hlidrun=zeros(n,1);
for j=1:N
    Y(j)=horner(c, hlidrun, X(j));
end
figure
plot(x,y,'*',X,Y)
xlabel('x'), ylabel('y')
title('Adferd minnstu fervika fyrir marglidu af stigi 2')
print
```

Hér kemur myndin sem beðið var um:


Adferd minnstu fervika fyrir marglidu af stigi 2
y
3
2
1
0
−1
−2
−3
0
1
2
3
4
5
6
x

KAFLI 4

Töluleg diffrun

Nanny's philosophy of life was to do what seemed like a good idea at the time, and do it as hard as possible. It had never let her down. – Terry Pratchett, Maskerade

4.1 Inngangur

4.1.1 Töluleg diffrun og heildun

Deildun og heildun eru meginaðgerðir stærðfræðigreiningarinnar.

Þess vegna er nauðsynlegt að geta nálgað

$$f'(a), f''(a), f'''(a), \ldots \quad \text{og} \quad \int_a^b f(x)\,dx,$$

þar sem f er fall sem skilgreint er á bili I sem inniheldur a og b.

4.1.2 Meginhugmynd í öllum nálgunaraðferðunum

Látum p vera margliðu sem nálgar f, og látum $r(x) = f(x) - p(x)$ tákna skekkjuna í nálgun á $f(x)$ með $p(x)$. Þá er

$$f'(x) = p'(x) + r'(x), \quad f''(x) = p''(x) + r''(x), \ldots$$

og

$$\int_a^b f(x)\,dx = \int_a^b p(x)\,dx + \int_a^b r(x)\,dx.$$

Nú þurfum við að gera tvennt:

1. Finna heppilegar nálgunarmargliður og reikna út

$$p'(a),\ p''(a), \ldots, \qquad \int_a^b p(x)\,dx$$

2. Meta skekkjurnar

$$r'(a),\ r''(a), \ldots \int_a^b r(x)\,dx$$

Byrjum á að leiða út nokkrar nálgunarformúlur með skekkjumati.

4.2 Aðferðirnar

Látum $f : I \to \mathbb{R}$ vera fall á bili $I \subset \mathbb{R}$ og a vera punkt í I. Afleiða f í punktinum a er skilgreind með

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

ef markgildið er til. Við skrifum því oft

$$f'(a) \approx \frac{f(a+h) - f(a)}{h}$$

Þessi nálgun er kölluð *frammismunur* því oftast hugsar maður sér að $h > 0$ og þá er $a + h$ lítið skref áfram frá a.

Við þurfum skekkjumat fyrir þessa formúlu ef við eigum að geta notað hana.

4.2.1 Frammismunur

Við fáum mat á skekkjuna í nálguninni með að skoða Taylor-margliðu f í a. Samkvæmt setningu Taylors er til ξ á milli a og $a + h$ þannig að

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}f''(\xi)h^2.$$

Þá fæst að skekkjan í nálgun á $f'(a)$ með

$$\frac{f(a+h) - f(a)}{h} = f[a, a+h]$$

er

$$e = f'(a) - \frac{f(a+h) - f(a)}{h} = -\frac{1}{2}f''(\xi)h$$

Með öðrum orðum

$$\min_{t\in[0,h]} -\frac{1}{2}f''(t)h \le e \le \max_{t\in[0,h]} -\frac{1}{2}f''(t)h.$$

Við sjáum því að $e = O(h)$ þegar $h \to 0$.

4.2.2 Bakmismunur

Við getum sett $a - h$ í stað $a + h$ í skilgreininguna á afleiðu. Þá fæst svokallaður *bakmismunur*

$$f'(a) \approx \frac{f(a) - f(a-h)}{h}$$

og ljóst er að sama skekkjumat gengur fyrir þessa nálgun og fyrir nálgun með frammismun.

4.2.3 Miðsettur mismunakvóti

Lítum nú á þriðja stigs Taylor nálgun

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \tfrac{1}{2}f''(a)h^2 + \tfrac{1}{6}f'''(\alpha)h^3, \\ f(a-h) &= f(a) - f'(a)h + \tfrac{1}{2}f''(a)h^2 - \tfrac{1}{6}f'''(\beta)h^3, \end{aligned}$$

þar sem α er á milli a og $a+h$ og β er á milli a og $a-h$.

Tökum nú mismuninn og fáum

$$f(a+h) - f(a-h) = f'(a)\cdot 2h + \tfrac{1}{6}\big(f'''(\alpha) + f'''(\beta)\big)h^3$$

Ef f''' er samfellt fall, þá gefur milligildissetningin okkur að til er ξ á milli α og β þannig að $f'''(\xi) = \frac{1}{2}(f'''(\alpha) + f'''(\beta))$

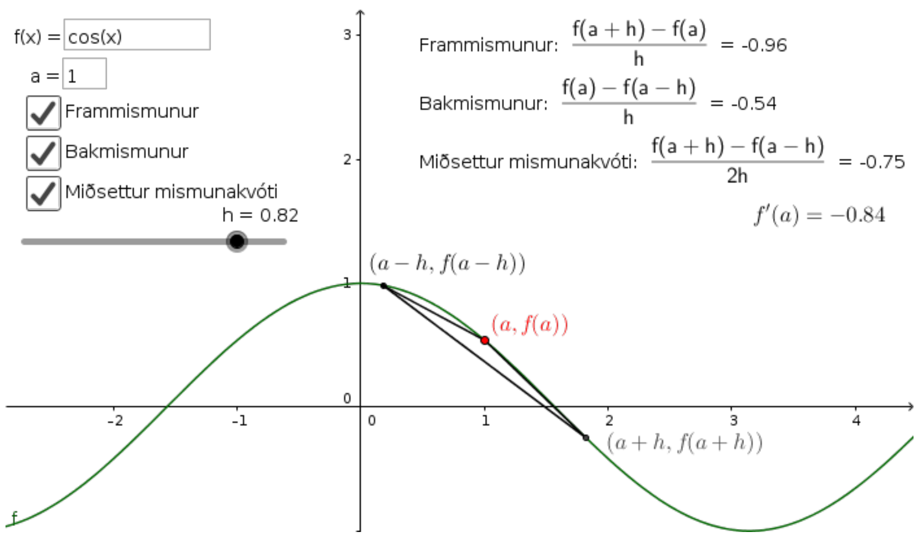
Niðurstaðan verður

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} - \tfrac{1}{6}f'''(\xi)h^2.$$

Þannig að skekkjan er

$$e = -\frac{1}{6}f'''(\xi)h^2,$$

og jafnframt er $e = O(h^2)$ þegar $h \to 0$.



4.2.4 Miðsettur mismunakvóti fyrir aðra afleiðu

Við getum útfært þessa sömu hugmynd til þess að reikna út aðra afleiðu, en þá byrjum við með fjórða stigs Taylor-nálgun

$$f(a+h) = f(a) + f'(a)h + \tfrac{1}{2}f''(a)h^2 + \tfrac{1}{6}f'''(a)h^3 + \tfrac{1}{24}f^{(4)}(\alpha)h^4,$$
$$f(a-h) = f(a) - f'(a)h + \tfrac{1}{2}f''(a)h^2 - \tfrac{1}{6}f'''(a)h^3 + \tfrac{1}{24}f^{(4)}(\beta)h^4,$$

þar sem α er á milli a og $a+h$ og β er á milli a og $a-h$.

Nú leggjum við saman og fáum

$$f(a+h) + f(a-h) = 2f(a) + f''(a)h^2 + \tfrac{1}{24}\big(f^{(4)}(\alpha) + f^{(4)}(\beta)\big)h^4.$$

Nú þurfum við að gefa okkur að $f^{(4)}$ sé samfellt fall, þá gefur milligildissetningin okkur að til er ξ á milli α og β þannig að $f^{(4)}(\xi) = \frac{1}{2}(f^{(4)}(\alpha) + f^{(4)}(\beta))$.

Niðurstaðan verður

$$f''(a) = \frac{f(a+h) + f(a-h) - 2f(a)}{h^2} - \tfrac{1}{12}f^{(4)}(\xi)h^2$$

Með Taylor-margliðum má leiða út fleiri nálgunarformúlur fyrir afleiður.

Við ætlum ekki að halda lengra í þessa átt heldur snúa okkur að almennu aðferðinni.

4.3 Skekkjumat

4.3.1 Almennt um nálganir á afleiðum

Ef $x_0, \dots, x_n$ eru punktar í I (hugsanlega með endurtekningum) og p er margliðan sem brúar f í þeim, þá er

$$f(x) = p(x) + r(x),$$

þar sem skekkjuliðurinn $r(x)$ er gefinn með formúlunni

$$r(x) = f[x_0, \dots, x_n, x](x - x_0) \cdots (x - x_n)$$

Ef við tökum $p'(a)$ sem nálgun á $f'(a)$ er skekkjan

$$r'(a) = f'(a) - p'(a).$$

4.3.2 Skekkjumat

Munið að formúlan fyrir afleiðu af margfeldi margra þátta er

$$\begin{gathered}(\varphi_1\varphi_2\varphi_3\cdots\varphi_m)'(a)\\ = \varphi_1'(a)\varphi_2(a)\varphi_3(a)\cdots\varphi_m(a) + \varphi_1(a)\varphi_2'(a)\varphi_3(a)\cdots\varphi_m(a) + \cdots\\ \cdots + \varphi_1(a)\varphi_2(a)\cdots\varphi_{m-1}(a)\varphi_m'(a)\end{gathered}$$

Horfum nú á skekkjuliðinn $r(x)$. Hann er svona margfeldi með $\varphi_1(x) = f[x_0, \dots, x_n, x]$, $\varphi_2(x) = x - x_0$, $\varphi_3(x) = x - x_1$ o.s.frv.

Athugum nú að ef a er einn af gefnu punktunum x_k, þá er $\varphi_{k+2}(x) = (x - x_k)$ sem gefur $\varphi_{k+2}(x_k) = 0$ og $\varphi_{k+2}'(x_k) = 1$.

Þetta segir okkur að ef við tökum $a = x_k$, þá eru allir liðirnir í summunni í hægri hliðinni 0 nema einn, þ.e. við sitjum eftir með þann sem inniheldur φ_{k+2}'.

Niðurstaðan verður því að skekkjan í nálgun á $f'(a)$ með $p'(a)$ er

$$\begin{aligned}f'(a) - p'(a) = r'(a) &= f[x_0, \dots, x_n, x_k] \prod_{\substack{j=0\\ j\neq k}} (x_k - x_j)\\ &= \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{\substack{j=0\\ j\neq k}} (a - x_j)\end{aligned}$$

þar sem $a = x_k$.

Hér notuðum við skekkjumatið fyrir Newton aðferðina sem segir að til er ξ á minnsta bilinu sem inniheldur $x_0, \dots, x_n, x_k$ sem uppfyllir

$$f[x_0, \dots, x_n, x_k] = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

4.3.3 Frammismunur

Nálgum f með fyrsta stigs brúunarmargliðunni gegnum punktana $(a, f(a))$ og $(a+h, f(a+h))$ (þ.e. $x_0 = a$ og $x_1 = a+h$),

$$f(x) = f[a] + f[a, a+h](x-a) + f[a, a+h, x](x-a)(x-a-h)$$

Af þessu leiðir formúlan sem við vorum áður komin með

$$f'(a) = f[a, a+h] + f[a, a+h, a](a-a-h) = \frac{f(a+h)-f(a)}{h} - \tfrac{1}{2}f''(\xi)h$$

Þar sem ξ er á milli a og $a+h$ og uppfyllir að $f[a, a+h, a] = f[a, a, a+h] = \frac{1}{2}f''(\xi)$. Hér erum við að notafæra okkur aftur skekkjumatið sem við sönnuðum í kaflanum um brúunarmargliður.

4.3.4 Miðsettur mismunakvóti

Tökum þriggja punkta brúunarformúlu með $a-h$, $a+h$ og a. Þá er

$$\begin{aligned} f(x) = {} & f[a-h] + f[a-h, a+h](x-a+h) \\ & + f[a-h, a+h, a](x-a+h)(x-a-h) \\ & + f[a-h, a+h, a, x](x-a+h)(x-a-h)(x-a) \end{aligned}$$

Athugum að afleiðan af annars stigs þættinum

$$x \mapsto (x-a+h)(x-a-h) = (x-a)^2 - h^2$$

er 0 í punktinum a og því er

$$\begin{aligned} f'(a) &= f[a-h, a+h] + f[a-h, a+h, a, a](-h^2) \\ &= \frac{f(a+h)-f(a-h)}{2h} - \tfrac{1}{6}f'''(\xi)h^2 \end{aligned}$$

Hér nýttum við okkur að til er ξ á milli $a-h$ og $a+h$ þannig að $f[a-h, a+h, a, a] = \frac{1}{6}f'''(\xi)$.

4.3.5 Miðsettur mismunakvóti fyrir aðra afleiðu

Áfram heldur leikurinn. Nú skulum við leiða aftur út formúluna fyrir nálgun á $f''(a)$ með miðsettum mismunakvóta

Þá tökum við þriggja punkta brúunarformúlu með $a-h$, $a+h$ og a með a tvöfaldan. Þá er

$$\begin{aligned} f(x) = {} & f[a-h] + f[a-h, a+h](x-a+h) \\ & + f[a-h, a+h, a](x-a+h)(x-a-h) \\ & + f[a-h, a+h, a, a](x-a+h)(x-a-h)(x-a) \\ & + f[a-h, a+h, a, a, x](x-a+h)(x-a-h)(x-a)^2 \end{aligned}$$

Gætum þess að halda liðnum $(x-a)$. Þá fáum við

$$\begin{aligned} f(x) = {} & f[a-h] + f[a-h, a+h](x-a+h) \\ & + f[a-h, a+h, a]\big((x-a)^2 - h^2)\big) \\ & + f[a-h, a+h, a, a]\big((x-a)^3 - h^2(x-a))\big) \\ & + f[a-h, a+h, a, a, x]\big((x-a)^4 - h^2(x-a)^2)\big) \end{aligned}$$

Nú þurfum við að reikna aðra afleiðu í punktinum a. Athugum að önnur afleiða af annars stigs þættinum

$$x \mapsto (x-a+h)(x-a-h) = (x-a)^2 - h^2$$

er fastafallið 2, önnur afleiða af þriðja stigs liðnum

$$x \mapsto (x-a)^3 - h^2(x-a)$$

er 0 í punktinum a og önnur afleiða af fjórða stigs liðnum

$$x \mapsto (x-a)^4 - h^2(x-a)^2$$

er fastafallið $-2h^2$.

Við höfum því

$$f''(a) = 2f[a-h, a+h, a] + f[a-h, a+h, a, a, a](-2h^2)$$

Nú er til punktur ξ á minnsta bili sem inniheldur $a-h$, $a+h$ og a þannig að $f[a-h, a+h, a, a, a] = \frac{1}{24} f^{(4)}(\xi)$.

Við þurfum að reikna út fyrri mismunakvótann

$$\begin{aligned} f[a-h, a+h, a] = f[a-h, a, a+h] &= \frac{f[a, a+h] - f[a-h, a]}{2h} \\ &= \frac{1}{2h}\left(\frac{f(a+h)-f(a)}{h} - \frac{f(a)-f(a-h)}{h}\right) \\ &= \frac{f(a+h)+f(a-h)-2f(a)}{2h^2} \end{aligned}$$

Við höfum því leitt aftur út formúluna

$$f''(a) = \frac{f(a+h)+f(a-h)-2f(a)}{h^2} - \tfrac{1}{12} f^{(4)}(\xi) h^2$$

4.4 Richardson útgiskun

Það ætti að vera ljóst að töluleg deildun er nokkuð óstöðug aðferð því ef skrefastærðin h er lítil eru tölurnar $f(a+h), f(a), f(a-h)$ nálægt hver annarri og við getum lent í styttingarskekkjum.

Því er ekki hægt að búast við að fá alltaf betri nálgun á $f'(a)$ við að minnka skrefalengdina h.

Leiðin er Richardson útgiskun (e. extrapolation), sem er aðferð til að bæta nálganir.

Til eru mjög almennar útgáfur þessarar aðferðar en við munum aðeins skoða þau sértilfelli sem nýtast okkur mest.

4.4.1 Útleiðsla á miðsettum mismunakvóta

Við skulum byrja á að að leiða aftur út formúluna fyrir miðsettann mismunakvóta til að fá betri upplýsingar um skekkjuliðinn. Fyrir fall f sem er nógu oft deildanlegt má beita Taylor til að skrifa

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \ldots + \frac{f^{(2n)}(a)}{(2n!)} h^{2n} + \frac{f^{(2n+1)}(a)}{(2n+1)!)} h^{2n+1} + O(h^{2n+2}) \\ f(a-h) &= f(a) - f'(a)h + \ldots + \frac{f^{(2n)}(a)}{(2n!)} h^{2n} - \frac{f^{(2n+1)}(a)}{(2n+1)!)} h^{2n+1} + O(h^{2n+2}) \end{aligned}$$

Ef við drögum seinni jöfnuna frá þeirri fyrri fæst

$$f(a+h) - f(a-h) = 2f'(a)h + 2\frac{f'''(a)}{3!}h^3 + \ldots + 2\frac{f^{(2n+1)}(a)}{(2n+1)!}h^{2n+1} + O(h^{2n+2})$$

svo ef við einangrum $f'(a)$ sjáum við að

$$f'(a) = R_1(h) + a_2h^2 + a_4h^4 + \ldots + a_{2n}h^{2n} + O(h^{2n+1})$$

þar sem

$$R_1(h) = \frac{f(a+h) - f(a-h)}{2h} \quad \text{og} \quad a_k = -\frac{f^{(k+1)}(a)}{(k+1)!}, \quad k = 2, 4, \ldots, 2n.$$

4.4.2 Helmingun á skrefinu

Hér er minnsta veldi í skekkjuliðnum h^2, svo nálgunin $f'(a) \approx R_1(h)$ er $O(h^2)$, eins og við höfum reyndar séð áður. Helmingum nú skrefalengdina h, þá fæst

$$f'(a) = R_1(h/2) + a_2\left(\frac{h}{2}\right)^2 + a_4\left(\frac{h}{2}\right)^4 + \ldots + a_{2n}\left(\frac{h}{2}\right)^{2n} + O(h^{2n+1}).$$

Nú berum við saman þessi tvö skref:

$$\begin{aligned} f'(a) &= R_1(h/2) + \tfrac{1}{4}a_2h^2 + a_4\left(\frac{h}{2}\right)^4 + \ldots + a_{2n}\left(\frac{h}{2}\right)^{2n} + O(h^{2n+1}), \\ f'(a) &= R_1(h) + a_2h^2 + a_4h^4 + \ldots + a_{2n}h^{2n} + O(h^{2n+1}) \end{aligned}$$

Margföldum efri jöfnuna með 4 og drögum þá síðari frá. Þá stendur eftir

$$\begin{aligned} 3f'(a) = 4R_1(h/2) - R_1(h) &+ a_4\left(\frac{4}{2^4} - 1\right)h^4 \\ &+ a_6\left(\frac{4}{2^6} - 1\right)h^6 + \ldots + a_{2n}\left(\frac{4}{2^{2n}} - 1\right)h^{2n} + O(h^{2n+1}) \end{aligned}$$

4.4.3 Fjórða stigs nálgun

Nú erum við komin með nýja formúlu:

$$f'(a) = R_2(h) + b_4h^4 + b_6h^6 + \ldots + b_{2n}h^{2n} + O(h^{2n+1})$$

þar sem

$$R_2(h) = \frac{4R_1(h/2) - R_1(h)}{3} \quad \text{og} \quad b_k = \frac{a_k}{3}\cdot\left(\frac{4}{2^k} - 1\right),\ k = 4, 6, \ldots, 2n.$$

Ef við berum þetta saman við jöfnuna sem við byrjuðum með

$$f'(a) = R_1(h) + a_2h^2 + a_4h^4 + \ldots + a_{2n}h^{2n} + O(h^{2n+1})$$

þá sjáum við að minnsta veldi í skekkjuliðnum er h^4, svo nálgunin $f'(a) \approx R_2(h)$ uppfyllir

$$f'(a) - R_2(h) = O(h^4)$$

og er því betri nálgun en áður.

Þetta ferli heitir *Richardson útgiskun*.

4.4.4 Hægt er að halda áfram útgiskun

Næsta takmark er að eyða liðnum $b_4 h^4$ úr þessari formúlu með því að líta á

$$f'(a) = R_2(h/2) + b_4\left(\frac{h}{2}\right)^4 + b_6\left(\frac{h}{2}\right)^6 + \ldots + b_{2n}\left(\frac{h}{2}\right)^{2n} + O(h^{2n+1})$$

Síðan stillum við þessari jöfnu upp með þeirri síðari

$$\begin{aligned} f'(a) &= R_2(h/2) + \tfrac{1}{16}b_4h^4 + \tfrac{1}{64}b_6h^6 + \ldots + \tfrac{1}{2^{2n}}b_{2n}h^{2n} + O(h^{2n+1}) \\ f'(a) &= R_2(h) + b_4h^4 + b_6h^6 + \ldots + b_{2n}h^{2n} + O(h^{2n+1}) \end{aligned}$$

Margföldum fyrri jöfnuna með 16 og drögum þá síðari frá

$$\begin{aligned} 15f'(a) = {} & 16R_2(h/2) - R_2(h) + b_6\left(\frac{16}{2^6} - 1\right)h^6 \\ & + b_8\left(\frac{16}{2^8} - 1\right)h^8 + \ldots + b_{2n}\left(\frac{16}{2^{2n}} - 1\right)h^{2n} + O(h^{2n+1}). \end{aligned}$$

4.4.5 Sjötta stigs skekkja

$$\begin{aligned} 15f'(a) = {} & 16R_2(h/2) - R_2(h) + b_6\left(\frac{16}{2^6} - 1\right)h^6 \\ & + b_8\left(\frac{16}{2^8} - 1\right)h^8 + \ldots + b_{2n}\left(\frac{16}{2^{2n}} - 1\right)h^{2n} + O(h^{2n+1}). \end{aligned}$$

Því er

$$f'(a) = R_3(h) + c_6h^6 + c_8h^8 \ldots + c_{2n}h^{2n} + O(h^{2n+1})$$

þar sem

$$R_3(h) = \frac{16R_2(h/2) - R_2(h)}{15}, \quad \text{og} \quad c_k = \frac{b_k}{15}\cdot\left(\frac{16}{2^k} - 1\right), \quad k = 6, 8, \ldots, 2n.$$

Nýja nálgunin uppfyllir

$$f'(a) - R_3(h) = O(h^6)$$

og er því enn betri en áður, en við þurfum líka að reikna út $R_1(h/4)$ til að reikna $R_2(h/2)$.

4.4.6 Almenn rakningarformúla

Richardson-útgiskunin heldur áfram og út kemur

$$R_{i+1}(h) = \frac{4^iR_i(h/2) - R_i(h)}{4^i - 1} = R_i(h/2) + \frac{R_i(h/2) - R_i(h)}{4^i - 1}$$

fyrir $(i+1)$-tu Richardson útgiskun og $R_{i+1}(h)$ uppfyllir að

$$f'(a) - R_{i+1}(h) = O(h^{2i+2}),$$

en á móti kemur að til að reikna út $R_{i+1}(h)$ þurfum við að hafa reiknað út tölurnar

$R_1(h)$, $R_1(h/2)$, $\ldots$, $R_1(h/2^i)$ auk
$R_2(h)$, $R_2(h/2)$, $\ldots$, $R_2(h/2^{i-1})$ og svo framvegis að
$\vdots$
$R_i(h)$ og $R_i(h/2)$.

Eins og áður sagði fara styttingarskekkjur á endanum að segja til sín í útreikningum á $R_1(h)$, svo einhver takmörk eru fyrir hversu margar Richardson útgiskanir er hægt að framkvæma.

4.4.7 Reiknirit

Útreikningarnir að ofan eru yfirleitt settir fram í töflu

$$\begin{array}{lllll} D(1,1) & & & & \\ D(2,1) & D(2,2) & & & \\ D(3,1) & D(3,2) & D(3,3) & & \\ \vdots & \vdots & \vdots & \ddots & \\ D(n,1) & D(n,2) & D(n,3) & \ldots & D(n,n) \end{array}$$

þar sem $D(i,j) = R_j(h/2^{i-j})$ og þar með

$$D(i,j) = \begin{cases} \dfrac{f(a+h/2^{i-1}) - f(a-h/2^{i-1})}{2 \cdot h/2^{i-1}}, & j = 1 \\ D(i,j-1) + \dfrac{D(i,j-1) - D(i-1,j-1)}{4^{j-1}-1}, & j > 1 \end{cases}$$

sem gerir okkur auðvelt að forrita Richardson útgiskun.

4.4.8 Skekkjumat

Finnum nú eftirámat fyrir $D(i,j)$ með stærðunum $D(i,j-1)$ og $D(i-1,j-1)$. Hér á eftir er $R_j(h/2)$ í hlutverki $D(i,j-1)$ og $R_i(h)$ í hlutverki $D(i-1,j-1)$ (h er helmingað þegar við förum niður um eina línu).

Munum að $R_i(h)$ uppfyllir að

$$f'(a) = R_j(h) + Kh^{2j} + O(h^{2j+1})$$

fyrir eitthvert K í $\mathbb{R}$ og að

$$f'(a) = R_j(h/2) + K\left(\frac{h}{2}\right)^{2j} + O(h^{2j+1})$$

Ef við tökum mismun á hægri og vinstri hliðum þessara jafna, þá fáum við

$$0 = R_j(h) - R_j(h/2) + K\left(1 - \frac{1}{2^{2j}}\right)h^{2j} + O(h^{2j+1})$$

og ef við einangrum K fæst

$$K = -\frac{4^j}{h^{2j}} \cdot \frac{R_j(h) - R_j(h/2)}{4^j - 1} + O(h^{2j+1}).$$

4.4.9 Útleiðsla á fyrirframmati

Þá er skekkjan í nálgun á $f'(a)$ með $R_j(h/2)$ jöfn

$$\begin{aligned} e_j(h/2) &= f'(a) - R_j(h/2) \\ &= K\left(\frac{h}{2}\right)^{2j} + O(h^{2j+1}) \\ &= -\frac{R_j(h) - R_j(h/2)}{4^j - 1} + O(h^{2j+1}) \\ &\approx -\frac{R_j(h) - R_j(h/2)}{4^j - 1}. \end{aligned}$$

Þar sem $R_j(h/2)$ er nálgun á $f'(a)$ af stigi $O(h^{2j+1})$, en $R_{j+1}(h)$ er nálgun á $f'(a)$ af stigi $O(h^{2i+3})$ getum við slegið á $e_{j+1}(h)$ með $e_j(h/2)$. Ef við lækkum vísinn $j+1$ um einn gefur það okkur matið

$$e_j(h) \approx \frac{R_{j-1}(h) - R_{j-1}(h/2)}{4^{j-1} - 1} = \frac{D(i, j-1) - D(i-1, j-1)}{4^{j-1} - 1}$$

sem er einmitt liðurinn í rakningarformúlunni fyrir $D(i, j)$.

4.4.10 Sýnidæmi

Látum $f(x) = x/(x^2 + 4)^{2/3}$ og $a = -1$. Byrjum með $h = 1$ og notum svo rakningarformúluna til þess að fylla út útgiskunartöfluna.

h	$D(i, 1)$	$D(i, 2)$	$D(i, 3)$	$D(i, 4)$
1 .	0.25000000			
0.5	0.25151838	0.25202451		
0.25	0.25104655	0.25088928	0.25081360	
0.125	0.25086355	0.25080254	0.25079676	0.25079649

Niðustaðan er: $f'(-1) \approx 0.2507964$, með eftirámat á skekkju $-3 \cdot 10^{-7}$.

Rétt gildi er 0.25079647217924889177.

Töluleg heildun

So much universe, and so little time. – Terry Pratchett

Gerum ráð fyrir að $x_0, x_1, \ldots, x_n$ séu punktar á bilinu $[a, b]$ og að við þekkjum gildi f í þessum punktum. Þá getum við fundið brúunarmargliðuna p_n gegnum punktana $(x_k, f(x_k))$ og skrifað

$$f(x) = p_n(x) + r_n(x),$$

þar sem skekkjan r_n er gefin með

$$r_n(x) = f[x_0, \ldots, x_n, x](x - x_0) \cdots (x - x_n).$$

Nú er auðvelt að reikna heildi margliða, svo við nálgum heildi f með

$$\int_a^b f(x)dx \approx I_n(f) := \int_a^b p_n(x)dx$$

og skekkjan í þessari nálgun er gefin með

$$e_n = \int_a^b r_n(x)dx.$$

Þessi aðferð er kölluð *Newton-Cotes-heildun*.

5.1 Aðferðirnar

5.1.1 Newton-Cotes heildun

Hugsum okkur að brúunarpunktarnir $x_0, \ldots, x_n$ séu ólíkir. Þá getum við skrifað p_n með Lagrange-margliðum

$$p_n(x) = \sum_{k=0}^{n} f(x_k)\ell_k(x), \quad \ell_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{(x - x_j)}{(x_k - x_j)},$$

og þá er heildi p_n jafnt

$$\int_a^b p_n(x)dx = \sum_{k=0}^{n} f(x_k)A_k, \quad \text{þar sem} \quad A_k = \int_a^b \ell_k(x)dx.$$

Athugið að gildi A_k veltur aðeins á brúunarpunktunum $x_0, \ldots, x_n$ en ekki gildum $f(x_k)$. Ef það á að heilda mörg föll yfir sama bil er því hægt að reikna gildi A_k í eitt skipti fyrir öll og endurnýta þau svo.

5.1.2 Sýnidæmi

Metum heildi $f(x) = e^{-x}\cos(x)$ og $g(x) = \sin(\frac{x^2}{2})$ yfir bilið $[0, 2]$ með að nota skiptipunktana $x_0 = 0$, $x_1 = 1$ og $x_2 = 2$. Lagrange-margliðurnar sem við eiga eru

$$\ell_0(x) = \frac{(x-1)(x-2)}{2}, \quad \ell_1(x) = -x(x-2), \quad \ell_2(x) = \frac{x(x-1)}{2}$$

svo við fáum að

$$A_0 = \frac{1}{2}\int_0^2 (x-1)(x-2)dx = \frac{1}{3}, \qquad A_1 = -\int_0^2 x(x-2)dx = \frac{4}{3},$$

$$A_2 = \frac{1}{2}\int_0^2 x(x-1)dx = \frac{1}{3}.$$

Nú eru stuðlarnir fundnir og því fáum við

$$\begin{aligned}\int_0^2 f(x)dx &\approx f(0)\frac{1}{3} + f(1)\frac{4}{3} + f(2)\frac{1}{3} \\ &= \frac{1 + 4e^{-1}\cos(1) + e^{-2}\cos(2)}{3} \approx 0.59581\end{aligned}$$

og

$$\begin{aligned}\int_0^2 g(x)dx &\approx g(0)\frac{1}{3} + g(1)\frac{4}{3} + g(2)\frac{1}{3} \\ &= \frac{4\sin(1/2) + \sin(2)}{3} \approx 0.91972.\end{aligned}$$

Gildi heildanna eru $\int_0^2 f(x)dx \approx 0.58969$ og $\int_0^2 g(x)dx \approx 0.99762$ með 5 réttum aukastöfum svo nálgunargildin verða að teljast nokkuð góð miðað við hversu lítið fór í þau.

5.1.3 Trapisureglan

Nú ætlum við að leiða út formúlur fyrir helstu reglum fyrir nálgun á heildum. Sú fyrsta er *trapisuregla*.

Veljum $x_0 = a$ og $x_1 = b$ sem skiptipunktana okkar. Þá er graf p_1 línustrikið gegnum $(a, f(a))$ og $(b, f(b))$,

$$p_1(x) = f(a)\ell_0(x) + f(b)\ell_1(x) = f(a)\frac{b-x}{b-a} + f(b)\frac{x-a}{b-a}$$

og vigtirnar eru

$$A_0 = \int_a^b \ell_0(x) = \frac{b-a}{2} = A_1,$$

svo

$$\int_a^b f(x)dx \approx \frac{b-a}{2}\left(f(a) + f(b)\right).$$

Trapisureglan er kölluð þessu nafni því með henni nálgum við heildi f með flatarmáli trapisunnar sem hefur hornpunktana $(a, 0)$, $(b, 0)$, $(b, f(b))$ og $(a, f(a))$.

5.1.4 Miðpunktsreglan

Enn einfaldari er *miðpunktsreglan*, þá veljum við aðeins einn skiptipunkt, $x_0 = \frac{1}{2}(a+b)$, og brúunarmargliðan verður fastamargliðan $p_0(x) = f(x_0)$. Þá er

$$\int_a^b f(x)dx \approx (b-a)\, f\left(\frac{a+b}{2}\right)$$

5.1.5 Regla Simpsons

Nú veljum við þrjá skiptipunkta, $x_0 = a$, $x_1 = b$ og $x_2 = \frac{1}{2}(a+b)$. Til einföldunar skulum við hliðra fallinu f um miðpunkt bilsins $m = \frac{1}{2}(a+b)$.

Við skilgreinum $\alpha = \frac{1}{2}(b-a)$ og $g(x) = f\big(x+m\big)$

Þá hliðrast a, m og b yfir í $-\alpha$, 0 og α og

$$\int_{-\alpha}^{\alpha} g(x)dx = \int_a^b f(x)dx.$$

Lagrange margliðurnar og vigtirnar eru

$$\begin{aligned}
l_0(x) &= \frac{(x-\alpha)x}{(-\alpha-\alpha)(-\alpha-0)} = \frac{(x-\alpha)x}{2\alpha^2} \\
A_0 &= \int_{-\alpha}^{\alpha} l_0(x)\,dx = \frac{\alpha}{3} \\
l_1(x) &= \frac{(x-(-\alpha))(x-0)}{(\alpha-(-\alpha))(\alpha-0)} = \frac{(x+\alpha)x}{2\alpha^2} \\
A_1 &= \int_{-\alpha}^{\alpha} l_1(x)\,dx = \frac{\alpha}{3} \\
l_2(x) &= \frac{(x-(\alpha))(x-\alpha)}{0-(-\alpha)(0-\alpha)} = \frac{(x+\alpha)(x-\alpha)}{-\alpha^2} \\
A_2 &= \int_{\alpha}^{\alpha} l_2(x)\,dx = \frac{4\alpha}{3}
\end{aligned}$$

Nálgunarformúlan verður þá

$$\begin{aligned}
\int_a^b f(x)\,dx = \int_{-\alpha}^{\alpha} g(x)\,dx &\approx \frac{\alpha}{3}g(-\alpha) + \frac{\alpha}{3}g(\alpha) + \frac{4\alpha}{3}g(0) \\
&= (b-a)\left(\frac{1}{6}f(a) + \frac{4}{6}f\left(\frac{a+b}{2}\right) + \frac{1}{6}f(b)\right)
\end{aligned}$$

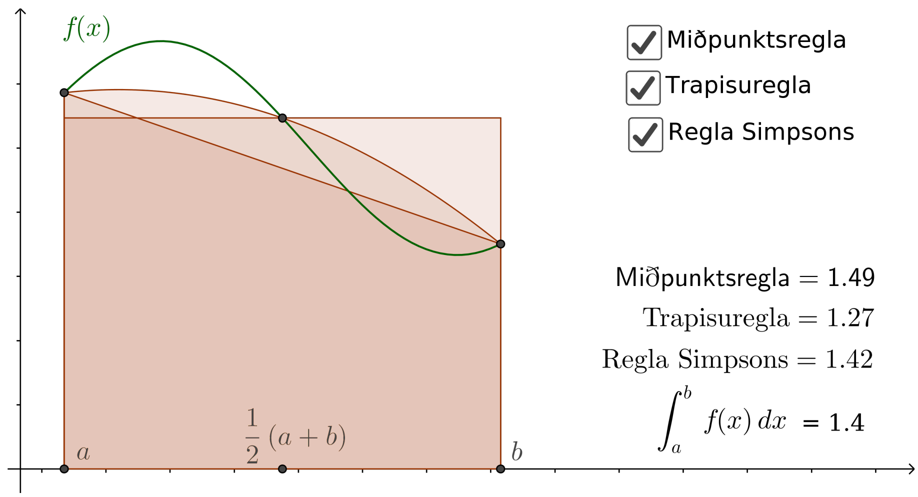
Ef við tökum brúunarmargliðu gegnum a, b og $\frac{1}{2}(a+b)$ með $\frac{1}{2}(a+b)$ tvöfaldan þá fáum við 3. stigs brúunarmargliðu

$$p_3(x) = p_2(x) + g[-\alpha,\alpha,0,0](x+\alpha)(x-\alpha)x$$

Heildið yfir seinni liðinn hægra megin er 0 því margliðan $(x+a)(x-a)x$ er oddstæð, en heildið yfir fyrri liðinn er

$$\frac{\alpha}{3}(g(-\alpha) + 4g(0) + g(\alpha)).$$

Út kemur því Simpson-regla.



5.2 Samsettar útgáfur

Sometimes the truth is arrived at by adding all the little lies together and deducting them from the totality of what is known. – Terry Pratchett, Going Postal

5.2.1 Inngangur

Þar sem Newton-Cotes heildun notar brúunarmargliður fylgja henni nokkur vandamál.

Ef okkur finnst nákvæmnin í nálguninni vera of lítil getum við ekki búist við að hún batni við að fjölga skiptipunktum; þá hækkar stig margliðunnar líklega sem orsakar sveiflukenndari hegðun.

Eins er ekki gott að halda sig við margliður af lægra stigi; ef bilið sem á að heilda yfir er stórt væri mikil tilviljun að 1., 2. eða 3. stigs brúunarmargliða nálgaði fallið vel á öllu bilinu.

Lausnin á þessu vandamáli er í sama anda og fyrir splæsibrúun. Við veljum skiptingu

$$a = x_0 < x_1 < \ldots < x_n = b$$

á bilinu $[a, b]$.

Um heildi gildir að

$$\int_a^b f(x)\,dx = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)\,dx$$

svo við getum nálgað heildi f á sérhverju litlu hlutbili $[x_{k-1}, x_k]$ með að heilda brúunarmargliðu af lágu stigi og lagt öll gildin saman til að fá nálgun á heildi f yfir allt bilið.

Þegar ákveðin regla er notuð til að nálga heildi f á sérhverju hlutbili er þetta kölluð *samsetta* útgáfa reglunnar. Einfalt er að leiða út samsettar útgáfur reglanna að ofan.

5.2.2 Samsetta trapisureglan

Á sérhverju hlutbili er

$$\int_{x_{k-1}}^{x_k} f(x)\,dx \approx \frac{x_k - x_{k-1}}{2}(f(x_{k-1}) + f(x_k))$$

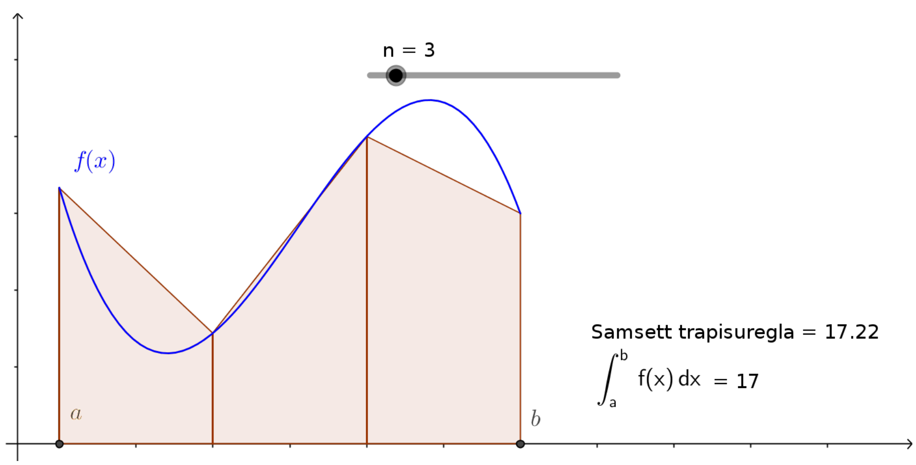
svo

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{n} \frac{x_k - x_{k-1}}{2}(f(x_{k-1}) + f(x_k)).$$

Ef öll hlutbilin eru jafn löng og $h = x_k - x_{k-1}$, þá fæst

$$\int_a^b f(x)\,dx$$

$$\approx h\left(\frac{1}{2}f(a) + f(a+h) + f(a+2h) + \cdots + f(a+(n-1)h) + \frac{1}{2}f(b)\right).$$



5.2.3 Samsetta miðpunktsreglan

Fljótséð er að

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{n}(x_k - x_{k-1}) f\left(\frac{x_{k-1} + x_k}{2}\right)$$

Ef öll hlutbilin eru jafn löng verður formúlan

$$\int_a^b f(x)\,dx \approx h \sum_{k=1}^{n} f\left(\frac{x_{k-1} + x_k}{2}\right)$$

5.2.4 Samsett regla Simpsons

Hér er venjan að velja $2n+1$ jafndreifða skiptipunkta og fá n jafn stór hlutbil. Þá er $h = \frac{b-a}{2n}$, $x_k = a + kh$ fyrir $k = 0, \dots, 2n$ og hlutbilin eru $[x_{2k-2}, x_{2k}]$ fyrir $k = 1, \dots, n$.

Á hverju hlutbili er

$$\int\limits_{x_{2k-2}}^{x_{2k}} f(x)\,dx \approx 2h\left(\frac{1}{6}f(x_{2k-2}) + \frac{4}{6}f(x_{2k-1}) + \frac{1}{6}f(x_{2k})\right)$$

svo að

$$\begin{aligned}\int\limits_a^b f(x)\,dx \approx &\sum_{k=1}^n \left(\frac{h}{3}\Big(f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})\Big)\right)\\ =&\frac{h}{3}\Big(f(a) + 4f(a+h) + 2f(a+2h) + 4f(a+3h) + 2f(a+4h)\\ &+ \cdots + 2f(a+(2n-2)h) + 4f(a+(2n-1)h) + f(b).\Big)\end{aligned}$$

5.3 Skekkjumat

5.3.1 Inngangur

Rifjum upp grunnhugmyndina að baki nálgunarformúlunum. Við veljum brúunarpunkta $x_0, \dots, x_n$ í $[a, b]$, látum p_n vera tilsvarandi brúunarmargliðu og skrifum

$$f(x) = p_n(x) + r_n(x)$$

þar sem $r_n(x) = f[x_0, \dots, x_n, x](x - x_0) \cdots (x - x_n)$. Þá er nálgunin

$$\int_a^b f(x)\,dx \approx \int_a^b p_n(x)\,dx$$

með skekkjuna

$$\int_a^b r_n(x)\,dx$$

Nú viljum við meta skekkjuheildið.

5.3.2 Meðalgildissetningin fyrir heildi

Við skekkjumatið í þessum kafla munum við þurfa að nota eftirafarandi setningu nokkrum sinnum.

Ef $G : [a, b] \to \mathbb{R}$ er samfellt fall og φ er heildanlegt fall sem skiptir ekki um formerki á bilinu $[a, b]$ þá er til tala $\eta \in [a, b]$ þannig að

$$\int_a^b G(x)\varphi(x)\,dx = G(\eta)\int_a^b \varphi(x)\,dx.$$

5.3.3 Trapisureglan

Við getum hliðrar sérhverju bili $[a, b]$ yfir í $[-\alpha, \alpha]$ þar sem $\alpha = (b - a)/2$, því er nóg fyrir okkur að skoða samhverf bil af gerðinni $[-\alpha, \alpha]$. Þetta er það sama og við gerðum þegar regla Simpsons var leidd út.

Samkvæmt 3.7.7 þá er

$$r_1(x) = f[-\alpha, \alpha, x](x + \alpha)(x - \alpha)$$

Athugum að

$$(x + \alpha)(x - \alpha) = (x^2 - \alpha^2)$$

skiptir ekki um formerki á bilinu $] - \alpha, \alpha[$. Þá gefur meðalgildissetningin fyrir heildi að til er $\eta \in [a, b]$ þannig að

$$\begin{aligned} \int_a^b r_1(x)\,dx &= f[-\alpha, \alpha, \eta] \int_{-\alpha}^{\alpha} (x^2 - \alpha^2)\,dx \\ &= \frac{f''(\xi)}{2!} \left(-\frac{4}{3}\alpha^3 \right) \\ &= \frac{-f''(\xi)}{2!} \frac{(b-a)^3}{6}, \qquad \xi \in [a, b] \end{aligned}$$

Niðurstaða:

$$\int_a^b f(x)\,dx = (b - a) \left(\frac{1}{2} f(a) + \frac{1}{2} f(b) \right) - \frac{1}{12} f''(\xi)(b - a)^3$$

5.3.4 Samsetta trapisureglan

Ef við lítum á samsettu trapisuregluna með jafna skiptingu þar sem hlutbilin eru $[x_i, x_{i+1}]$, þá fáum við fyrir hvert hlutbil skekkjuna

$$-\frac{h^3}{12} f''(\xi_i), \qquad \xi_i \in [x_i, x_{i+1}]$$

Ef við leggjum skekkjurnar saman og beitum milligildissetningunni á f'' þá fæst að til er $\xi \in [a, b]$ þannig að $f''(\xi) = \sum_{i=1}^n f''(\xi_i)/n$. Þá fáum við a'

$$\int_a^b f(x)\,dx = T(h) - \frac{h^2}{12}(b - a) f''(\xi), \qquad \xi \in [a, b]$$

að því gefnu að $f \in C^2[a, b]$.

Athugið að hér er $T(h)$ útkoman úr samsettu Trapisureglunni með jafna skiptingu $h = \frac{b-a}{n}$.

5.3.5 Miðpunktsregla

Til einföldunar skoðum við áfram bilið $[-\alpha, \alpha]$. Veljum miðpunktinn sem tvöfaldan brúunarpunkt

$$\begin{aligned} p_1(x) &= f(0) + f'(0)x \\ r_1(x) &= f[0, 0, x]x^2 \end{aligned}$$

Athugum að heildið af $f'(0)x$ yfir $[-\alpha,\alpha]$ er 0. Nú skiptir x^2 ekki um formerki og því gefur meðalgildisreglan fyrir heildi að til er $\eta \in [-\alpha,\alpha]$ þannig að

$$\begin{aligned}\int_a^b r_1(x)\,dx &= \int_{-\alpha}^{\alpha} f[0,0,x]x^2\,dx \\ &= f[0,0,\eta]\int_{-\alpha}^{\alpha} x^2\,dx \\ &= \frac{f''(\xi)}{2!}2\frac{\alpha^3}{3} \\ &= \frac{(b-a)^3}{24}\cdot f''(\xi)\end{aligned}$$

Þar sem ξ fæst úr skekkjumatinu fyrir brúunarmargliður.

5.3.6 Samsetta miðpunktsreglan

Fyrir hvert bil fáum við skekkjulið:

$$\frac{h^3}{24}\cdot f''(\xi_i)$$

Leggjum saman skekkjuliðina og beitum milligildissetningunni, þá fæst að til er ξ þannig að:

$$\int_a^b f(x)\,dx = h\sum_{i=1}^{n} f\left(a+(i-\frac{1}{2})h\right) + \frac{b-a}{24}f''(\xi)h^2$$

5.3.7 Regla Simpsons

$$\int_a^b f(x)\,dx \approx (b-a)\left(\frac{1}{6}f(a)+\frac{4}{6}f\left(\frac{1}{2}(a+b)\right)+\frac{1}{6}f(b)\right)$$

Leiddum út þessa formúlu með því að taka brúunarmargliðu $p_3(x)$ með punktana $-\alpha,\alpha,0,0$. Skekkjan er

$$f(x)-p_3(x) = f[-\alpha,\alpha,0,0,x](x+\alpha)(x-\alpha)x^2$$

þar með er skekkjan í formúlu Simpsons:

$$\int_{-\alpha}^{\alpha} f[-\alpha,\alpha,0,0,x](x+\alpha)(x-\alpha)x^2\,dx$$

Fallið $x\mapsto (x+\alpha)(x-\alpha)x^2 = (x^2-\alpha^2)x^2$ er ≤ 0 á $[-\alpha,\alpha]$. Þar með gefur meðalgildissetningin fyrir heildi að til er $\eta\in[-\alpha,\alpha]$ þannig að skekkjan er

$$\begin{aligned}&f[-\alpha,\alpha,0,0,\eta]\int_{-\alpha}^{\alpha}(x^2-\alpha^2)x^2\,dx \\ &= \frac{f^{(4)}(\xi)}{4!}\cdot\frac{(-4)}{15}\cdot\alpha^5 = \frac{-f^{(4)}(\xi)}{90}\left(\frac{b-a}{2}\right)^5, \qquad \xi\in[a,b]\end{aligned}$$

Þar sem ξ fæst úr skekkjumatinu fyrir brúunarmargliður.

5.3.8 Samsett regla Simpsons

Skiptum $[a, b]$ í n jafnlöng bil og látum h vera helming hlutbillengdarinnar,

$$h = \frac{(b-a)}{2n}.$$

Þá er

$$\begin{aligned}\int\limits_a^b f(x)\,dx \approx & \sum_{k=1}^{n}\left(\frac{h}{3}\Big(f(x_{2k-2}) + 4f(x_{2k-1}) + f(x_{2k})\Big)\right)\\ =&\frac{h}{3}\Big(f(a) + 4f(a+h) + 2f(a+2h) + 4f(a+3h) + 2f(a+4h)\\ &+ \cdots + 2f(a+(2n-2)h) + 4f(a+(2n-1)h) + f(b)\Big)\end{aligned}$$

Ef við beitum skekkjumatinu á sérhvert bilanna þá fáum við

$$\frac{-f^{(4)}(\xi_i)}{90}h^5$$

sem skekkju með $\xi_i \in [x_i, x_i + 1]$. Heildarskekkjan verður

$$-\sum_{i=1}^{n}\frac{f^{(4)}(\xi_i)}{90}h^5 = \frac{-h^5}{90}\cdot\sum_{i=1}^{n}f^{(4)}(\xi_i)$$

Nú gefur meðalgildisreglan að til er $\xi \in [a, b]$ þannig að

$$f^{(4)}(\xi) = \frac{1}{n}\sum_{i=1}^{n}f^{(4)}(\xi_i)$$

Nú er $nh = \frac{(b-a)}{2}$ þar með er skekkjan:

$$\frac{-h^5}{90}\cdot nf^{(4)}(\xi) = \frac{-(b-a)}{180}f^{(4)}(\xi)\cdot h^4$$

Ef við táknum útkomuna úr samsettu Simpsonsreglunni fyrir $h = \frac{b-a}{2n}$ með $S(h)$ þá fæst að til er $\xi \in [a, b]$ þannig að

$$\int_a^b f(x)\,dx = S(h) - \frac{(b-a)}{180}f^{(4)}(\xi)h^4$$

5.4 Romberg-útgiskun

Á sama hátt og við gátum bætt nálgun okkar á afleiðu falls með að nota Richardson útgiskun getum við bætt nálgun á heildi.

Aðferðin virkar í aðalatriðum eins fyrir heildi og afleiður, en til að fá sem bestar upplýsingar um samleitni hennar skulum við leiða út formúluna fyrir trapisureglunni aftur.

5.4.1 Euler-Maclauren-formúlan

Fyrir samfellt fall $f : [0, 1] \to \mathbb{R}$ sem er $2n$-sinnum samfellt deildanlegt gildir Euler-Maclauren formúlan

$$\begin{aligned}\int\limits_0^1 f(t)\,dt =&\frac{1}{2}\left(f(0) + f(1)\right) + \sum_{k=1}^{n-1}A_{2k}\left(f^{(2k-1)}(0) - f^{(2k-1)}(1)\right)\\ &- A_{2n}f^{(2n)}(\xi), \qquad \xi \in [0, 1]\end{aligned}$$

Hér eru stuðlarnir A_k þannig að $k!A_k$ verði Bernoulli-talan númer k. Þessar tölur eru stuðlar í veldaröðinni

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} A_k x^k$$

Athugasemd: Það þarf að hafa töluvert fyrir því að sanna þessa formúlu og því sleppum við því hér.

5.4.2 Afleiðing af Euler-Maclaurin-formúlunni

Látum nú $f : [a, b] \to \mathbb{R}$ vera $2n$-sinnum samfellt deildanlegt fall. Ef við búum til skiptingu $a = x_0 < x_1 < \cdots < x_n = b$ með jöfn hlutbil $h = x_{i+1} - x_i$ og beitum síðan Euler-Maclauren formúlunni á $g(t) = f(x_i + ht)$ fæst

$$\begin{aligned}
\int_{x_i}^{x_{i+1}} f(x)\,dx =& h \int_0^1 \underbrace{f(x_i + ht)}_{g(t)}\,dt \\
=& h\left(\frac{1}{2} f(x_i) + \frac{1}{2} f(x_{i+1})\right) \\
&+ \sum_{k=1}^{n-1} A_{2k} h^{2k} \left(f^{(2k-1)}(x_i) - f^{(2k-1)}(x_{i+1})\right) - A_{2n} h^{2n+1} f^{(2n)}(\xi_i),
\end{aligned}$$

þar sem $\xi_i \in [x_i, x_{i+1}]$.

Nú innleiðum við

$$\begin{aligned}
T(h) &:= \sum_{i=0}^{n-1} h\left(\frac{1}{2} f(x_i) + \frac{1}{2} f(x_{i+1})\right) \\
&= h\left(\frac{1}{2} f(a) + f(a+h) + \cdots + f(a + (n-1)h) + \frac{1}{2} f(a + nh)\right)
\end{aligned}$$

og fáum síðan:

$$\begin{aligned}
\int_a^b f(x)\,dx =& T(h) + \sum_{k=1}^{n-1} A_{2k} h^{2k} \left(f^{(2k-1)}(a) - f^{(2k-1)}(b)\right) \\
&- A_{2n} h^{2n+1} \sum_{i=0}^{n-1} f^{(2n)}(\xi_i)
\end{aligned}$$

Nú gefur milligildissetningin að til er $\xi \in [a, b]$ þannig að

$$\frac{1}{n} \sum_{k=0}^{n-1} f^{(2n)}(\xi_i) = f^{(2n)}(\xi)$$

Notum okkur nú að $nh = b - a$ og fáum að

$$\begin{aligned}
\int_a^b f(x)\,dx =& T(h) + \sum_{k=1}^{n-1} A_{2k} h^{2k} \left(f^{(2k-1)}(a) - f^{(2k-1)}(b)\right) \\
&- A_{2n} h^{2n} (b - a) f^{(2n)}(\xi).
\end{aligned}$$

Niðurstaðan er að samsetta trapisureglan er

$$\int_a^b f(x)\,dx = T(h) + c_2 h^2 + c_4 h^4 + \cdots + c_{2m-2} h^{2m-2} + c_{2n} h^{2m} f^{(2m)}(\xi)$$

5.4.3 Ítrekun á samsettu trapisureglunni með helmingun

Hugsum okkur nú að við viljum reikna út $T(h_j)$ fyrir $h_j = (b-a)/2^j$, $j = 1, 2, \dots$ og að við viljum nýta öll fallgildi í $T(h_{j-1})$ til að reikna út $T(h_j)$. Rakningarformúlan er

$$T(h_j) = \frac{1}{2}T(h_{j-1}) + h_j \sum_{k=1}^{2^{j-1}} f(a + (2k-1)h_j)$$

Athugið að hér er bilinu $[a, b]$ skipt í 2^j hlutbil.

> **Aðvörun:** Þetta er gott að nota ef forrita á Romberg-heildun til þess að spara útreikninga þegar fyrsti dálkurinn er reiknaður. Það er hins vegar ekki nauðsynlegt að nota þetta og þetta tengist ekki beint Romberg aðferðinni.

5.4.4 Reikniritið fyrir Romberg-heildun

Romberg-heildun er hugsuð nákvæmlega eins og Richardson-útgiskunin: Við reiknum út línu fyrir línu í töflunni:

$$\begin{array}{cccccc}
i & & & & & \\
1 & R(1,1) & & & & \\
2 & R(2,1) & R(2,2) & & & \\
3 & R(3,1) & R(3,2) & R(3,3) & & \\
4 & R(4,1) & R(4,2) & R(4,3) & R(4,4) & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

þar sem

$$\begin{aligned}
R(i,1) &= T(h_i) \qquad i = 1, 2, \dots \\
R(i,j) &= \frac{4^{j-1}R(i,j-1) - R(i-1,j-1)}{4^{j-1}-1}.
\end{aligned}$$

Með þessu fæst $\int\limits_a^b f(x)\,dx = R(k,k) + O(h_k^{2k})$, þar sem k er síðasta línan sem við reiknum í töflunni að ofan.

5.4.5 Skekkjumat í Romberg heildun

Skekkjumatið er hægt að finna með nákvæmlega sama hætti í fyrir Richardson útgiskuna. Þ.e. við getum notað síðustu viðbót sem eftirámat fyrir skekkjuna, þetta mat er

$$e \approx \frac{1}{4^{j-1}-1}\left(R(i,j-1) - R(i-1,j-1)\right)$$

þegar þessi stærð er komin niður fyrir fyrirfram gefin skekkjumörk er hætt.

Einnig er hægt að nota

$$e \approx \frac{1}{2^{j-1}}\left(R(i,j-1) - R(i-1,j-1)\right),$$

sem gefur heldur varfærnislegra mat.

Athugasemd: Athugið að það er ekki nauðsynlegt að hafa h_1 sem allt bilið $[a, b]$, það er ekkert sem kemur í veg fyrir það að við byrjum með $h_1 = \frac{b-a}{m}$, og helmingum svo; $h_2 = \frac{b-a}{2m}$, $h_3 = \frac{b-a}{4m}$, ... Þannig að almennt þá er $h_j = \frac{b-a}{2^{j-1}m}$.

KAFLI 6

Upphafsgildisverkefni

In the beginning there was nothing, which exploded. – Terry Pretchett

6.1 Inngangur

6.1.1 Fyrsta stigs afleiðujafna með upphafsgildi

Gefið fall f á einhverju svæði U í $\mathbb{R}^2$ sem inniheldur (t_0, x_0) þá er *fyrsta stigs afleiðujafna með upphafsgildi* verkefni á forminu

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0. \end{cases}$$

Við segjum að fall $x(t)$ sé lausn á þessu verkefni ef x er skilgreint á bili I, sem er þannig að

- $t_0 \in I$,
- $(t, x(t)) \in U$ fyrir öll $t \in I$,
- $x'(t) = f(t, x(t))$ fyrir öll $t \in I$, og
- $x(t_0) = x_0$.

6.1.2 Útskýring og dæmi

Ástæðan fyrir því að við notum t fyrir breytuna og x fyrir fall hér, er að breytan t lýsir oft tíma og lausnin getur þá t.d. verið staðsetning sem fall af tíma.

Eðlilegt er að hugsa um afleiðujöfnuna þannig að fallið f geymi upplýsingar um þau lögmál sem kerfið að hagar sér eftir, upphafsskilyrðið $x(t_0) = x_0$ segir í hvaða stöðu kerfið er þegar það er sett af stað og lausnin $x(t)$ lýsir hegðun kerfisins með tíma.

Í upphafi námskeiðsins, dæmi 1.2, skoðuðum við dæmi um eldflaug sem skotið er á loft. Þar er notum við reyndar v í stað x því jafnan lýsir hraða (e. velocity) en ekki staðsetningu. Jafnan var

$$v' = \frac{5000 - (300 - 10t)g - 0,1v^2 + 10v}{300 - 10t}, \qquad v(0) = 0.$$

Hér er því $f(t, v) = \frac{5000-(300-10t)g-0,1v^2+10v}{300-10t}$, $t_0 = 0$ og $x_0 = 0$.

6.1.3 Tilvist og ótvíræðni lausna

$$\begin{cases} x' = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

Ef f er samfellt, þá er alltaf til lausn á einhverju bili I. (Setning Peano)

Ef f uppfyllir Lipschitz-skilyrði með tilliti til x, þ.e.a.s. til er fasti C þannig að

$$|f(t,x_1) - f(t,x_2)| \leq C|x_1 - x_2|$$

fyrir öll (t,x_1) og (t,x_2) í grennd um (t_0,x_0) þá er lausnin ótvírætt ákvörðuð. (Setning Picard)

6.1.4 Upphafsgildisverkefni fyrir hneppi

Með því að nota vigra og vigurgild föll má útfæra upphafsgildisverkefni fyrir hneppi:

$$\begin{cases} \mathbf{x}' = \mathbf{f}(t,\mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

þar sem

$$\mathbf{f} : U \to \mathbb{R}^n, \qquad U \subset \mathbb{R}^{n+1}, \quad (t_0, \mathbf{x}_0) \in U.$$

Athugið að við skrifum $\mathbf{x}(t)$ og $\mathbf{f}(t,\mathbf{x})$ sem dálkvigra,

$$\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_n(t)]^T \quad \text{og} \quad \mathbf{f}(t,\mathbf{x}) = [f_1(t,\mathbf{x}), \ldots, f_n(t,\mathbf{x})]^T$$

6.1.5 Jöfnur af stigi > 1 og jafngild hneppi

Aðferðirnar sem við munum skoða eru eingöngu fyrir fyrsta stigs afleiðujöfnur, sem þýðir að jöfnurnar sem við leysum innihalda bara x' en ekki $x'', x''', \ldots$. Hins vegar þá getum við leyst afleiðujöfnur af hærra stigi með því að umrita þær yfir í jafngilt fyrsta stigs hneppi.

Ef við höfum m-stigs diffurjöfnu

$$\begin{gathered} u^{(m)} = g(t, u, u', u'', \ldots, u^{(m-1)}) \\ u(t_0) = u_0, \quad u'(t_0) = u_1, \quad \ldots, \quad u^{m-1}(t_0) = u_{(m-1)} \end{gathered}$$

þar sem g er gefið fall og $t, u_0, \ldots, u_{m-1}$ eru gefnar tölur þá er jafngilt hneppi er fengið með því að setja

$$\begin{aligned} x_1 &= u, \\ x_2 &= u', \\ x_3 &= u'', \\ &\vdots \\ x_m &= u^{(m-1)} \end{aligned}.$$

Þá fæst hneppið

$$\mathbf{x}' = \begin{pmatrix} x_1' & = x_2 \\ x_2' & = x_3 \\ \vdots & \vdots \\ x_{m-1}' & = x_m \\ x_m' & = g(t, x_1, \ldots, x_m) \end{pmatrix} = \mathbf{f}(t,\mathbf{x})$$

með upphafsskilyrðið $\mathbf{x}(t_0)^T = [u_0, u_1, u_2, \ldots, u_m]^T$.

Fyrsta hnitið í lausn hneppisins, x_1, gefur þá lausn, u á upprunalegu m-ta stigs afleiðujöfnunni.

6.1.6 Tilvist og ótvíræðni lausna á hneppum

Tilvistar- og ótvíræðnisetningar Peanos og Picards eru þær sömu fyrir hneppi

$$\begin{cases} \mathbf{x}' = \mathbf{f}(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

Við þurfum bara að setja norm $\|\cdot\|$ í stað tölugildis $|\cdot|$ í öllum ójöfnum og þar með talið í Lipschitz-skilyrðinu.

6.1.7 Ritháttur

Til einföldunar á rithætti skulum við skrifa lausnarvigurinn $\mathbf{x}$ og vörpunina $\mathbf{f}$ sem x og f og láta eins og við séum að leysa fyrsta stigs afleiðujöfnu.

Við veljum gildi $t_0 < t_1 < \cdots < t_j < \cdots$ og reiknum út *nálgunargildi* w_j á gildi lausnarinnar $x(t_j)$ í punktinum t_j. Gildið $w_0 = x(t_0)$ er rétta upphafsgildi lausnarinnar

Talan t_j kallast j-ti *tímapunkturinn* og talan $h_j = t_j - t_{j-1}$ nefnist j-ta *tímaskrefið*.

Skekkjan á tíma t_j er þá $e_j = x(t_j) - w_j$.

6.1.8 Grunnhugmyndin í nálgunaraðferðum

Ef við heildum lausn afleiðujöfnunnar yfir tímabilið $[t, t+h]$, þá fáum við að hún uppfyllir jöfnuna

$$x(t+h) = x(t) + \int_t^{t+h} f(\tau, x(\tau))\, d\tau = x(t) + h \int_0^1 f(t+sh, x(t+sh))\, ds.$$

Ef við setjum $t = t_{j-1}$ inn í þessa jöfnu, þá fáum við

$$\frac{x(t_j) - x(t_{j-1})}{h_j} = \int_0^1 f(t_{j-1} + sh_j, x(t_{j-1} + sh_j))\, ds$$

Nálgunaraðferðirnar snúast allar um að gera einhvers konar nálgun á heildinu í hægri hliðinni

$$\int_0^1 f(t_{j-1} + sh_j, x(t_{j-1} + sh_j))\, ds \approx \varphi(f, t_0, \ldots, t_j, w_0, \ldots, w_j)$$

og leysa síðan w_j út úr jöfnunni

$$\frac{w_j - w_{j-1}}{h_j} = \varphi(f, t_0, \ldots, t_j, w_0, \ldots, w_j)$$

6.1.9 Beinar og óbeinar aðferðir

Nálgunaraðferð sem byggir á jöfnunni

$$\frac{w_j - w_{j-1}}{h_j} = \varphi(f, t_0, \ldots, t_j, w_0, \ldots, w_j)$$

er nefnist *bein aðferð* (e. explicit method) ef w_j kemur ekki fyrir í í hægri hliðinni.

Annars nefnist hún *óbein aðferð* eða *fólgin aðferð* (e. implicit method).

Ef aðferðin er bein og við höfum reiknað út $w_0, \dots, w_{j-1}$, þá fáum við rakningarformúlu, þannig að $w_j \approx x(t_j)$ er reiknað út

$$w_j = w_{j-1} + h_j \varphi(f, t_0, \dots, t_j, w_0, \dots, w_{j-1})$$

6.1.10 Eins skrefs aðferðir og fjölskrefaaðferðir

Nálgunaraðferð sem byggir á jöfnunni

$$\frac{w_j - w_{j-1}}{h_j} = \varphi(f, t_{j-1}, t_j, w_{j-1}, w_j)$$

er nefnist *eins skrefs aðferð* (e. one step method) og er þá vísað til þess að fallið í hægri hliðinni er einungis háð gildum á síðasta tímaskrefinu.

er af gerðinni

$$\frac{w_j - w_{j-1}}{h_j} = \varphi(f, t_{j-2}, t_{j-1}, t_j, w_{j-2}, w_{j-1}, w_j)$$

Almennt er k *-skrefa aðferð* af gerðinni

$$\frac{w_j - w_{j-1}}{h_j} = \varphi(f, t_{j-k}, \dots, t_j, w_{j-k}, \dots, w_j)$$

Fjölskrefaðferð er k-skrefa aðferð með $k \geq 2$.

6.2 Aðferðir með fasta skrefastærð

6.2.1 Aðferð Eulers

Rifjum upp að lausnin uppfyllir

$$\begin{aligned} x(t+h) - x(t) &= \int_t^{t+h} x'(\tau)\, d\tau = \int_t^{t+h} f(\tau, x(\tau))\, d\tau \\ &= h \int_0^1 f(t+sh, x(t+sh))\, ds \end{aligned}$$

Billengdin í síðasta heildinu er 1, svo við tökum einföldustu nálgum sem hugsast getur en það er gildið í vinstri endapunkti $f(t, x(t))$. Fyrir lítil h fæst því

$$x(t+h) \approx x(t) + hf(t, x(t)).$$

Við þekkjum $w_0 = x(t_0)$, svo með þessu getum við fikrað okkur áfram og fengið runu nálgunargilda $w_0, w_1, w_2, \dots$ þannig að

$$w_j = w_{j-1} + h_j f(t_{j-1}, w_{j-1}).$$

6.2.2 Aðferð Eulers: Matlab-forrit

```
function w = euler(f,t,alpha);
%   function w = euler(f,t,alpha)
% Aðferð Eulers fyrir afleiðujöfnuhneppi
%         x'(t)=f(t,x(t)), x(0)=alpha.
% Inn fara: f - fallið f
%           t - vigur með skiptingu á t-ás.
%           alpha - upphafsgildið í t(1).
% Út koma:  w - fylki með nálgunargildunum.

N = length(t);
m = length(alpha);
w = zeros(m,N);
w(:,1) = alpha;
for j=2:N
   w(:,j) = w(:,j-1)+(t(j)-t(j-1))*f(t(j-1),w(:,j-1));
end
```

6.2.3 Aðferð Eulers: Dæmi

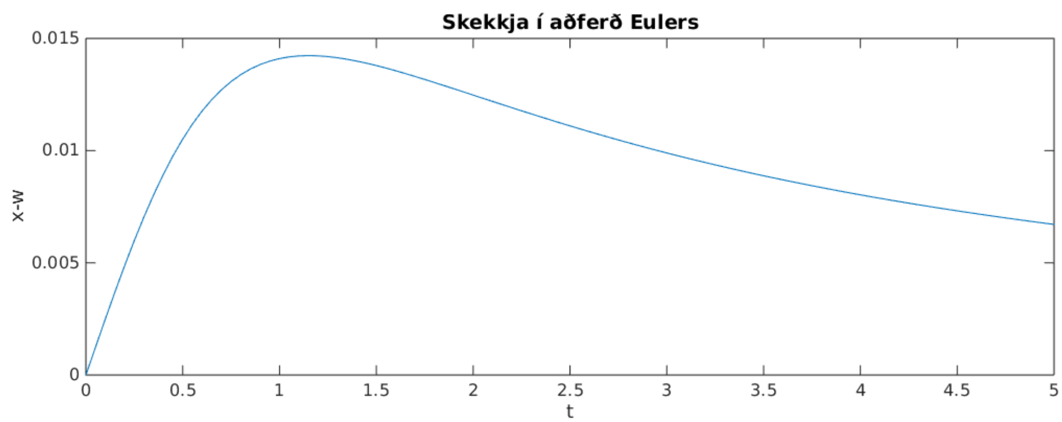
Prófum aðferð Eulers á afleiðujöfnunni

$$x' = \frac{t}{x}, \qquad x(0) = 1$$

Við sjáum að rétt lausn er $x(t) = \sqrt{t^2+1}$.

Notum 101 jafndreifð tímagildi á bilinu [0,5]. Þá er skekkjan

```
>> f = @(t,x) t./x;
>> t=linspace(0,5,101);
>> w=euler(f,t,1);
>> plot(t,sqrt(t.^2+1) - w)
>> title('Skekkja í aðferð Eulers'); xlabel('t'); ylabel('x-w');
```



6.2.4 Endurbætt aðferð Eulers

Í aðferð Eulers nálguðum við heildið $\int_0^1 f(t+sh, x(t+sh))\,ds$ með margfeldi af billengdinni og fallgildinu í vinstri endapunkti.

Við getum endurbætt þessa nálgun með því að taka einhverja nákvæmari tölulega nálgun á heildinu til dæmis miðpunktsaðferð

Nálgunarformúlan verður þá

$$\int_0^1 f(t+sh, x(t+sh))\,ds \approx f(t+\tfrac{1}{2}h, x(t+\tfrac{1}{2}h)).$$

Nú er vandamálið að við höfum nálgað $x(t_{j-1})$ með w_{j-1} en höfum ekkert nálgunargildi á $x(t_{j-1}+\frac{1}{2}h_j)$.

Við grípum þá til fyrsta stigs Taylor nálgunar

$$\begin{aligned} x(t_j+\tfrac{1}{2}h_j) &= x(t_{j-1}) + x'(t_{j-1})\big(\tfrac{1}{2}h_j\big) + \tfrac{1}{2}x''(\xi)\big(\tfrac{1}{2}h_j\big)^2 \\ &\approx w_{j-1} + \tfrac{1}{2}h_j f(t_{j-1}, w_{j-1}). \end{aligned}$$

Endurbætt aðferð Eulers er þá í tveim skrefum; við reiknum

$$\tilde{w}_j = w_{j-1} + \tfrac{1}{2}h_j f(t_{j-1}, w_{j-1})$$

og fáum svo nálgunargildið

$$w_j = w_{j-1} + h_j f\left(t_{j-1} + \tfrac{1}{2}h_j, \tilde{w}_j\right)$$

6.2.5 Aðferð Heun

Lítum nú á aðra aðferð þar sem við nálgum heildið með trapisuaðferð.

$$\int_0^1 f(t+sh, x(t+sh))\,ds \approx \tfrac{1}{2}\big(f(t,x(t)) + f(t+h, x(t+h))\big).$$

Af þessu leiðir að nálgunarformúlan á að vera

$$w_j = w_{j-1} + \tfrac{1}{2}h_j\big(f(t_{j-1}, w_{j-1}) + f(t_j, w_j)\big)$$

Þetta er greinilega óbein aðferð svo við verðum að byrja á nálgun á w_j, með

$$w_j \approx x(t_j) = x(t_{j-1}+h_j) \approx x(t_{j-1}) + h_j x'(t_{j-1}) = x(t_{j-1}) + h_j f(t_{j-1}, w_{j-1})$$

Þetta nýja afbrigði af aðferð Eulers nefnist *aðferð Heun*. Hún er í tveim skrefum: Við reiknum fyrst

$$\tilde{w}_j = w_{j-1} + h_j f(t_{j-1}, w_{j-1})$$

og fáum svo nálgunargildið

$$w_j = w_{j-1} + \tfrac{1}{2}h_j\big(f(t_{j-1}, w_{j-1}) + f(t_j, \tilde{w}_j)\big)$$

6.2.6 Forsagnar- og leiðréttingarskref

Endurbætt aðferð Eulers og aðferð Heun eru leiðir til þess að vinna úr óbeinum aðferðum, þar sem rakningarformúlan fyrir nálgunargildin er af gerðinni

$$w_j = w_{j-1} + h_j\varphi(f, t_{j-1}, t_j, w_{j-1}, w_j)$$

og okkur vantar eitthverja nálgun á w_j til þess að stinga inn í hægri hlið þessarar jöfnu. Við skiptum þessu tvö skref:

Við beitum einhverri beinni aðferð til þess að reikna út

$$\tilde{w}_j = w_{j-1} + h_j\psi(f, t_{j-1}, t_j, w_{j-1})$$

Setjum

$$w_j = w_{j-1} + h_j\varphi(f, t_{j-1}, t_j, w_{j-1}, \tilde{w}_j).$$

Svona aðferðir kallast *Runge-Kutta aðferðir*. Fyrra skrefið, þegar $\tilde{w}_j$ er reiknað út kallast *forsagnarskref* og seinna skrefið kallast *leiðréttingarskref*.

6.2.7 Annars stigs Runge-Kutta-aðferð

Lítum aftur á verkefnið

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}$$

og skoðum 2. stigs Taylor liðun á lausninni x í punkti t. Innleiðum fyrst smá rithátt til styttingar, setjum

$$x = x(t), \quad f_t' = \frac{\partial f}{\partial t}(t, x(t)), \quad f = f(t, x(t)), \quad f_x' = \frac{\partial f}{\partial x}(t, x(t)).$$

Keðjureglan gefur

$$x''(t) = \frac{d}{dt} f(t, x(t)) = f_t' + f_x' x'(t) = f_t' + f\, f_x'.$$

Taylor-liðun lausnarinnar er

$$\begin{aligned} x(t+h) &= x + hx'(t) + \frac{1}{2}h^2 x''(t) + O(h^3) \\ &= x + hf + \frac{1}{2}h^2(f_t' + ff_x') + O(h^3) \\ &= x + \frac{1}{2}hf + \frac{1}{2}h(f + hf_t' + (hf)f_x') + O(h^3) \end{aligned}$$

Nú sjáum við að síðasti liðurinn er 1. stigs Taylor liðun f með miðju (t, x) skoðuð í punktinum $(t+h, x+hf)$, því

$$f(t+h, x+hf) = f + hf_t' + (hf)f_x' + O(h^2)$$

og þar með er

$$x(t+h) = x(t) + \frac{1}{2}hf(t, x) + \frac{1}{2}hf(t+h, x+hf) + O(h^3).$$

Þessi formúla liggur til grundvallar 2. stigs Runge-Kutta-aðferð: Með henni fáum við nálgunarrunu $w_0, w_1, w_2, \dots$ þannig að $w_0 = x(0)$ og

$$w_j = w_{j-1} + \tfrac{1}{2}(F_1 + F_2), \quad j = 1, 2, \dots$$

þar sem

$$F_1 = h_j f(t_{j-1}, w_{j-1}), \quad \text{og} \quad F_2 = h_j f(t_j, w_{j-1} + F_1)$$

og eins og alltaf er $w_j \approx x(t_j)$.

6.2.8 Klassíska (fjórða stigs) Runge-Kutta aðferðin

Algengasta Runge-Kutta aðferðin er klassíska Runge-Kutta aðferðin. Þetta er fjórða stigs aðferð, sem þýðir að staðarskekkjan er $O(h^5)$ og heildarskekkjan er $O(h^4)$, .

$$w_j = w_{j-1} + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

þar sem

$$\begin{aligned} k_1 &= hf(t_{j-1}, w_{j-1}) \\ k_2 &= hf\left(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{k_1}{2}\right) \\ k_3 &= hf\left(t_{j-1} + \frac{h}{2}, w_{j-1} + \frac{k_2}{2}\right) \\ k_4 &= hf(t_{j-1} + h, w_{j-1} + k_3). \end{aligned}$$

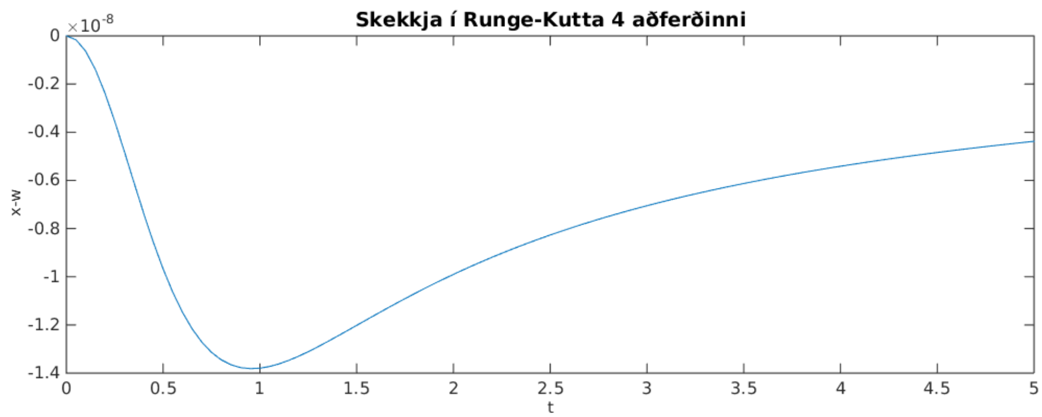
Ef $f(t, x)$ er bara fall af t, þ.e. óháð x, þá svarar þetta til þess að meta heildið φ með Simpson-reglunni.

6.2.9 Klassíska Runge-Kutta aðferðin: Dæmi

Skoðum nú sama dæmi og þegar við prófuðum aðferð Eulers.

Þá gefa eftirfarandi skipanir mynd af skekkjunni.

```
>> f = @(t,x) t./x;
>> [w,t]=rk4(f,0,1,5,100);
>> plot(t,sqrt(t.^2+1) - w)
```



Þetta er töluvert betra en aðferð Eulers sem skilaði skekkju af stærðargráðunni 10^{-2}.

6.3 Skekkjumat, samleitni og stöðugleiki

```
+++Mr. Jelly! Mr. Jelly!+++
+++Error At Address: 14, Treacle Mine Road, Ankh-Morpork+++
+++MELON MELON MELON+++
+++Divide By Cucumber Error. Please Reinstall Universe And Reboot +++
+++Whoops! Here Comes The Cheese! +++
+++Oneoneoneoneoneoneone+++
```

–villuskilaboð tölvunnar Hex í Interesting Times eftir Terry Pratchett

6.3.1 Skekkja

Fyrir eins skrefs aðferð skilgreinum við *staðarskekkju* við tímann t_n sem

$$\tau_n = \frac{x(t_n) - x(t_{n-1})}{h_n} - \varphi(f, t_{n-1}, t_n, x(t_{n-1}), x(t_n))$$

Hér er réttu lausninni stungið inn í nálgunarformúluna. Munum að hún uppfyllir

$$\frac{x(t_n) - x(t_{n-1})}{h_n} = \int_0^1 f(t_{n-1} + sh_n, x(t_{n-1} + sh_n))\, ds$$

Viljum geta metið τ_n sem fall af h_n, t.d.

$$\tau_n = O(h_n^k)$$

Almennt batna aðferðir eftir því sem veldisvísirinn k í staðarskekkjunni verður stærri.

Staðarskekkja er hlutfallsleg skekkja við að fara úr w_{n-1} yfir í w_n. Einnig má skoða uppsafnaða skekkju frá upphafstímanum t_0, hún er skilgreind með $e_n = x(t_j) - w_j$ og kallast *heildarskekkja*.

6.3.2 Staðarskekkja í aðferð Eulers

Aðferð Eulers er sett fram með formúlunni

$$w_n = w_{n-1} + h_n f(t_{n-1}, w_{n-1})$$

Staðarskekkjan er því

$$\begin{aligned}\tau_n &= \frac{x(t_n) - x(t_{n-1})}{h_n} - f(t_{n-1}, x(t_{n-1})) \\ &= \frac{x(t_n) - x(t_{n-1}) - x'(t_{n-1})h_n}{h_n} \\ &= \frac{\frac{1}{2}x''(\xi_n)h_{n-1}^2}{h_n} = \frac{1}{2}x''(\xi_n)h_{n-1} = O(h_n)\end{aligned}$$

Aðferð Eulers er því fyrsta stigs aðferð.

6.3.3 Stýring á staðarskekkju og breytileg skrefastærð

Hingað til þá höfum við ekki fengið neinar upplýsingar til að finna heppilegustu skrefastærð. Eftir því sem skrefastærðin er minni er staðarskekkjan sennilega minni, en þá komumst við hægar yfir og það er hætta á að heildarskekkjan hækki við að taka mörg skref. Í Aðferðir með breytilega skrefastærð munum við reyna að stilla skrefastærðina þannig að við tökum eins stór skref og mögulegt en þó þannig að staðarskekkjan sé ekki of há. Þá munum við þurfa eftirfarandi útleiðslu.

Hugsum okkur að við höfum tvær beinar nálgunaraðferðir

$$w_n = w_{n-1} + h_n\varphi(f, t_{n-1}, t_n, w_{n-1})$$

og

$$\tilde{w}_n = w_{n-1} + h_n\tilde{\varphi}(f, t_{n-1}, t_n, w_{n-1})$$

Skilgreinum tilsvarandi staðarskekkjur

$$\tau_n(h_n) = k_1 h_n^{\alpha_1} + o(h_i^{\alpha_1})$$

og

$$\tilde{\tau}_n(h_n) = k_2 h_n^{\alpha_2} + o(h_i^{\alpha_2}),$$

þar sem $\alpha_2 > \alpha_1$. Við tímann t_{n-1} hafa nálgunargildin $w_0, \dots, w_{n-1}$ hafi verið valin samkvæmt fyrri aðferðinni.

Meiningin að velja næsta tímapunkt t_n og þar með tímaskref h_n þannig að $\tau_n(h_n) \leq \delta$, en að $\tau_n(h_n)$ haldi sig sem næst δ, þar sem δ er gefið efra mark á staðarskekkjunni í fyrri aðferðinni.

Stærðin δ er kölluð *þolmörk* (e. tolerance) fyrir staðarskekkjuna og er oft táknuð með TOL.

Við byrjum á að setja $h = h_n$ inn í báðar aðferðirnar og bera útkomurnar saman

$$w_n = w_{n-1} + h\varphi(f, t_{n-1}, t_{n-1} + h, w_{n-1})$$

$$\tilde{w}_n = \tilde{w}_{n-1} + h\tilde{\varphi}(f, t_{n-1}, t_{n-1} + h, w_{n-1})$$

Við látum $\hat{w}_n$ tákna rétt gildi lausnarinnar á upphafsgildisverkefninu

- $x'(t) = f(t, x(t)),$

- $x(t_{n-1}) = w_{n-1}$,

í punktinum $t_{n-1} + h$.

Þá höfum við

$$\begin{aligned}\tau_n(h) &= \frac{\hat{w}_n - w_{n-1}}{h} - \varphi(f, t_{n-1}, t_{n-1} + h, w_{n-1}) \\ &= \frac{\hat{w}_n - w_{n-1} - h\varphi(f, t_{n-1}, t_{n-1} + h, w_{n-1})}{h} = \frac{\hat{w}_n - w_n}{h}\end{aligned}$$

og eins fæst

$$\begin{aligned}\tilde{\tau}_n(h) &= \frac{\hat{w}_n - w_{n-1}}{h} - \tilde{\varphi}(f, t_{n-1}, t_{n-1} + h, w_{n-1}) \\ &= \frac{\hat{w}_n - w_{n-1} - h\tilde{\varphi}(f, t_{n-1}, t_{n-1} + h, w_{n-1})}{h} = \frac{\hat{w}_n - \tilde{w}_n}{h}.\end{aligned}$$

Nú tökum við mismuninn og skilgreinum

$$\begin{aligned}\varepsilon &= \left|\frac{\tilde{w}_n - w_n}{h}\right| = |\tau_n(h) - \tilde{\tau}_n(h)| \\ &= |k_1|h^{\alpha_1} + o(h^{\alpha_1}) \approx |k_1|h^{\alpha_1}\end{aligned}$$

Munum að hér er skreflengdin $h = h_n$. Þessi nálgunarformúla gefur okkur möguleika á því að meta fastann

$$|k_1| \approx \frac{\varepsilon}{h_n^{\alpha_1}}.$$

6.3.4 Mat á skrefastærð

Segjum nú að við viljum halda staðarskekkjunni innan markanna $\delta/2$ og hafa skreflengdina í næsta skrefi $h_n = qh_{n-1}$, þá höfum við nálgunarjöfnuna

$$|\tau_n(qh_{n-1})| \approx |k_1|(qh_{n-1})^{\alpha_1} = \varepsilon q^{\alpha_1} \approx \frac{\delta}{2}.$$

Við tökum

$$q = \left(\frac{\delta}{2\varepsilon}\right)^{1/\alpha_1}$$

veljum síðan skrefstærðina $h_n = qh_{n-1}$ og reiknum út næsta gildi

$$w_n = w_{n-1} + h_n\varphi(f, t_{n-1}, t_n, w_{n-1})$$

6.4 Aðferðir með breytilega skrefastærð

Dæmi um aðferðir sem notast við breytilega skrefastærð.

- Einfaldast væri að nota Heun aðferðina (annars stigs) til að meta skrefastærðina í Euler aðferðinni (fyrsta stigs).
- Algengasta aðferðin er Runge-Kutta-Fehlberg (RKF45) sem notar 5. stigs nálgun til þess að meta staðarskekkjuna í 4. stigs aðferð.
- Endurbót á RKF45 er Runge-Kutta-Verner (RKV56) sem notar 6. stigs aðferð til að meta skekkjuna í 5. stigs aðferð.
- Fleiri aðferðir: Bogacki–Shampine (3. og 2. stigs), Cash–Karp (5. og 4. stigs) og Dormand–Prince (5. og 4. stigs).

6.4.1 Reiknirit fyrir Runge-Kutta-Fehlberg (RKF45)

$$\begin{aligned}\tilde{w}_j &= w_{j-1}\frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6 \\ w_j &= w_{j-1} + \frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5\end{aligned}$$

þar sem

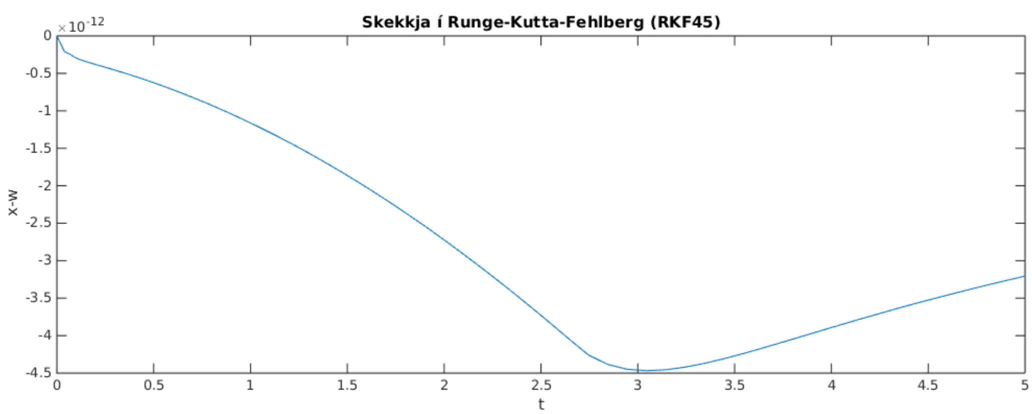
$$\begin{aligned}k_1 &= hf(t_{j-1}, w_{j-1}) \\ k_2 &= hf\left(t_{j-1} + \frac{1}{4}h, w_{j-1} + \frac{1}{4}k_1\right) \\ k_3 &= hf\left(t_{j-1} + \frac{3}{8}h, w_{j-1} + \frac{3}{32}k_1 + \frac{9}{32}k_2\right) \\ k_4 &= hf\left(t_{j-1} + \frac{12}{13}h, w_{j-1} + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3\right) \\ k_5 &= hf\left(t_{j-1} + h, w_{j-1} + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3 - \frac{845}{4104}k_4\right) \\ k_6 &= hf\left(t_{j-1} + \frac{1}{2}h, w_{j-1} - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5\right)\end{aligned}$$

6.4.2 Runge-Kutte-Fehlberg (RKF45) prófuð

Höldum áfram með dæmi sem við beittum aðferð Eulers og klassísku Runge-Kutta hér á undan

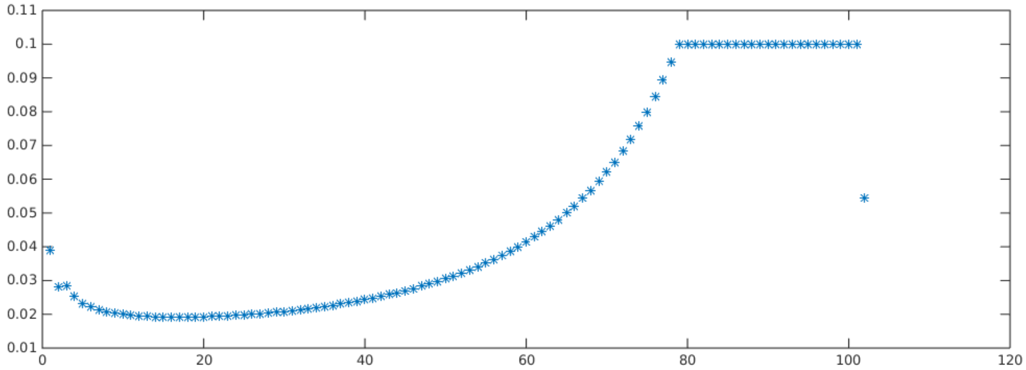
Þá gefur eftirfarandi mynd af skekkjunni. Hér er 0.01 minnsta leyfilega skrefastærðin, 0.1 stærsta leyfilega skrefastærðin og þolmörkin eru 10^{-10}.

```
>> f = @(t,x) t./x;
>> [w,t] = rkf45(f,0,1,5,[0.01,0.1,1E-10]);
>> plot(t,sqrt(t.^2+1) - w)
```



Hér á undan þá notðum við þolmörkin 10^{-10} sem skilaði okkur 103 misstórum tímagildum á bilinu $[0, 5]$. Svona getum við teiknað upp stærðina á tímaskrefunum.

```
>> plot(t(2:end)-t(1:end-1),'*')
```



6.5 Fjölskrefaaðferðir

Þær aðferðir sem við höfum séð eiga allar sameiginlegt að ákvarða nálgunargildi w_n aðeins út frá gildinu w_{n-1} næst á undan. Hægt er að nota fleiri gildi w_{n-1}, w_{n-2}, $\ldots$ og fá þannig betri nákvæmni, en aðferðirnar verða að sama skapi flóknari í notkun.

Eins og alltaf höfum við verkefnið

$$\begin{cases} x'(t) = f(t, x(t)) \\ x(t_0) = w_0 \end{cases}$$

og viljum nálga gildi lausnarinnar x á bili $[a, b]$ þar sem $a = t_0$ eða $b = t_0$. Látum t_0, t_1, …, t_n vera skiptingu á bilinu $[a, b]$ og gerum til einföldunar ráð fyrir að hún hafi jafna billengd $h = t_j - t_{j-1}$ fyrir $j = 1, \ldots, n$.

6.5.1 k-skrefa Adams-Bashforth aðferð

Við vitum að lausnin x uppfyllir

$$x(t_n) - x(t_{n-1}) = \int_{t_{n-1}}^{t_n} f(t, x(t))\,dt$$

Skrifum nú

$$f(t, x(t)) = P_{k-1}(t) + R_{k-1}(t)$$

þar sem

$$P_{k-1}(t) = \sum_{j=1}^{k} f(t_{n-j}, x(t_{n-j})) \cdot \ell_{k-1,j}(t)$$

er brúunarmargliðan gegnum punktana $(t_{n-k}, x(t_{n-k}))$, $(t_{n+1-k}, x(t_{n+1-k}))$, …, $(t_{n-1}, x(t_{n-1}))$, þ.e. gegnum síðustu k punkta á undan $(t_n, x(t_n))$.

Þetta eru k punktar og því er aðferðin kölluð k-skrefa aðferð.

Munum að til er ξ þannig að

$$R_{k-1}(t) = \frac{f^{(k)}(\xi, x(\xi))}{k!} \prod_{j=1}^{m} (t - t_{n-j}).$$

Við nálgum nú heildið af f yfir bilið $[t_{n-1}, t_n]$ með heildi P_{k-1} og fáum

$$w_{i+1} = w_i + \int_{t_i}^{t_{i+1}} P_{k-1}(t)\,dt$$

og með beinum útreikningum má sjá að skekkjan í þessari nálgun er $O(h^{k+1})$. Þessir útreikninga flækjast auðvitað eftir því sem k stækkar.

Augljóslega getum við ekki notað k skrefa Adams-Bashforth aðferðir um leið og við sjáum upphafsgildisverkefni, því við þurfum k ágiskunargildi $w_0, w_1, \ldots, w_{k-1}$ til að byrja að nota aðferðina. Þessi gildi má fá með hverri sem er af aðferðunum sem við höfum séð hingað til.

Ákveðin sértilfelli Adams-Bashforth aðferðanna eru meira notuð en önnur, það eru tveggja, þriggja og fjögurra skrefa aðferðirnar. Áhugasömum verður ekki skotaskuld úr að leiða út formúlurnar fyrir þær, en við birtum bara niðurstöðurnar.

Til styttingar skilgreinum við $f_j = f(t_j, w_j)$.

6.5.2 Tveggja skrefa Adams-Bashforth-aðferð

Þegar gildin w_{n-1} og w_{n-2} hafa verið fundin fæst næsta nálgunargildi með

$$w_n = w_{n-1} + h\big(\tfrac{3}{2} f_{n-1} - \tfrac{1}{2} f_{n-2}\big)$$

og skekkjan í nálguninni er $O(h^3)$.

6.5.3 Forrit fyrir tveggja skrefa Adams-Bashforth-aðferð

Aðferðin er útfærð í forritinu hér að neðan; það skýrir sig að mestu sjálft en við skulum taka eftir þrennu:

(i) Við krefjumst þess að notandinn gefi nálgunargildi á x(t(2)), þetta gerum við því til eru margar mismunandi aðferðir til að fá slíkt gildi og þær henta mis vel hverju sinni.

(ii) Við gerum ekki sérstaklega ráð fyrir að jafnt bil sé á milli stakanna í vigrinum t þó við höfum gert það hingað til. Það var aðeins gert til að einfalda útreikninga; aðferðin virkar nákvæmlega eins ef það er ekki jafnt bil á milli stakanna, svo sjálfsagt er að forrita hana þannig.

(iii) Við lágmörkum fjölda skipta sem við reiknum gildi f með að geyma alltaf gildið frá síðustu ítrun og nota það aftur, þetta getur sparað nokkurn tíma í útreikningum ef f er flókið fall.

```
function w = adams_bashforth_2(f,t,x1,x2)
%   w = adams{_}bashforth{_}2(f,t,x1,x2)
% Nálgar lausn upphafsgildisverkefnisins
%   x' = f(t,x)
%   x(t(1)) = x1
% í punktunum í t með 2ja þrepa Adams-Bashforth aðferð.
% Stakið x2 er nálgunargildi á x(t(2)).

N = length(t);  M = length(x1); w = zeros(M,N);
% Upphafsstillum gildi f(t,x) og w
fx1 = f(t(1),x1); fx2 = f(t(2),x2);
w(:,1) = x1; w(:,2) = x2;
for i=3:N
  % Reiknum nálgunargildi
  h = t(i)-t(i-1);
  w(:,i) = w(:,i-1) + (h/2)*(3*fx2 - fx1);
  fx1 = fx2; fx2 = f(t(i),w(:,i));
end
```

6.5.4 Þriggja skrefa Adams-Bashforth

Gefin w_{n-1}, w_{n-2} og w_{n-3} fæst næsta nálgunargildi með

$$w_n = w_{n-1} + h(\tfrac{23}{12} f_{n-1} - \tfrac{16}{12} f_{n-2} + \tfrac{5}{12} f_{n-2})$$

og staðarskekkjan er $O(h^4)$

6.5.5 Fjögurra skrefa Adams-Bashforth

Þegar við þekkjum w_{n-1}, w_{n-2}, w_{n-3} og w_{n-4} reiknum við næsta gildi með

$$w_n = w_{n-1} + h\big(\tfrac{55}{24} f_{n-1} - \tfrac{59}{24} f_{n-2} + \tfrac{37}{24} f_{n-3} - \tfrac{9}{24} f_{n-4}\big)$$

og skekkjan í nálguninni er $O(h^5)$.

6.6 Greining á samleitni og stöðugleika

Lítum aftur á upphafsgildisverkefnið okkar

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = w_0. \end{cases}$$

Við hugsum okkur að nálgun sé fundin í tímapunktunum

$$a = t_0 < t_1 < t_2 < \cdots < t_N = b.$$

Við táknum nálgunargildi á $x(t_j)$ með w_j. Það er gefið með

$$w_n = w_{n-1} + h_n \varphi(f, t_0, \ldots, t_n, w_0, \ldots, w_n)$$

þar sem fallið $\varphi(f, t_0, \ldots, t_n, w_0, \ldots, w_n)$ er skilgreint með einhverjum hætti.

Við köllum þetta *nálgunaraðferðina sem fallið φ gefur af sér.*

6.6.1 Skekkja

Skekkja (e. error) eða *heildarskekkja* (e. total error) í nálgun á $x(t_n)$ með w_n er

$$e_n = x(t_n) - w_n,$$

og *staðarskekkja* (e. local truncation error) nálgunaraðferðarinnar við tímann t_n er

$$\tau_n = \frac{x(t_n) - x(t_{n-1})}{h_n} - \varphi(f, t_0, \ldots, t_n, x(t_0), \ldots, x(t_n))$$

Athugasemd: Munið að hér er *rétta lausnin* sett inn í nálgunaraðferðina.

6.6.2 Samleitni

Hugsum okkur nú að fjöldi tímapunktanna N stefni á óendanlegt. Við segjum að nálgunaraðferðin φ sé *samleitin* ef

$$\lim_{N\to\infty} \max_{1\le n\le N} |e_n| = 0$$

þar sem $e_n = x(t_n) - w_n$ táknar skekkjuna í n-ta tímaskrefinu.

6.6.3 Samræmi

Við segjum að nálgunaraðferðin φ *samræmist* upphafsgildisverkefninu ef um sérhvern tímapunkt t_{n-1} gildir að

$$\lim_{h_n\to 0} \tau_n$$
$$= \lim_{t_n\to t_{n-1}} \left(\frac{x(t_n) - x(t_{n-1})}{t_n - t_{n-1}} - \varphi(f, t_0, \dots, t_n, x(t_0), \dots, x(t_n)) \right) = 0$$

6.6.4 Samræmi endurbættu Euler-aðferðarinnar

Munum að endurbætta Euler-aðferðin er

$$w_n = w_{n-1} + h_n f(t_{n-1} + \tfrac{1}{2}h_n, w_{n-1} + \tfrac{1}{2}hf(t_{n-1}, w_{n-1}))$$

sem gefur staðarskekkjuna

$$\tau_n = \frac{x(t_{n-1} + h_n) - x(t_{n-1})}{h_n}$$
$$-f(t_{n-1} + \tfrac{1}{2}h_n, x(t_{n-1}) + \tfrac{1}{2}h_n f(t_{n-1}, x(t_{n-1}))).$$

Nú hugsum við okkur að t_{n-1} sé haldið föstu og látum billengdina $h_n = t_n - t_{n-1}$ stefna á 0. Þá fæst

$$\lim_{h_n\to 0} \tau_n = x'(t_{n-1}) - f(t_{n-1}, x(t_{n-1})) = 0$$

Þetta segir okkur að endurbætta Euler-aðferðin samræmist upphafsgildisverkefninu.

6.6.5 Samræmi beinna eins skrefs aðferða

Þessi röksemdafærla alhæfist á allar beinar eins skrefs aðferðir, því staðarskekkja þeirra er

$$\tau_n = \frac{x(t_{n-1} + h_n) - x(t_{n-1})}{h_n} - \varphi(f, t_{n-1}, t_{n-1} + h_n, x(t_{n-1}))$$

Nú er eðlilegt að gefa sér að φ sé samfellt fall og þá verður markgildið af staðarskekkjunni

$$x'(t_{n-1}) - \varphi(f, t_{n-1}, t_{n-1}, x(t_{n-1}))$$
$$= f(t_{n-1}, x(t_{n-1})) - \varphi(f, t_{n-1}, t_{n-1}, x(t_{n-1})).$$

Eins skrefs aðferðin sem fallið φ gefur af sér er því stöðug ef og aðeins ef

$$\varphi(f, t_{n-1}, t_{n-1}, x(t_{n-1})) = f(t_{n-1}, x(t_{n-1})).$$

6.6.6 Stöðugleiki

Gerum nú ráð fyrir að upphafsgildinu w_0 sé breytt í $\tilde{w}_0$ og að $\tilde{x}(t)$ uppfylli

$$\begin{cases} \tilde{x}'(t) = f(t, \tilde{x}(t)), \\ \tilde{x}(t_0) = \tilde{w}_0. \end{cases}$$

Lítum síðan á tilsvarandi nálgunarrunu

$$\tilde{w}_n = \tilde{w}_{n-1} + h_n\varphi(f, t_0, \dots, t_n, \tilde{w}_0, \dots, \tilde{w}_n).$$

Við segjum að nálgunaraðferðin sem φ gefur af sér sé *stöðug* ef til er fall $k(t) > 0$ þannig að

$$|\tilde{w}_n - w_n| \le k(t_n)|\tilde{w}_0 - w_0|, \qquad n = 1, 2, 3 \dots.$$

6.6.7 Lipschitz-samfelldni

Rifjum nú upp að við gerum ráð fyrir að fallið $f(t,x)$ sé skilgreint á svæði D sem inniheldur

$$\{(t,x) \in \mathbb{R}^2\,;\, a \le t \le b, x \in \mathbb{R}\}.$$

Við segjum að f sé *Lipschitz samfellt á D með tilliti til x* ef til er fasti C_f þannig að

$$|f(t,x) - f(t,y)| \le C_f|x-y|, \qquad x, y \in \mathbb{R}.$$

Hugsum okkur að $\varphi(f,s,t,x)$ sé fall sem gefur af sér beina eins skrefs nálgunaraðferð fyrir upphafsgildisverkefnið $x'(t) = f(t,x(t))$ með $x(t_0) = w_0$.

Við segjum að φ sé *Lipschitz-samfellt með tilliti til x* ef um sérhvert Lipschitz-samfellt fall f, tölur $s,t \in [a,b]$ og $x,y \in \mathbb{R}$ gildir að til er fasti L_φ þannig að

$$|\varphi(f,s,t,x) - \varphi(f,s,t,y)| \le L_\varphi|x-y|, \qquad x,y \in \mathbb{R}.$$

6.6.8 Setning um stöðugleika og samleitni

Gefum okkur jafna skiptingu á tímabilinu $[a,b]$, $t_n = a + nh$, þar sem $n = 0,1,2,\dots,N$ og $h = (b-a)/N$.

Ef fallið φ er Lipschitz-samfellt með tilliti til x með Lipschitz-fastann L_φ, þá gildir:

1. Eins skrefs aðferðin sem φ gefur af sér er stöðug,

$$|\tilde{w}_n - w_n| \le e^{L_\varphi(t_n-a)}|\tilde{w}_0 - w_0|, \qquad n = 1,2,3,\dots.$$

2. Ef til eru fastar c og p þannig að staðarskekkjan uppfyllir $|\tau_n| \le c\,h^p$, fyrir öll $n = 1,2,3,\dots$ og $h \in]0,h_0]$, þá er aðferðin samleitin og við höfum

$$|e_n| = |x(t_n) - w_n| \le \frac{ch^p}{L_\varphi}\left(e^{L_\varphi(t_n-a)} - 1\right).$$

Jaðargildisverkefni

The pen is mightier than the sword if the sword is very short, and the pen is very sharp. – Terry Pratchett

7.1 Inngangur

7.1.1 Jaðargildisverkefni

Við ætlum að finna nálgunarlausnir á verkefnum af gerðinni

$$\begin{aligned} y'' = f(x, y, y'), \qquad a \le x \le b, \\ \alpha_1 y(a) + \alpha_2 y'(a) = \alpha_3, \\ \beta_1 y(b) + \beta_2 y'(b) = \beta_3. \end{aligned}$$

Lausn á verkefninu er þá fall $y(x) : [a, b] \to \mathbb{R}$ sem er þannig að y uppfyllir

- afleiðujöfnuna $y''(x) = f(x, y(x), y'(x))$,
- jaðarskilyrðin $\alpha_1 y(a) + \alpha_2 y'(a) = \alpha_3$ í $x = a$ og
- jaðarskilyrðin $\beta_1 y(b) + \beta_2 y'(b) = \beta_3$ í $x = b$.

Afleiðujafnan er sögð vera línuleg ef f er á forminu

$$y'' = p(x)y' + q(x)y + r(x), \qquad x \in [a, b].$$

7.1.2 Þrjár tegundir jaðarskilyrða

Venjulega eru jaðarskilyrðin flokkuð í þrjá flokka.

(i)	Dirichlet-jaðarskilyrði: (Fallsjaðarskilyrði:)	$y(a) = \alpha, \quad y(b) = \beta$
(ii)	Neumann-jaðarskilyrði: (Afleiðujaðarskilyrði:) (Flæðisjaðarskilyrði:)	$y'(a) = \alpha, \quad y'(b) = \beta$
(iii)	Robin-jaðarskilyrði: (Blandað jaðarskilyrði:)	$\alpha_1 y(a) + \alpha_2 y'(a) = \alpha_3$ $\beta_1 y(b) + \beta_2 y'(b) = \beta_3$ $(\alpha_1, \alpha_2) \neq (0, 0)$

Athugasemd: Athugið að Robin jaðarskilyrði með $\alpha_2 = 0$ (eða $\beta_2 = 0$) er Dirichlet skilyrði með $\alpha = \alpha_3/\alpha_1$ (eða

$\beta = \beta_3/\beta_1$).

Athugið að Robin jaðarskilyrði með $\alpha_1 = 0$ (eða $\beta_1 = 0$) er Neumann skilyrði með $\alpha = \alpha_3/\alpha_2$ (eða $\beta = \beta_3/\beta_2$).

7.2 Dirichlet-jaðarskilyrði

7.2.1 Skiptipunktar

Gefum okkur jafna skiptingu á bilinu $[a, b]$, $x_j = a + hj$, $j = 0, \dots, N$ þar sem $h = (b - a)/N$. Þá er

$$a = x_0 < x_1 < x_2 < \cdots < x_{N-1} < x_N = b.$$

Við nefnum x_j *skiptipunkta* skiptingarinnar.

Punktarnir $a = x_0$ og $b = x_N$ nefnast *endapunktar* skiptingarinnar og x_j, með $j = 1, \dots, N-1$, nefnast *innri punktar* skiptingarinnar.

Við ætlum aðeins að nálga lausnir á línulegum jöfnum

$$y'' = p(x)y' + q(x)y + r(x), \qquad x \in [a, b],$$

Við munum reiknum út nálgun á réttu lausninni $y(x)$ í skiptipunktunum x_j. Rétta gildið í punktinum x_j táknum við með y_j og nálgunargildið með w_j,

$$y_j = y(x_j) \approx w_j.$$

Eins skrifum við

$$p_j = p(x_j), \qquad q_j = q(x_j), \qquad r_j = r(x_j).$$

Aðvörun: Ólíkt upphafsgildisverkefnunum í kaflanum á undan þá táknum við breytuna hér með x og fallgildið með y. Þetta er eðlilegur ritháttur hér því í jaðargildisverkefnum þá er y oftast fall af staðsetningu en ekki tíma, t.d. hiti í röri, sveigja burðarbita, o.s.fr.

7.2.2 Línulegar afleiðujöfnur

Nú leiðum við út nálgunarjöfnur, eina fyrir hvern innri skiptipunkt. Við byrjum á því að stinga punkti x_j inn í afleiðujöfnuna

$$\big\{y''(x) = p(x)y'(x) + q(x)y(x) + r(x)\big\}_{x=x_j}.$$

Næst skiptum við afleiðunum y'' og y' út fyrir miðsettan mismunakvóta fyrir aðra afleiðuna og miðsettan mismunakvóta fyrir fyrstu afleiðuna. Þá fæst

$$\frac{y_{j+1} - 2y_j + y_{j-1}}{h^2} + O(h^2) = p_j\frac{y_{j+1} - y_{j-1}}{2h} + q_j y_j + r_j + O(h^2).$$

Nú fellum við niður leifaliðina og setjum nálgunargildin í stað réttu gildanna

$$\frac{w_{j+1} - 2w_j + w_{j-1}}{h^2} = p_j\frac{w_{j+1} - w_{j-1}}{2h} + q_j w_j + r_j$$

Hér fáum við eina jöfnu fyrir sérhvern innri skiptipunkt $j = 1, \dots, N-1$.

7.2.3 Dirichlet-jaðarskilyrði

Við erum komin með $N-1$ nálgunarjöfnu til þess að finna $N+1$ nálgunargildi $w_0, \dots, w_N$ fyrir $y_0, \dots, y_N$.

Ef við erum að leysa línulegt jaðargildisverkefni með Dirichlet-jaðarskilyrðum,

$$\begin{aligned} y'' &= p(x)y' + q(x)y + r(x), \qquad a \le x \le b, \\ y(a) &= \alpha \quad \text{og} \quad y(b) = \beta, \end{aligned}$$

þá fæst nálgunin með því að leysa línulega jöfnuhneppið

$$\begin{aligned} w_0 &= \alpha, \\ \frac{w_{j+1} - 2w_j + w_{j-1}}{h^2} &= p_j \frac{w_{j+1} - w_{j-1}}{2h} + q_j w_j + r_j, \qquad j = 1, \dots, N-1, \\ w_N &= \beta. \end{aligned}$$

7.2.4 Jafngild framsetning á hneppinu

Við lítum aftur á línulegu nálgunarjöfnurnar

$$\frac{w_{j+1} - 2w_j + w_{j-1}}{h^2} = p_j \frac{w_{j+1} - w_{j-1}}{2h} + q_j w_j + r_j.$$

Margföldum alla liði með $-h^2$ og röðum síðan óþekktu stærðunum vinstra mengin jafnaðarmerkisins. Þá fæst línulega jöfnuhneppið

$$\big(-1 - \tfrac{1}{2}hp_j\big)w_{j-1} + \big(2 + h^2 q_j\big)w_j + \big(-1 + \tfrac{1}{2}hp_j\big)w_{j+1} = -h^2\, r_j$$

fyrir $j = 1, 2, 3, \dots, N-1$.

7.2.5 Línulega jöfnuhneppið á fylkjaformi

Nú er hægt að skrifa jöfnuhneppið á fylkjaformi

$$A\mathbf{w} = \mathbf{b}$$

Hér er

$$A = \begin{bmatrix} 1 & 0 & & & & & & \\ l_1 & d_1 & u_1 & & & & & \\ & l_2 & d_2 & u_2 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & l_{N-2} & d_{N-2} & u_{N-2} & \\ & & & & & l_{N-1} & d_{N-1} & u_{N-1} \\ & & & & & & 0 & 1 \end{bmatrix}$$

þar sem stuðlarnir l_j, d_j og u_j eru gefnir með

$$\begin{aligned} l_j &= -1 - \tfrac{1}{2}hp_j \\ d_j &= 2 + h^2 q_j \\ u_j &= -1 + \tfrac{1}{2}hp_j \end{aligned}$$

og vigrarnir eru

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \cdot \\ \cdot \\ \cdot \\ w_{N-2} \\ w_{N-1} \\ w_N \end{bmatrix} \qquad \text{og} \qquad \mathbf{b} = \begin{bmatrix} \alpha \\ -h^2 r_1 \\ -h^2 r_2 \\ \cdot \\ \cdot \\ \cdot \\ -h^2 r_{N-2} \\ -h^2 r_{N-1} \\ \beta \end{bmatrix}$$

Við þekkjum allar tölurnar í A og $\mathbf{b}$, þannig að við getum leyst jöfnuhneppið og með því fundið nálgunargildin $\mathbf{w}$.

7.3 Neumann og Robin -jaðarskilyrði

7.3.1 Felugildi

Við skulum gera ráð fyrir að rétta lausnin $y(x)$ uppfylli blandað jaðarskilyrði í $x = a$,

$$\alpha_1 y(a) + \alpha_2 y'(a) = \alpha_3.$$

Til þess að líkja eftir afleiðujöfnunni í punktinum $x = a$ þá hugsum við okkur að við bætum við einum skiptipunkti $x_{-1} = a - h$ og látum w_f tákna ímyndað gildi lausnarinnar í x_{-1}.

Svona punktur x_{-1} utan við skiptinguna er kallaður *felupunktur* við skiptinguna og ímyndað gildi w_f í felupunkti er kallað *felugildi*.

Takið eftir því að lausnin er ekki til í felupunktinum, en við reiknum eins og w_f sé gildi hennar þar.

Mismunajafnan sem líkir eftir afleiðujöfnunni í punktinum x_0 er

$$\big(-1 - \tfrac{1}{2}hp_0\big)w_f + \big(2 + h^2 q_0\big)w_0 + \big(-1 + \tfrac{1}{2}hp_0\big)w_1 = -h^2 r_0$$

Mismunajafnan sem líkir eftir jaðarskilyrðinu er

$$\alpha_1 w_0 + \alpha_2 \frac{w_1 - w_f}{2h} = \alpha_3.$$

7.3.2 Jafna fyrir felugildið

Jafnan sem líkir eftir jaðarskilyrðinu er:

$$\alpha_1 w_0 + \alpha_2 \frac{w_1 - w_f}{2h} = \alpha_3.$$

Út úr henni leysum við

$$w_f = w_1 - \frac{2h}{\alpha_2}\big(\alpha_3 - \alpha_1 w_0\big)$$

Við stingum síðan þessu gildi inn í jöfnuna sem líkir eftir afleiðujöfnunni

$$\big(-1 - \tfrac{1}{2}hp_0\big)w_f + \big(2 + h^2 q_0\big)w_0 + \big(-1 + \tfrac{1}{2}hp_0\big)w_1 = -h^2 r_0$$

Út fæst fyrsta jafna hneppisins

$$\left(2 + h^2 q_0 - \big(2 + hp_0\big)h\frac{\alpha_1}{\alpha_2}\right)w_0 - 2w_1 = -h^2 r_0 - \big(2 + hp_0\big)h\frac{\alpha_3}{\alpha_2}.$$

Með því að innleiða felupunkt $x_{N+1} = b + h$ hægra megin við skiptinguna, tilsvarandi felugildi w_f og leysa saman tvær jöfnur þá fáum við síðustu jöfnu hneppisins

$$-2w_{N-1} + \left(2 + h^2 q_N + \big(2 - hp_N\big)h\frac{\beta_1}{\beta_2}\right)w_N = -h^2 r_N - \big(2 - hp_N\big)h\frac{\beta_3}{\beta_2}$$

Við erum því aftur komin með $(N+1) \times (N+1)$-jöfnuhneppi.

7.3.3 Hneppið á fylkjaformi

$$A\mathbf{w} = \mathbf{b}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & & & & & & & \\ l_1 & d_1 & u_1 & & & & & & \\ & l_2 & d_2 & u_2 & & & & & \\ & & \cdot & \cdot & \cdot & & & & \\ & & & \cdot & \cdot & \cdot & & & \\ & & & & \cdot & \cdot & \cdot & & \\ & & & & & l_{N-2} & d_{N-2} & u_{N-2} & \\ & & & & & & l_{N-1} & d_{N-1} & u_{N-1} \\ & & & & & & & a_{N+1,N} & a_{N+1,N+1} \end{bmatrix}$$

Þar sem stuðlarnir l_j, d_j og u_j fyrir $j = 1, 2, 3 \ldots, N-1$ eru þeir sömu og áður.

$$\begin{aligned} l_j &= -1 - \tfrac{1}{2}hp_j \\ d_j &= 2 + h^2 q_j \\ u_j &= -1 + \tfrac{1}{2}hp_j \end{aligned}$$

7.3.4 Fyrsta og síðasta lína hneppisins

$$a_{11} = \begin{cases} 1, & \text{Dirichlet í } x = a : \alpha_1 \neq 0, \alpha_2 = 0, \\ d_0 & \text{Neumann í } x = a : \alpha_1 = 0, \alpha_2 \neq 0, \\ d_0 + 2hl_0\alpha_1/\alpha_2 & \text{Robin í } x = a : \alpha_2 \neq 0. \end{cases}$$

$$a_{12} = \begin{cases} 0, & \text{Dirichlet í } x = a : \alpha_1 \neq 0, \alpha_2 = 0, \\ -2, & \text{annars.} \end{cases}$$

$$a_{N+1,N+1} = \begin{cases} 1, & \text{Dirichlet í } x = b : \beta_1 \neq 0, \beta_2 = 0, \\ d_N & \text{Neumann í } x = b : \beta_1 = 0, \beta_2 \neq 0, \\ d_N - 2hu_N\beta_1/\beta_2 & \text{Robin í } x = a : \beta_2 \neq 0. \end{cases}$$

$$a_{N+1,N} = \begin{cases} 0, & \text{Dirichlet í } x = b : \beta_1 \neq 0, \beta_2 = 0, \\ -2 & \text{annars.} \end{cases}$$

7.3.5 Hægri hlið hneppisins

$$\mathbf{b} = \begin{bmatrix} b_1 \\ -h^2 r_1 \\ -h^2 r_2 \\ \cdot \\ \cdot \\ \cdot \\ -h^2 r_{N-2} \\ -h^2 r_{N-1} \\ b_{N+1} \end{bmatrix}$$

$$b_1 = \begin{cases} \alpha = \alpha_3/\alpha_1, & \text{Dirichlet í } x = a : \alpha_1 \neq 0, \alpha_2 = 0, \\ -h^2 r_0 + 2hl_0\alpha_3/\alpha_2 & \text{Neumann í } x = a : \alpha_1 = 0, \alpha_2 \neq 0, \\ -h^2 r_0 + 2hl_0\alpha_3/\alpha_2 & \text{Robin í } x = a : \alpha_2 \neq 0. \end{cases}$$

$$b_{N+1} = \begin{cases} \beta = \beta_3/\beta_1, & \text{Dirichlet í } x = a : \beta_1 \neq 0, \beta_2 = 0, \\ -h^2 r_N - 2hu_N\beta_3/\beta_2 & \text{Neumann í } x = a : \beta_1 = 0, \beta_2 \neq 0, \\ -h^2 r_N - 2hu_N\beta_3/\beta_2 & \text{Robin í } x = a : \beta_2 \neq 0. \end{cases}$$

7.3.6 Samantekt

Gildi lausnarinnar $y(x)$ á línulega jaðargildisverkefninu

$$\begin{gathered} y'' = p(x)y' + q(x)y + r(x), \qquad a \leq x \leq b, \\ \alpha_1 y(a) + \alpha_2 y'(a) = \alpha_3, \\ \beta_1 y(b) + \beta_2 y'(b) = \beta_3 \end{gathered}$$

í punktunum $x_j = a + jh$, þar sem $h = (b - a)/N$ og $j = 0, \ldots, N$, eru nálguð með

$$w_j \approx y(x_j) = y_j$$

Dálkvigurinn

$$\mathbf{w} = [w_0, w_1, \ldots, w_N]^T$$

er lausn á línulegu jöfnuhneppi $A\mathbf{w} = \mathbf{b}$.

Stuðlum $(N + 1) \times (N + 1)$ fylkisins A og $(N + 1)$-dálkvigursins $\mathbf{b}$ hefur verið lýst hér að framan.

Jöfnuhneppi

Why do you go away? So that you can come back. So that you can see the place you came from with new eyes and extra colors. And the people there see you differently, too. Coming back to where you started is not the same as never leaving. – Terry Pratchett

8.1 Línuleg jöfnuhneppi

8.1.1 Verkefnið

Verkefnið í þessum kafla er eftirfarandi. Gefið $n \times n$ fylki A og n-vigur $\mathbf{b}$ þá leitum við að vigri $\mathbf{x}$ þannig að

$$A\mathbf{x} = \mathbf{b}.$$

8.1.2 Lausnir

Við höfum almennt tvær leiðir til þess að leysa línuleg jöfnuhneppi:

- Gauss-eyðing og innsetning.
- Reikna andhverfu A, A^{-1}, en þá er

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

8.1.3 Fjöldi aðgerða

- Gauss-eyðing fyrir $n \times n$ fylki krefst um það bil $\frac{2}{3}n^3$ reikniaðgerða. Innsetningin krefst svo n^2 aðgerða til viðbótar. Samanlagður fjöldi aðgerða er því

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n.$$

- Það að reikna A^{-1} krefst hins vegar $2n^3 - 2n^2 + n$ aðgerða og margföldunin $A^{-1}b$, krefst $2n^2 - n$ aðgerða til viðbótar. Samanlagður fjöldi aðgerða er því

$$2n^3.$$

Hér er greinilega gáfulegra að nota Gauss-eyðingu. Auk þess er ekki skynsamlegt að ætla að reikna andhverfuna ef ákveða fylkisins er 0 eða nálægt 0. Almennt þá forðumst við eins og mögulegt er að reikna A^{-1}.

8.1.4 Vandamál með stöðugleika

Skoðum jöfnuhneppið

$$\begin{bmatrix} \varepsilon & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Nákvæm lausn er $x_1 = 1 + \frac{\varepsilon}{1-\varepsilon}, x_2 = 1 - \frac{\varepsilon}{1-\varepsilon}$.

Ef hins vegar ε er minna en nákvæmnin í tölvunni sem erum að vinna á, þá gefur Gauss-eyðing í tölvu, þar sem eytt er með 1. línu, svarið $x_1 = 0, x_2 = 1$.

Ef línunum væri víxlað, þá gæfi tölvan hins vegar $x_1 = 1, x_2 = 1$ sem er miklu nær réttu svari.

Aðvörun: Það er alveg ljóst að megum ekki framkvæma Gauss-eyðingu blindandi því þá geta magnast upp styttingarskekkjur sem skemma lausnina okkar.

8.2 Vending (e. pivoting)

8.2.1 Inngangur

Það sem olli vandræðum í dæminu hér á undan var það að forystustuðull fyrstu línunnar var hlutfallslega miklu minni en forystustuðull annarrar línu.

Lausnin felst í því að víxla á línum þannig að við þurfum ekki að notast við litla forystustuðla.

8.2.2 Hlutvending (e. partial pivoting)

Í grófum dráttum: Í umferð i í Gauss-eyðingunni þá athugum við hvort tölugildi forystustuðla línanna fyrir neðan línu i eru stærri en forystustuðull línu i, ef svo er þá víxlum við á þeirri línu og línu i.

Það er, í i-tu ítrun Gauss-eyðingar þá látum við $M_i = \max_{i \leq j \leq n} |a_{ji}|$. Ef $|a_{ii}| < M_i$ þá víxlum við á línu i og fyrstu línunni fyrir neðan sem hefur forystustuðul með tölugildi jafnt og M_i. Það er, ef $j_0 = \min\{j; i \leq j \leq n \text{ og } a_{ji} = M_i\}$ þá víxlum við á línu i og j_0.

8.2.3 Vankantar

Hlutvending virkar oft vel en getur búið til skekkju þar sem hún tekur bara tillit til forystustuðlanna í hverri línu.

Ef jöfnuhneppið

$$\begin{aligned} 0.7x_1 + 1725x_2 &= 1739 \\ 0.4352x_1 - 5.433x_2 &= 3.271 \end{aligned}$$

er leyst með fjögra stafa nákvæmni og hlutvendingu þá fæst $x_2 = 1.001$ og $x_1 = 17.14$ en rétt svar er $x_2 = 1$ og $x_1 = 20$. Það er því ljóst að hlutvending tekst ekki alltaf til að halda skekkju í skefjum.

8.2.4 Sköluð hlutvending (e. scaled partial pivoting)

Ein leið til að koma í veg fyrir þessi vandræði er sköluð hlutvending.

Skilgreinum vigurinn s sem heldur utan um „stærð” línanna í A,

$$s_i = \max_{1\leq j\leq n} |a_{ij}|.$$

Látum dálkvigurinn $\mathbf{r}$ halda utan um það hvernig við umröðum línunum í A. Byrjum með

$$\mathbf{r} = [1\ 2\ 3\ 4\ \dots\ n]^T.$$

8.2.5 Athugasemd

Við munum uppfæra $\mathbf{r}$ eftir þörfum en breytum ekki s (of dýrt ef n er stórt).

Í ítrun i þá látum við

$$M_i = \max_{i\leq j\leq n} \frac{|a_{r_j i}|}{s_{r_j}},$$

og látum j_0 vera minnsta j þannig að hámarkinu er náð,

$$\frac{|a_{r_{j_0} i}|}{s_{r_{j_0}}} = M_i.$$

Ef $i < j_0$ þá skiptum við á línum i og j_0, þ.e.

$$\mathbf{r} = [\dots\ i\ \dots\ j_0\ \dots]^T \text{ breytist í } \mathbf{r} = [\dots\ j_0\ \dots\ i\ \dots]^T.$$

Þannig að í ítrun i þá segir i-ta stakið í $\mathbf{r}$ okkur hvaða línu við eigum að eyða með.

Aðvörun: Athugið að hér umröðum við ekki línunum í fylkinu því það er miklu ódýrara að nota vigurinn $\mathbf{r}$ til þess að halda utanum í hvaða röð á að framkvæma Gauss-eyðinguna.

8.3 Fylkjastaðall

8.3.1 Að mæla fjarlægð milli hluta

Á rauntalnalínunni þá mælum við fjarlægð með tölugildinu, þannig að fjarlægðin á milli x og y er gefin með $d(x, y) = |x - y|$.

Í $\mathbb{R}^n$ þá finnst okkur evklíðski staðallinn náttúrulegur, enda svarar hann til þess að mæla fjarlægð milli punkta með reglustiku;

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

Þetta er hins vegar ekki eina leiðin til þess mæla fjarlægð í $\mathbb{R}^n$ og, eins og við sjáum fljótlega, ekki endilega réttari en aðrar aðferðir.

Almennt viljum við geta mælt „fjarlægð” á milli allra þeirra hluta sem við erum skoða, hvort sem það eru margliður, föll eða fylki. Tilgangurinn er að geta metið hversu langt nálgunin okkar er frá réttu gildi og hversu stór skekkjan er í samanburði við stærð hlutarins sem við erum að vinna með.

8.3.2 Vigurstaðall

Fall $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ kallast *vigurstaðall* (e. vector norm) ef fyrir öll $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ og $\alpha \in \mathbb{R}$ gildir eftirfarandi:

1. $\|\mathbf{x}\| \geq 0$
2. $\|\mathbf{x}\| = 0$ ef og aðeins ef $\mathbf{x} = 0$
3. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Athugasemd: Tölugildisfallið á $\mathbb{R}$ er greinilega staðall.

8.3.3 Dæmi um staðla

- ℓ_2 staðallinn: Einnig kallaður evklíðska fjarlægðin, er gefinn með

$$\|\mathbf{x}\|_2 = \left(\sum_{j=1}^{n} x_j^2\right)^{\frac{1}{2}} = \sqrt{\mathbf{x}\cdot\mathbf{x}}.$$

- ℓ_∞ staðallinn:

$$\|\mathbf{x}\|_\infty = \max_{1\leq j\leq n} |x_j|.$$

- ℓ_p staðlar: Almennt ef $1 \leq p < \infty$ þá skilgreinum við

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{\frac{1}{p}}.$$

8.3.4 Fylkjastaðall

Fylkjastaðall (e. matrix norm) er fall $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$, þannig að fyrir öll $A, B \in \mathbb{R}^{n\times n}$ og $\alpha \in \mathbb{R}$ gildir

1. $\|A\| \geq 0$
2. $\|A\| = 0$ ef og aðeins ef $A = 0$
3. $\|\alpha A\| = |\alpha|\|A\|$
4. $\|A + B\| \leq \|A\| + \|B\|$
5. $\|AB\| \leq \|A\|\|B\|$

Athugasemd: Ef þessi skilgreining er borin saman við skilgreininguna á *vigurstaðli* fyrir vigurrúm þá sjáum við að eini raunverulegi munurinn er skilyrði 5.

8.3.5 Fylkjastaðall skilgreindur út frá vigurstaðli

Látum $\|\cdot\|$ vera vigurstaðal. Fallið $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ sem skilgreint er með

$$\|A\| = \max_{\|\mathbf{x}\|\neq 0} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|},$$

kallast *náttúrulegi fylkjastaðallinn* sem $\|\cdot\|$ gefur af sér.

Athugasemd: Það þarf að sýna, og er ekki mjög erfitt, að þessi fylkjastaðall uppfyllir öll skilyrðin í skilgreiningu hér á undan og er því sannarlega fylkjastaðall.

Athugasemd: Ef $\|\cdot\|$ er náttúrulegur fylkjastaðall þá gildir að fyrir öll fylki A og alla vigra $\mathbf{x}$ að

$$\underbrace{\|A\mathbf{x}\|}_{\text{vigurstaðll}} \leq \underbrace{\|A\|}_{\text{fylkjastaðall}} \cdot \underbrace{\|\mathbf{x}\|}_{\text{vigurstaðall}}$$

8.3.6 Dæmi um fylkjastaðal

Fyrir sérhvern ℓ_p staðal fáum við fylkjastaðal $\|\cdot\|_p$.

$\|\cdot\|_\infty$: Einfaldastur er staðallinn sem tilheyrir ℓ_∞, en hann uppfyllir

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|.$$

8.3.7 Eigingildi

Látum A vera fylki. Ef tala λ (hugsanlega tvinntala) og vigur $\mathbf{x}$ uppfylla

$$A\mathbf{x} = \lambda\mathbf{x},$$

þá kallast λ *eigingildi* A, og $\mathbf{x}$ *eiginvigur* A.

Athugið að eigingildi A eru nákvæmlega rætur kennimargliðu A, $t \mapsto \det(A - It)$.

8.3.8 Róf og rófgeisli

Mengi allra eigingilda A er kallað *róf* A (e. spectrum) og er táknað með $\sigma(A)$.

Rófgeisli (e. spectral radius) fylkisins A er talan

$$\rho(A) = \max_{\lambda\in\sigma(A)} |\lambda|.$$

8.3.9 Setning um rófgeisla

Látum A vera fylki, þá gildir eftirfarandi

- $\|A\|_2 = \sqrt{\rho(A^T A)}$
- $\rho(A) \leq \|A\|$ fyrir sérhvern náttúrulegan fylkjastaðal $\|\cdot\|$
- Fyrir sérhvert $\varepsilon > 0$ þá er til náttúrulegur fylkjastaðall $\|\cdot\|$ þannig að $\|A\| \leq \rho(A) + \varepsilon$.

8.4 Skekkjumat og ástandstala

8.4.1 Hvernig á að mæla skekkju

Gerum ráð fyrir að A sé andhverfanlegt fylki, $\mathbf{b}$ vigur og að við séum að leita að lausn $\mathbf{x}$ á

$$A\mathbf{x} = \mathbf{b}.$$

Ef við höfum nálgun $\tilde{\mathbf{x}}$ þannig að *leifin* (e. residual) $\mathbf{r} = A\tilde{\mathbf{x}} - \mathbf{b}$ er lítil, hvað getum við þá sagt um *skekkjuna* (e. error) $\mathbf{e} = \mathbf{x} - \tilde{\mathbf{x}}$? Er hún endilega lítil?

Sjáum að svo er ekki, skekkjan getur verið hlutfallslega miklu meiri heldur en leifin.

Skoðum jöfnuhneppið

$$\begin{bmatrix} 1 & 2 \\ 0.99 & 1.99 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Sjáum að rétt lausn er $[-1\ 1]^T$. Ef við skilgreinum $\tilde{x} = [1\ 0]^T$ þá er

$$\mathbf{e} = [-2\ 1]^T \qquad \|\mathbf{e}\|_\infty = 2.$$

Hins vegar þá er leifin

$$\mathbf{r} = A\tilde{\mathbf{x}} - \mathbf{b} = \begin{bmatrix} 1 & 2 \\ 0.99 & 1.99 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ -0.01 \end{bmatrix}$$

þannig að $\|\mathbf{e}\|_\infty = 0.01$. Hér er skekkjan 200 sinnum stærri en leifin sem sýnir að lítil leif þarf ekki endilega að hafa í för með sér litla skekkju.

8.4.2 Skekkjumat

Við höfum fjórar jöfnur

$$\mathbf{b} = A\mathbf{x}, \quad \mathbf{x} = A^{-1}\mathbf{b}, \quad \mathbf{r} = A(\tilde{\mathbf{x}} - \mathbf{x}) = A\mathbf{e}, \quad \text{og} \quad \mathbf{e} = A^{-1}\mathbf{r}$$

og þær gefa okkur fjórar ójöfnur fyrir tilsvarandi staðal:

$$\|\mathbf{b}\| \le \|A\|\|\mathbf{x}\|, \quad \|\mathbf{x}\| \le \|A^{-1}\|\|\mathbf{b}\|, \quad \|\mathbf{r}\| \le \|A\|\|\mathbf{e}\|, \quad \|\mathbf{e}\| \le \|A^{-1}\|\|\mathbf{r}\|$$

Við getum tengt tvær síðustu ójöfnurnar saman í mat á skekkjunni

$$\frac{1}{\|A\|} \cdot \|\mathbf{r}\| \le \|\mathbf{e}\| \le \|A^{-1}\|\|\mathbf{r}\|$$

og með því að nota fyrstu tvær ójöfnurnar fæst mat á hlutfallslegri skekkju

$$\frac{1}{\|A\|\|A^{-1}\|}\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \le \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \le \|A\|\|A^{-1}\|\frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

Nú skilgreinum við *ástandstölu fylkisins* A með

$$\kappa(A) = \|A\|\|A^{-1}\|.$$

8.4.3 Ástandstala fylkis og mat á hlutfallslegri skekkju

Með ástandstölunni verður mat okkar á hlutfallslegu skekkjunum að

$$\frac{1}{\kappa(A)} \cdot \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} \le \frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \le \kappa(A) \cdot \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|}$$

Þetta segir okkur að ef ástandstala fylkisins er stór er mögulegt að skekkjan verði miklu stærri en leifin, samanber dæmið hér á undan þar sem ástandstalan er 1197.

Aðvörun: Athugið að skilgreiningin

$$\kappa(A) = \|A\|\|A^{-1}\|.$$

er *mjög* háð því hvaða staðal við veljum. Þó höfum við að fyrir alla fylkjastaðla gildir

$$1 = \|I\| = \|AA^{-1}\| \le \|A\|\|A^{-1}\| = \kappa(A)$$

8.4.4 Áhrif gagnaskekkju

Hugsum okkur nú að við viljum leysa jöfnuhneppi $A\mathbf{x} = \mathbf{b}$, en vegna skekkju í stuðlum jöfnuhneppisins leysum við annað hneppi $\tilde{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$.

Við skilgreinum *gagnaskekkjur* $\delta A = \tilde{A} - A$ og $\delta\mathbf{b} = \tilde{\mathbf{b}} - \mathbf{b}$ og ætlum að nota þær til þess að meta skekkjuna $\mathbf{e} = \delta\mathbf{x} = \tilde{\mathbf{x}} - \mathbf{x}$.

Við verðum að gera ráð fyrir að $\|\delta A\| \le 1/\|A^{-1}\|$ sem tryggir að fylkið $\tilde{A}$ sé andhverfanlegt.

Nú stillum við upp jöfnuhneppinu $\tilde{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ á forminu

$$(A + \delta A)(\mathbf{x} + \delta\mathbf{x}) = (\mathbf{b} + \delta\mathbf{b})$$

sem jafngildir

$$\delta\mathbf{x} = A^{-1}\big(\delta\mathbf{b} - (\delta A)\mathbf{x} - (\delta A)(\delta x)\big)$$

Af þessari jöfnu leiðir ójafnan

$$\|\delta\mathbf{x}\| \le \|A^{-1}\|\big(\|\delta\mathbf{b}\| + \|\delta A\|\|\mathbf{x}\| + \|\delta A\|\|\delta\mathbf{x}\|\big)$$

Einangrum nú $\|\delta\mathbf{x}\|$,

$$\|\delta\mathbf{x}\| \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|} \cdot \big(\|\delta\mathbf{b}\| + \|\delta A\|\|\mathbf{x}\|\big)$$

deilum með $\|\mathbf{x}\|$ báðum megin, margföldum síðan með $\|A\|$ í teljara og nefnara í hægri hliðinni,

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{\|A\|\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|} \cdot \left(\frac{\|\delta\mathbf{b}\|}{\|A\|\|\mathbf{x}\|} + \frac{\|\delta A\|}{\|A\|}\right)$$

Samkvæmt skilgreiningu er $\kappa(A) = \|A\|\|A^{-1}\|$ og við höfum auk þess ójöfnuna $\|\mathbf{b}\| \le \|A\|\|\mathbf{x}\|$, en það gefur matið á hlutfallslegu skekkjunni sem við sækjumst eftir

$$\frac{\|\delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \frac{\kappa(A)}{1 - \kappa(A) \cdot (\|\delta A\|/\|A\|)} \cdot \left(\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|} + \frac{\|\delta A\|}{\|A\|}\right).$$

8.5 LU-þáttun

8.5.1 Nokkrar skilgreiningar

1. Fylkið A nefnist *neðra þríhyrningsfylki* ef öll stök fyrir ofan hornalínuna í A eru 0, þ.e. $a_{ij} = 0$ ef $i < j$.
2. Fylkið A nefnist *efra þríhyrningsfylki* ef öll stökin neðan við hornalínuna eru 0, þ.e. $a_{ij} = 0$ ef $i > j$.
3. Fylkið A nefnist *bandfylki* (e. striped matrix) ef til er $\beta \leq n-2$ þannig að $a_{ij} = 0$ ef $|i-j| > \beta$. Minnsta talan β sem uppfyllir þetta skilyrði kallst á *bandvídd* fylkisins A.
4. Ef A er bandfylki með bandvíddina 1, þá nefnist A *þríhornalínufylki*.
5. Fylkið A er sagt vera *samhverft* ef $a_{ij} = a_{ji}$ fyrir öll i og j.
6. Fylkið A er sagt vera *jákvætt ákvarðað* ef $\mathbf{x}^T A\mathbf{x} > 0$ gildir fyrir alla vigra $\mathbf{x} \neq 0$ í $\mathbb{R}^n$.

8.5.2 Úrlausn á jöfnuhneppi með neðra þríhyrningsfylki

Ef A er neðra þríhyrningsfylki, þá er úrlausn jöfnuhneppisins $A\mathbf{x} = \mathbf{b}$ auðveld, því hneppið er þá af gerðinni

$$\begin{aligned}
&a_{11}x_1 &&= b_1,\\
&a_{21}x_1 + a_{22}x_2 &&= b_2,\\
&a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &&= b_3,\\
&\vdots \qquad \vdots \qquad \vdots &&\vdots\\
&a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 \cdots + a_{nn}x_n &&= b_n,
\end{aligned}$$

Við getum rakið okkur niður línurnar og leyst út stærðirnar $x_1, \dots, x_n$ hverja á eftir annarri

$$\begin{aligned}
x_1 &= b_1/a_{11},\\
x_2 &= (b_2 - a_{21}x_1)/a_{22},\\
x_3 &= (b_3 - a_{31}x_1 - a_{32}x_2)/a_{33}\\
\vdots \quad & \quad \vdots \qquad \vdots\\
x_n &= (b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{n,n-1}x_{n-1})/a_{nn}.
\end{aligned}$$

8.5.3 Fjöldi aðgerða

Nú skulum við telja saman fjölda reikningsaðgerða sem þarf til þess að framkvæma þessa útreikninga.

Við lítum á samlagningu og frádrátt sem sömu aðgerðina. Við þurfum enga samlagningu til að reikna út x_1, eina til þess að reikna út x_2, tvær til þess að reikna x_3 og þannig áfram upp í $n-1$ samlagningu til þess að reikna út x_n.

Heildarfjöldinn er því

$$1 + 2 + \cdots + n - 1 = \tfrac{1}{2}n(n-1) = \tfrac{1}{2}n^2 - \tfrac{1}{2}n.$$

Fjöldi margfaldana er sá sami.

Við þurfum hins vegar aðeins eina deilingu til þess að reikna út hverja af stærðunum $x_1, \dots, x_n$.

Heildarfjöldi reikniaðgerða við úrlausn á línulegu jöfnuhneppi $Ax = b$, þar sem A er neðra þríhyrningsfylki er því

$$\tfrac{1}{2}n^2 - \tfrac{1}{2}n + \tfrac{1}{2}n^2 - \tfrac{1}{2}n + n = n^2.$$

8.5.4 Úrlausn á jöfnuhneppi með efra þríhyrningsfylki

Hugsum okkur nú að A sé efra þríhyrningsfylki. Þá verður jöfnuhneppið

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1, \\ a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2, \\ a_{33}x_3 + \cdots + a_{3n}x_n &= b_3, \\ \vdots \quad & \quad \vdots \\ a_{nn}x_n &= b_n, \end{aligned}$$

Við getum rakið okkur upp línurnar og fundið $x_n, x_{n-1}, \ldots, x_1$ hverja af annarri

$$\begin{aligned} x_n &= b_n/a_{nn}, \\ x_{n-1} &= (b_{n-1} - a_{n-1,n}x_n)/a_{n-1,n-1}, \\ \vdots \quad & \quad \vdots \qquad\qquad \vdots \\ x_1 &= (b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1,n}x_n)/a_{11}. \end{aligned}$$

Aðgerðafjöldinn er sá sami og í úrlausn neðra þríhyrningshneppisins.

8.5.5 Línuaðgerðir

Gerum ráð fyrir að við séum að ryðja 4×4 fylki með Gauss-eyðingu og að við séum búin með fyrsta dálkinn. Næsta skref er að nota línu 2 til þess að losna við stökin í sætum (3,2) og (4,2).

$$A = \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 0 & a & \cdot & \cdot \\ 0 & b & \cdot & \cdot \\ 0 & c & \cdot & \cdot \end{bmatrix}$$

Lína 3, l_3, verður þá að $l_3 - \frac{b}{a}l_2$, og lína 4, l_4, verður að $l_4 - \frac{c}{a}l_2$.

Þessar tvær aðgerðir má einnig framkvæma með því að margfalda fylkið að ofan frá vinstri með fylkinu

$$M_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{b}{a} & 1 & 0 \\ 0 & -\frac{c}{a} & 0 & 1 \end{bmatrix}$$

8.5.6 Ný sýn á Gauss-eyðingu

Það að ryðja fylkið eins og hér á undan, felst því í því að margfalda A frá vinstri með þremur fylkjum M_1, M_2 og M_3 sem eru þannig að M_i er einingafylkið nema í sætum $(i+1, i), \ldots, (n, i)$ eru tölur sem eru hugsanlega frábrugðnar 0. Fylkið M_i sér þá um að eyða í dálk i.

Athugum að Gauss-eyðing skilar fylki U á efra þríhyrningsformi.

Við getum því skrifað

$$M_3M_2M_1A = U$$

Almennt, fyrir $n \times n$ fylki þá getum við skrifað

$$M_{n-1} \cdots M_2M_1A = U,$$

þar sem M_i eru fylki eins og lýst er hér að ofan.

8.5.7 Nánar um M_i

Við sjáum að ef

$$M_i = \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & & \\ & & 1 & & & & \\ & & m_{i+1,i} & 1 & & & \\ & & \cdot & & \cdot & & \\ & & \cdot & & & \cdot & \\ & & m_{n,i} & & & & 1 \end{bmatrix},$$

þá er

$$M_i^{-1} = \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & & \\ & & 1 & & & & \\ & & -m_{i+1,i} & 1 & & & \\ & & \cdot & & \cdot & & \\ & & \cdot & & & \cdot & \\ & & -m_{n,i} & & & & 1 \end{bmatrix}.$$

Eins þá er auðvelt að sjá að

$$M_i^{-1}M_j^{-1} = \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & & \\ & & 1 & & & & \\ & & -m_{i+1,i} & \cdot & & & \\ & & \cdot & & 1 & & \\ & & \cdot & & -m_{j+1,j} & \cdot & \\ & & \cdot & & \cdot & & \cdot & \\ & & -m_{n,i} & & -m_{n,j} & & & 1 \end{bmatrix}$$

Það er

$$M_1^{-1}M_2^{-1}\cdots M_{n-1}^{-1} = \begin{bmatrix} 1 & & & & & \\ -m_{2,1} & 1 & & & & \\ -m_{3,1} & -m_{3,2} & 1 & & & \\ & -m_{4,2} & -m_{4,3} & 1 & & \\ & \cdot & & & \cdot & \\ -m_{n,1} & -m_{n,2} & -m_{n,3} & \cdot & -m_{n,n-1} & 1 \end{bmatrix}$$

8.5.8 LU-þáttun

Þetta hefur í för með sér að ef við skilgreinum $L = M_1^{-1}M_2^{-1}\cdots M_{n-1}^{-1}$ þá er

$$A = LU$$

eða

$$A = \begin{bmatrix} 1 & & & & & & & \\ -m_{2,1} & 1 & & & & & & \\ -m_{3,1} & -m_{3,2} & 1 & & & & & \\ & -m_{4,2} & -m_{4,3} & 1 & & & & \\ & \cdot & & & \cdot & & & \\ & \cdot & & & & \cdot & & \\ & \cdot & & & & & \cdot & \\ -m_{n,1} & -m_{n,2} & -m_{n,3} & \cdot & \cdot & \cdot & -m_{n,n-1} & 1 \end{bmatrix} U$$

Þannig að með því að framkvæma Gauss-eyðingu á A og halda utanum aðgerðirnar (M_i'in) og niðurstöðuna U þá fæst LU-þáttun á A.

8.5.9 LU-þáttun og sköluð hlutvending

Aðferðin hér að framan gerði ráð fyrir að stak $a_{i,i}$ yrði aldrei 0 (þá getum við ekki notað þá línu til þess að eyða). Eins hugsuðum við ekkert út í styttingarskekkjur sem við búum til.

Ef við framkvæmum Gauss-eyðinguna með skalaðri hlutvendingu þá ráðum við bót á báðum þessum atriðum, því þá veljum við aldrei línu með forystustuðul 0 og við minnkum skekkjuna eins og hefur komið fram áður.

Aðvörun: Þegar við notum skalaða hlutvendingu þá uppfylla fylkin L og U ekki endilega $LU = A$. Þess í stað fæst

$$LU = PA$$

þar sem fylkið P umraðar línunum í A í samræmi við umröðunarvigrinn $\mathbf{r}$. Það er, stökin í P eru 0, nema $p_{i,r_i} = 1$.

8.5.10 Jöfnuhneppi leyst með LU-þáttun

Við skiptum nú úrlausnarferlinu á $A\mathbf{x} = \mathbf{b}$ í þrjú skref

(i) LU **-þáttun:** Reiknum út neðra þríhyrningsfylki L og efra þríhyrningsfylki U með skalaðri hlutvendingu. Höldum utanum $\mathbf{r}$ (og þar með P). Þá er

$$LU = PA.$$

(ii) **Forinnsetning:** Leysum $L\mathbf{y} = P\mathbf{b}$.

(iii) **Endurinnsetning:** Leysum $U\mathbf{x} = \mathbf{y}$.

Lausnin sem við leitum að er þá $\mathbf{x}$, því

$$P\mathbf{b} = L\mathbf{y} = UL\mathbf{x} = PA\mathbf{x},$$

sem er jafngilt því að $\mathbf{b} = A\mathbf{x}$

8.5.11 Fjöldi reikniaðgerða fyrir LU-þáttun

Heildarfjöldi reikningsaðgerða til þess að framkvæma LU-þáttunina er

$$\frac{2}{3}n^3 - \frac{1}{2}n^2 - \frac{7}{6}n.$$

Liðir (ii) og (iii) krefjast svo $n^2 + n^2 = 2n^2$ aðgerða til viðbótar. Samanlagður fjöldi aðgerða er því

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n.$$

Ef n er stór tala, segjum $n = 1000$, þá er fyrsti liðurinn lang stærstur og við getum slegið á aðgerðafjöldann með $\frac{2}{3}n^3$.

Þetta er töluvert betra heldur en að reikna A^{-1} og svo $\mathbf{x} = A^{-1}\mathbf{b}$, en þá er heildafjöldi aðgerða $2n^3$.

8.5.12 Mörg jöfnuhneppi

Ef við þurfum að leysa mörg jöfnuhneppi með sama stuðlafylkið þá koma kostir LU-þáttunar vel í ljós.

Gefið A og $\mathbf{b}_1, \dots, \mathbf{b}_m$ þá leitum við að vigrum $\mathbf{x}_1, \dots, \mathbf{x}_m$ þannig að

$$A\mathbf{x}_i = \mathbf{b}_i, \qquad \text{fyrir } i = 1, \dots, m.$$

Við þurfum bara að framkvæma LU-þáttunina einu sinni, en innsetningarnar í lið (ii) og (iii) framkvæmum við m-sinnum. Heildar fjöldi aðgerða er þá

$$\frac{2}{3}n^3 + (2m - \frac{1}{2})n^2 - (m - \frac{1}{6})n.$$

8.6 Fastapunktsaðferðir fyrir línuleg jöfnuhneppi

8.6.1 Ítrekunaraðferðir til þess að leysa línuleg jöfnuhneppi

Munum að samanlagður fjöldi reikniaðgerða sem þarf til þess að leysa $n \times n$ línulegt jöfnuhneppi $A\mathbf{x} = \mathbf{b}$ með Gauss-eyðingu, for- og endurinnsetningu er $\sim \frac{2}{3}n^3$.

Ef jöfnuhneppið er jafngilt hneppinu

$$\mathbf{x} = T\mathbf{x} + \mathbf{c}$$

þá getum við sett upp fastapunktsferð til þess að leysa þetta hneppi með því að giska á eitthvert nálgunargildi $\mathbf{x}^{(0)}$ fyrir lausnina og ítra síðan með formúlunni

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c}, \qquad k = 0, 1, 2, \dots,$$

í þeirri von að runan $(\mathbf{x}^{(k)})$ stefni á réttu lausnina $\mathbf{x}$ á upprunalega jöfnuhneppinu.

Það þarf $n^2 - n$ aðgerðir til þess að reikna út margfeldið $T\mathbf{v}$ fyrir vigur $\mathbf{v} \in \mathbb{R}^n$ og því getum við komist upp með að taka $\approx \frac{2}{3}n$ ítrekanir áður en heildaraðgerðafjöldinn er kominn upp fyrir aðgerðafjöldann í Gauss-eyðingu, ásamt for- og endurinnsetningu.

8.6.2 Fastapunktsítrekun til þess að leysa línuleg jöfnuhneppi

Við ætlum nú að gera ráð fyrir að jafnan $A\mathbf{x} = \mathbf{b}$ sé jafngild

$$\mathbf{x} = T\mathbf{x} + \mathbf{c}$$

giskum á eitthvert nálgunargildi $\mathbf{x}^{(0)}$ fyrir lausnina $\mathbf{x}$ og skilgreinum síðan rununa

$$\mathbf{x}^{(k+1)} = T\mathbf{x}^{(k)} + \mathbf{c}, \qquad k = 0, 1, 2, \dots,$$

Allt er nú undir því komið að $n \times n$ fylkið T sé vel valið.

8.6.3 Fastapunktsítrekun - skekkjumat

Við skilgreinum nú skekkjuna í k-ta ítrekunarskrefinu $\mathbf{e}^{(k)} = \mathbf{x} - \mathbf{x}^{(k)}$. Þá gildir formúlan

$$\mathbf{e}^{(k)} = T\mathbf{e}^{(k-1)} = T^2\mathbf{e}^{(k-2)} = \cdots = T^k\mathbf{e}^{(0)}$$

sem við höfum áður séð í athugun okkar á fastapunktsaðferðinni.

Nú beitum við

$$\|\mathbf{e}^{(k)}\| \leq \|T^k\|\|\mathbf{e}^{(0)}\| \leq \|T\|^k\|\mathbf{e}^{(0)}\|$$

Við höfum $\mathbf{x}^{(1)} - \mathbf{x}^{(0)} = T\mathbf{x}^{(0)} + \mathbf{c} - \mathbf{x}^{(0)}$ og $\mathbf{c} = \mathbf{x} - T\mathbf{x}$ og þar með

$$\mathbf{x}^{(1)} - \mathbf{x}^{(0)} = T(\mathbf{x}^{(0)} - \mathbf{x}) - (\mathbf{x}^{(0)} - \mathbf{x}) = -(\mathbf{e}^{(0)} - T\mathbf{e}^{(0)}) = -(I - T)\mathbf{e}^{(0)}.$$

Þetta gefur jöfnuna

$$\mathbf{e}^{(0)} = -(I - T)^{-1}\big(\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\big).$$

Með smá útreikningi má sýna fram á að ef $\|T\| < 1$, þá er

$$\|(I - T)^{-1}\| \leq \frac{1}{1 - \|T\|}.$$

Við vorum komin með ójöfnurnar

$$\|\mathbf{e}^{(k)}\| \leq \|T\|^k\|\mathbf{e}^{(0)}\|$$

og niðurstaðan verður því

$$\|\mathbf{e}^{(k)}\| \leq \frac{\|T\|^k}{1 - \|T\|}\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|$$

Sem þýðir að fastapunktsaðferðin er samleitin þegar $\|T\| < 1$.

8.6.4 Skilyrði fyrir samleitni

Munum nú að $\rho(T)$ er rófgeisli fylkisins T sem er samkvæmt skilgreiningu tölugildi á stærsta eigingildi fylkisins T.

Rifjum líka upp að samkvæmt *setningu um rófgeisla* gildir fyrir sérhvern náttúrlegan fylkjastaðal $\|\cdot\|$ að $\rho(T) \leq \|T\|$ og að fyrir sérhvert $\varepsilon > 0$ gildir að hægt er að finna náttúrlegan fylkjastaðal þannig að

$$\|T\| \leq \rho(T) + \varepsilon.$$

Sérstaklega gildir í tilfellinu $\rho(T) < 1$ að til er náttúrlegur fylkjastaðall $\|\cdot\|$ þannig að $\|T\| < 1$.

Þetta þýðir að fastapunktsaðferðin er samleitin ef $\rho(T) < 1$.

8.6.5 Skiptingaraðferð (e. splitting method)

Við viljum setja upp fastapunktsaðferð til þess að leysa línulega jöfnuhneppið $A\mathbf{x} = \mathbf{b}$ með því að umrita jöfnuna yfir í jafngilda línulega jöfnu

$$\mathbf{x} = T\mathbf{x} + \mathbf{c}.$$

Gerum ráð fyrir að $A = M - N$ þar sem M er andhverfanlegt fylki. Þá jafngildir $A\mathbf{x} = \mathbf{b}$ jöfnunni $M\mathbf{x} = N\mathbf{x} + \mathbf{b}$ og fastapunktsjafnan er

$$\mathbf{x} = M^{-1}N\mathbf{x} + M^{-1}\mathbf{b},$$

þar sem $T = M^{-1}N$ og $\mathbf{c} = M^{-1}\mathbf{b}$.

Þessi leið til þess að umrita línulega jöfnuhneppið $A\mathbf{x} = \mathbf{b}$ yfir í jafngilda hneppið $\mathbf{x} = T\mathbf{x}+\mathbf{c}$ nefnist *skiptingaraðferð* fyrir línulega jöfnuhneppið $A\mathbf{x} = \mathbf{b}$.

8.6.6 Jacobi-aðferð

Við skrifum $A = D - L - U$, þar sem D er hornalínufylkið með hornalínu A, L er neðra þríhyrningsfylki og U er efra þríhyrningsfylki

Við tökum $M = D$ og $N = L + U$ og fáum þá $T = D^{-1}(L + U)$ og $\mathbf{c} = D^{-1}\mathbf{b}$.

Þessi skiptingaraðferð er nefnd *Jacobi-aðferð*.

Rakningarformúlan er

$$\mathbf{x}^{(k+1)} = D^{-1}(L + U)\mathbf{x}^{(k)} + D^{-1}\mathbf{b}.$$

Ef við skrifum hana hnit fyrir hnit, þá fáum við fyrir $i = 1, 2, \dots, n$,

$$x_i^{(k+1)} = \frac{1}{a_{i,i}}\left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k)} - \sum_{j=i+1}^{n} a_{i,j}x_j^{(k)}\right)$$

8.6.7 Gauss-Seidel-aðferð

Augljós endurbót á Jacobi-aðferðinni er að nota gildið $x_i^{(k+1)}$ fyrir $i < j$ um leið og það hefur verið reiknað.

Við það breytist rakningarformúlan í Jacobi-aðferð í

$$x_i^{(k+1)} = \frac{1}{a_{i,i}}\left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{i,j}x_j^{(k)}\right)$$

Þetta svarar til þess að við veljum skiptingu á A með $M = D - L$ og $N = U$ og þar með að

$$T = (D - L)^{-1}U \qquad \text{og} \qquad \mathbf{c} = (D - L)^{-1}\mathbf{b}$$

8.6.8 SOR-aðferð (e. successive over-relaxation)

Það er hægt að hraða samleitni í Gauss-Seidel-aðferð með því að taka vegið meðaltal af gildinu $x_i^{(k+1)}$ sem kemur út úr Gauss-Seidel reikniritinu og næsta gildi á undan með vægisstuðli sem við táknum með ω.

Formúlan verður

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{i,i}}\left(b_i - \sum_{j=1}^{i-1} a_{i,j}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{i,j}x_j^{(k)}\right)$$

Þetta svarar til þess að við veljum

$$M = \frac{1}{\omega}D - L \quad og \quad N = \left(\frac{1}{\omega} - 1\right)D + U$$

og þar með að

$$T = \left(\frac{1}{\omega}D - L\right)^{-1}\left(\left(\frac{1}{\omega} - 1\right)D + U\right) \quad \text{og} \quad \mathbf{c} = \left(\frac{1}{\omega}D - L\right)^{-1}\mathbf{b}$$

8.6.9 Setning um samleitni Gauss-Seidel-aðferðarinnar

- Gerum ráð fyrir að A sé samhverft rauntölufylki með öll hornalínustökin jákvæð. Þá er Gauss-Seidel aðferðin samleitin ef og aðeins ef A er jákvætt ákvarðað.
- Ef fylkið A er jákvætt ákvarðað, þá er Gauss-Seidel-aðferð samleitin fyrir sérhvert val á upphafságiskun $\mathbf{x}^{(0)}$.

8.7 Newton-aðferð fyrir jöfnuhneppi

Að lokum þá skulum við skoða aðeins ólínuleg jöfnuhneppi.

Látum $f_k : I \to \mathbb{R}$, $k = 1, \dots, n$, þar sem I er svæði í $\mathbb{R}^n$ vera samfelld föll. Það getur komið sér vel að geta leyst ólínuleg jöfnuhneppi af gerðinni

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \qquad\vdots \\ f_n(x_1, \dots, x_n) = 0 \end{cases}$$

Svo heppilega vill til að aðferð Newtons virkar næstum óbreytt fyrir slík hneppi.

8.7.1 Jacobi-fylki

Skilgreinum $\mathbf{f} : I \to R^n$ með

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x}))$$

og gerum ráð fyrir að allar hlutafleiðurnar $\dfrac{\partial f_k}{\partial x_j}$ séu til og séu samfelldar.

Táknum Jacobi-fylki $\mathbf{f}$ með $\mathbf{f}'$, það er

$$\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \frac{\partial f_n}{\partial x_2}(x) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

8.7.2 Aðferð Newtons fyrir hneppi

Lausn á hneppinu hér að ofan er vigur $\mathbf{r}$ þannig að $\mathbf{f}(\mathbf{r}) = 0$.

Ef $\mathbf{r}$ er lausn hneppisins og $\mathbf{x}_0 \in \mathbb{R}^n$ er upphafságiskun á $\mathbf{r}$ má sjá að runan $(\mathbf{x}_n)$, þar sem

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h}_n^T$$

og $\mathbf{h}_n^T$ er lausn á jöfnuhneppinu

$$\mathbf{f}'(\mathbf{x}_n)\mathbf{h}_n^T = -\mathbf{f}(\mathbf{x}_n)$$

stefnir á lausnina $\mathbf{r}$.

Við getum metið skekkjuna með

$$\mathbf{e}_{n+1} \approx \|\mathbf{x}_{n+1} - \mathbf{x}_n\|$$

og forritið okkar helst næstum óbreytt, við þurfum aðeins að skipta abs út fyrir skipunina norm.

8.7.3 Matlab-forrit fyrir hneppi

```
function x = newtonNullHneppi(f,df,x0,epsilon)
%
%   x = newtonNullHneppi(f,df,x0,epsilon)
%
% Nálgar núllstöð fallsins f:Rn --> Rn með aðferð Newtons.
% Fallið df er Jacobi-fylki f, x0 er upphafságiskun
% á núllstöð og epsilon er tilætluð nákvæmni.
% x0 verður að vera dálkvigur og f verður að
% skila dálkvigrum

x = x0; mis = -df(x)\f(x);
% Ítrum meðan ástæða er til
while (norm(mis) >= epsilon)
    x = x + mis;
    mis = -df(x)\f(x);
end
```

Grafið $y = e^x$ sker lokaða ferilinn sem gefinn er með jöfnunni $x^4 + y^2 = 1$ í tveimur punktum. Notið aðferð Newtons til þess að nálga hnit þeirra með 5 aukastafa nákvæmni.

Það er alveg augljóst að punkturinn $\mathbf{r} = (0, 1)$ gefur lausn á jöfnuhneppinu. Við látum eins og ekkert sé og giskum á $\mathbf{x}^{(0)} = (0.5, 0.75)$

Hér er

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} y - e^x \\ x^4 + y^2 - 1 \end{bmatrix}$$

og

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} -e^x & 1 \\ 4x^3 & 2y \end{bmatrix}$$

Forritið hér að ofan gefur svo eftirfarandi töflu.

	n		$x^{(n)}$		$y^{(n)}$	$\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|$
	0		0.50000000000000		0.75000000000000	
1	0.17270262414568	1.10909912528477	0.18447541668198			
2	0.01946538693088	1.00638822766059	0.02028257413953			
3	0.00020831857772	1.00002048613263	0.00020930163934			
4	0.00000002190660	1.00000000020984	0.00000002190761			
5	0.00000000000000	1.00000000000000	0.00000000000000			
6	-0.00000000000000	1.00000000000000	0.00000000000000			

Tökum nú fyrir hinn skurðpunktinn og notum upphafságiskun $(-0.8, 0.25)$.

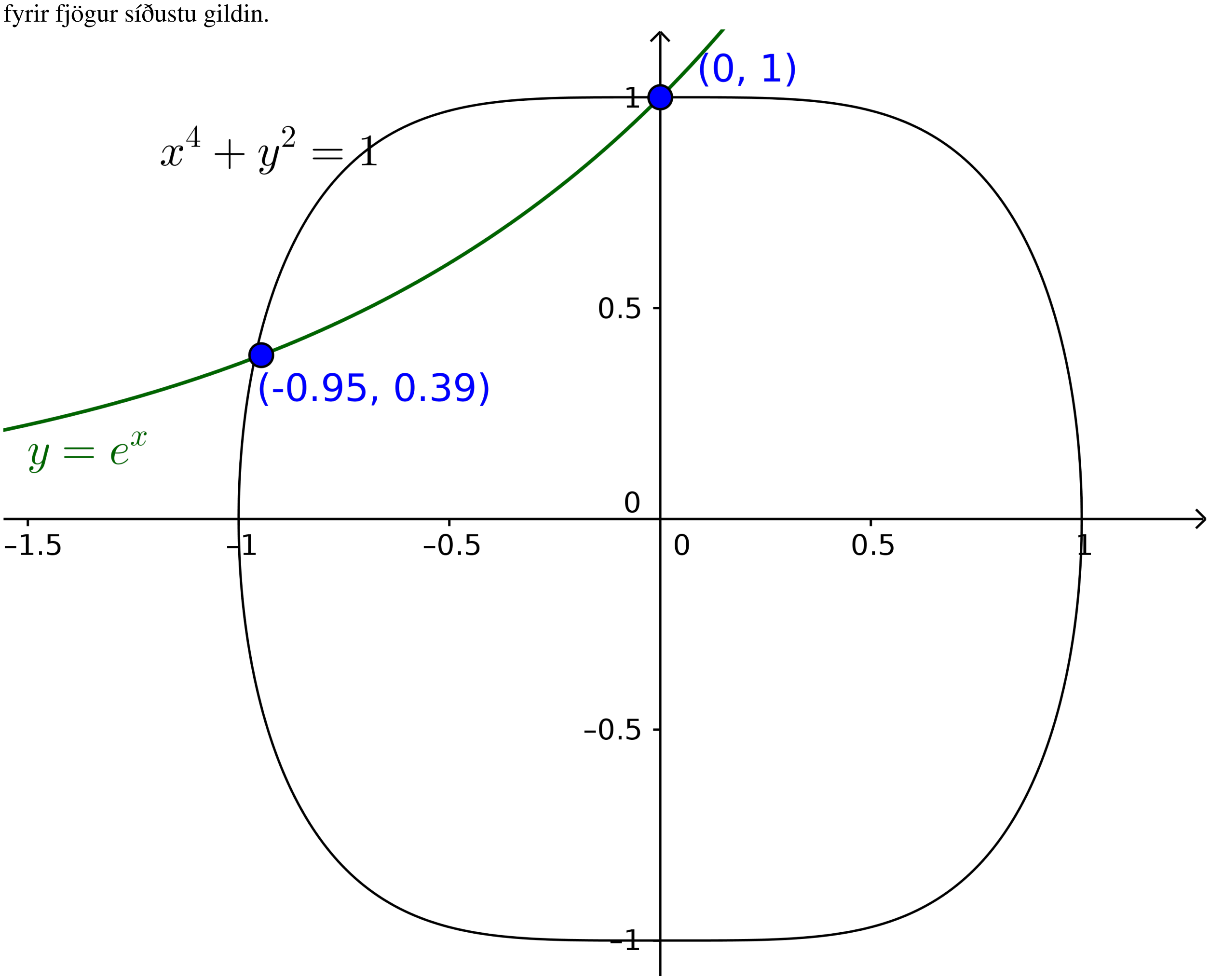
n	$x^{(n)}$	$y^{(n)}$	$\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|$
0	-0.80000000000000	0.25000000000000	
1	-1.03486380522268	0.34379785380788	0.07380487278262
2	-0.96968875917544	0.37842981331349	0.00919771982977
3	-0.96137076039507	0.38235523639344	0.00014098927991
4	-0.96124395918305	0.38241687590740	0.00000003252214
5	-0.96124392995056	0.38241689016050	0.00000000000000
6	-0.96124392995055	0.38241689016050	0.00000000000000

Nálgun okkar á skurðpunkti ferlanna í vinstra hálfplaninu er $(-0.96124392995055, 0.38241689016050)$

Í töfluna var ekki hægt að koma fyrir athugun á samleitnistiginu en hlutfallið

$$\frac{\|e_{n+1}\|}{\|e_n\|^2} \approx \frac{\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|}{\|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|^2} = 1.6$$

fyrir fjögur síðustu gildin.



Eigingildisverkefni

- *You can't give her that! It's not safe!*
- *It's a sword. They're not meant to be safe.*
- *She's a child!*
- *It's educational.*
- *What if she cuts herself?*
- *That will be an important lesson.*

– Terry Pratchett, Hogfather

9.1 Eigingildi og eiginvigrar

9.1.1 Nálgun á eigingildum og eiginvigrum

Skilgreining Látum A vera $n \times n$ fylki. Munum að $\lambda \in \mathbb{C}$ nefnist *eigingildi* fylkisins A ef til er $\mathbf{v} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ þannig að

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Vigurinn $\mathbf{v}$ nefnist þá *eiginvigur* fylkisins A og við segjum að hann svari til eigingildisins λ.

Athugasemd: Eigingildi fylkisins A eru nákvæmlega núllstöðvar kennimargliðunnar

$$p_A(z) = \det(zI - A), \qquad z \in \mathbb{C}.$$

Athugasemd: Ef $\mathbf{v}$ er eiginvigur fylkisins A, þá er $\alpha\mathbf{v}$ einnig eiginvigur fyrir sérhvert $\alpha \in \mathbb{C} \setminus \{\mathbf{0}\}$.

9.1.2 Gróf staðsetning á eigingildum: Skífusetning Gerschgorins

Skilgreinum

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|,$$

sem er summan af tölugildum stakanna í línu i *utan hornalínunnar* og látum

$$R_i = \{z \in \mathbb{C}\,;\, |z - a_{ii}| \le r_i\}$$

tákna skífuna með miðju í a_{ii} og geislann r_i. Þá gildir

1. Öll eigingildi A liggja í sammengi skífanna R_i.
2. Ef k af skífunum R_i mynda samanhangandi svæði U í $\mathbb{C}$ sem er sundlægt við hinar $n-k$ skífurnar, þá inniheldur U nákvæmlega k eigingildi.

9.1.3 Fylgisetning

Athugum fyrst að þar sem einingarfylkið er samhverft þá er $\det(A^T - \lambda I) = \det(A - \lambda I)$ og þar af leiðandi eru eigingildi A^T og A þau sömu.

Með því að beita Skífusetningu Gerschgorins á fylkið A^T fæst að við getum notað skífur með geisla sem eru sumur tölugilda stakanna í hverjum dálki að hornalínunni undanskilinni. Þetta fæst af því að summa yfir línu í A^T jafngildir summu yfir dálk í A.

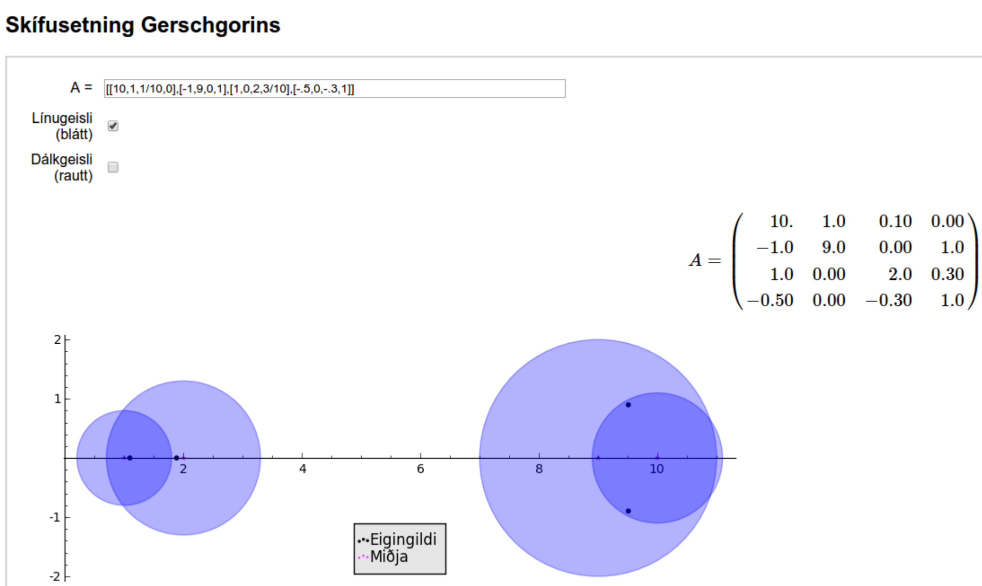
Nánar tiltekið, ef

$$c_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ji}|,$$
$$C_i = \{z \in \mathbb{C}\,;\, |z - a_{ii}| \le c_i\}$$

þá gildir að

1. Öll eigingildi A liggja í sammengi skífanna C_i.
2. Ef k af skífunum C_i mynda samanhangandi svæði U í $\mathbb{C}$ sem er sundlægt við hinar $n-k$ skífurnar, þá inniheldur U nákvæmlega k eigingildi.

9.1.4 Dæmi: Setning Gerschgorins



9.1.5 Eiginvigragrunnar

Nokkrar staðreyndir um eigingildi og eiginvigra:

1. Eiginvigrar sem svara til ólíkra eigingilda eru línulega óháðir.
2. Eiginvigrar sem svara til eins ákveðins eigingildis λ spanna hlutrúm í $\mathbb{C}^n$.
3. Við segjum að fylkið A sé *hornalínugeranlegt* ef til eru eigingildi $\lambda_1, \lambda_2, \ldots, \lambda_n$ og tilsvarandi eiginvigrar $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sem mynda grunn í $\mathbb{R}^n$.Þá er hægt að skrifa

$$A = T\Lambda T^{-1}$$

þar sem Λ er hornalínufylki með eigingildin $\lambda_1, \ldots, \lambda_n$ á hornalínunni og T er $n \times n$ fylki þannig að dálkur nr. k í því samanstendur af hnitum $\mathbf{v}_k$ miðað við staðalgrunninn í $\mathbb{R}^n$.
4. Ef fylkið A er samhverft, þá er það hornalínugeranlegt.

9.2 Veldaaðferð

9.2.1 Veldaaðferð

Hugsum okkur nú að við A sé hornalínugeranlegt og að við röðum eigingildunum á hornalínu Λ í minnkandi röð eftir tölugildi

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

Tökum einhvern vigur $\mathbf{x}^{(0)}$ og lítum á liðun hans í eiginvigra

$$\mathbf{x}^{(0)} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$$

Skilgreinum síðan rununa $\left(\mathbf{x}^{(m)}\right)$ með ítruninni

$$\mathbf{x}^{(m)} = A\mathbf{x}^{(m-1)}.$$

Við fáum þá

$$\begin{aligned}
\mathbf{x}^{(1)} = A\mathbf{x}^{(0)} &= \alpha_1 A\mathbf{v}_1 + \cdots + \alpha_n A\mathbf{v}_n \\
&= \alpha_1 \lambda_1 \mathbf{v}_1 + \cdots + \alpha_n \lambda_n \mathbf{v}_n,
\end{aligned}$$

$$\begin{aligned}
\mathbf{x}^{(2)} = A\mathbf{x}^{(1)} &= \alpha_1 \lambda_1 A\mathbf{v}_1 + \cdots + \alpha_n \lambda_n A\mathbf{v}_n, \\
&= \alpha_1 \lambda_1^2 \mathbf{v}_1 + \cdots + \alpha_n \lambda_n^2 \mathbf{v}_n \\
&\vdots \\
\mathbf{x}^{(m)} &= \alpha_1 \lambda_1^m \mathbf{v}_1 + \cdots + \alpha_n \lambda_n^m \mathbf{v}_n
\end{aligned}$$

Síðasti vigurinn gefur $\mathbf{x}^{(m)} = A^m \mathbf{x}^{(0)}$

$$\mathbf{x}^{(m)} = \lambda_1^m \big(\alpha_1 \mathbf{v}_1 + (\lambda_2/\lambda_1)^m \alpha_2 \mathbf{v}_{2+} \cdots + (\lambda_n/\lambda_1)^m \alpha_n \mathbf{v}_n\big)$$

Hnit númer i í þessum vigri er:

$$x_i^{(m)} = \lambda_1^m \big(\alpha_1 v_{1,i} + (\lambda_2/\lambda_1)^m \alpha_2 v_{2,i} + \cdots + (\lambda_n/\lambda_1)^m \alpha_n v_{n,i}\big)$$

Hugsum okkur nú að $|\lambda_1| > |\lambda_2|$. Þá fæst:

$$\frac{x_i^{(m)}}{x_i^{(m-1)}} = \frac{\lambda_1^m\big(\alpha_1 v_{1,i} + O((\lambda_2/\lambda_1)^m)\big)}{\lambda_1^{m-1}\big(\alpha_1 v_{1,i} + O((\lambda_2/\lambda_1)^{m-1})\big)}$$

Ef við höfum $\alpha_1 v_{1,i} \neq 0$, þá er niðurstaðan

$$\frac{x_i^{(m)}}{x_i^{(m-1)}} = \lambda_1 \frac{\big(1 + O((\lambda_2/\lambda_1)^m)\big)}{\big(1 + O((\lambda_2/\lambda_1))^{m-1}\big)} \to \lambda_1 \quad \text{þegar} \quad m \to \infty.$$

Skoðum aftur

$$\mathbf{x}^{(m)} = \lambda_1^m\big(\alpha_1\mathbf{v}_1 + (\lambda_2/\lambda_1)^m\alpha_2\mathbf{v}_2 + \cdots + (\lambda_n/\lambda_1)^m\alpha_n\mathbf{v}_n\big)$$

Ef $|\lambda_1| > |\lambda_2|$, þá gildir fyrir $j > 1$ að $(\lambda_j/\lambda_1)^m \to 0$ þegar $m \to \infty$ og

$$\lim_{m\to\infty} \frac{\mathbf{x}^{(m)}}{\lambda_1^m} = \alpha_1\mathbf{v}_1.$$

Þannig að ef $\mathbf{x}^{(0)}$ var valinn í upphafi þannig að $\alpha_1 \neq 0$, þá skilar þetta eiginvigrinum $\alpha_1\mathbf{v}_1$ fyrir eigingildið λ_1.

9.2.2 Reiknirit til þess að ákvarða stærsta eigingildi fylkis

Þegar við reiknum $\mathbf{x}^m$ eins og hér að framan þá er ekki ólíklegt að við lendum í undir- eða yfirflæðisvillum ef lengd $\mathbf{x}$ (skv. einhverjum staðli) stefnir á 0 eða $+\infty$. Til þess að ráða bót á þessu þá stöðlum við vigurinn í hverju skrefi með ℓ_∞ staðlinum á eftirfarandi hátt.

Við veljum $\mathbf{x}^{(0)}$ með einhverjum hætti og skilgreinum síðan

$$\mathbf{x}^{(m)} = \frac{A\mathbf{x}^{(m-1)}}{\|A\mathbf{x}^{(m-1)}\|_\infty}.$$

Þetta jafngildir því að setja

$$\mathbf{y}^{(m)} = A\mathbf{x}^{(m-1)}, \quad \text{og svo} \quad \mathbf{x}^{(m)} = \frac{\mathbf{y}^{(m)}}{y_{p_m}^{(m)}}$$

þar sem p_m er númerið á því hniti í vigrinum $\mathbf{y}^{(m)}$ sem hefur stærst tölugildi, það er hnit p_m uppfyllir

$$\|A\mathbf{x}^{m-1}\| = |y_{p_m}^{(m)}| = \|\mathbf{y}^{(m)}\|_\infty = \max_{1\le j\le n} |y_j^{(m)}|.$$

Ef mörg númer uppfylla þetta skilyrði, þá tökum við bara p_m sem lægsta gildið á j þar sem jafnaðarmerki gildir (það skiptir ekki máli fyrir nálgunina á hvaða j við veljum).

Þá fæst að

$$\begin{aligned}
\mathbf{x}^{(m)} &= \frac{A\mathbf{x}^{(m-1)}}{y_{p_m}^{(m)}} \\
&= \frac{A^2\mathbf{x}^{(m-2)}}{y_{p_m}^{(m)}\ y_{p_{m-1}}^{(m-1)}} \\
&= \frac{A^3\mathbf{x}^{(m-3)}}{y_{p_m}^{(m)}\ y_{p_{m-1}}^{(m-1)}\ y_{p_{m-2}}^{(m-2)}} \\
&\ \vdots \\
&= \frac{A^m\mathbf{x}^{(0)}}{y_{p_m}^{(m)}\ \ldots\ y_{p_1}^{(1)}}
\end{aligned}$$

Með því að skrifa $A^m\mathbf{x}^{(0)}$ með eiginvigrum eins og hér fyrir ofan

$$\mathbf{x}^{(m)} = \lambda_1^m\big(\alpha_1\mathbf{v}_1 + (\lambda_2/\lambda_1)^m\alpha_2\mathbf{v}_{2+}\cdots + (\lambda_n/\lambda_1)^m\alpha_n\mathbf{v}_n\big)$$

þá fæst

$$\mathbf{x}^{(m)} = \frac{\lambda_1^m}{|y_{p_m}^{(m)}|\ \ldots\ |y_{p_1}^{(1)}|}\big(\alpha_1\mathbf{v}_1 + (\lambda_2/\lambda_1)^m\alpha_2\mathbf{v}_{2+}\cdots + (\lambda_n/\lambda_1)^m\alpha_n\mathbf{v}_n\big)$$

Stuðlarnir við $\mathbf{v}_2, \ldots, \mathbf{v}_n$ stefna allir á 0 þannig að það er ljóst að $\mathbf{x}^{(m)}$ stefnir á eiginvigur sem tilheyrir λ_1. Nánar tiltekið stefnir $\mathbf{x}^{(m)}$ á eiginvigur sem hefur lengdina 1 í ℓ_∞ staðlinum, samkvæmt skilgreiningunni á $\mathbf{x}^{(m)}$.

Þar sem $\mathbf{x}^{(m)}$ er um það bil eiginvigur þá er $\mathbf{y}^{(m)} = A\mathbf{x}^{(m-1)} \approx \lambda_1\mathbf{x}^{(m-1)}$ sem segir okkur að ef við veljum p_{m-1}-ta hnitið úr jöfnunni þá fæst

$$y_{p_{m-1}}^{(m)} \approx \lambda_1$$

því p_{m-1}-ta hnitið í $\mathbf{x}^{(m-1)}$ er 1.

9.2.3 Samleitni

Nú kemur í ljós að $y_{p_{m-1}}^{(m)}$ stefnir á λ_1. Auk þess stefnir $\mathbf{x}^{(m)}$ á eiginvigur sem svarar til λ_1 og hefur lengdina 1 í l_∞ staðlinum.

Í útreikningum skilgreinum við því rununa $\lambda^{(m)} = y_{p_{m-1}}^{(m)}$. Við gefum okkur síðan þolmörk á skekkju TOL og reiknum úr runurnar þar til eitt af stoppskilyrðunum gildir:

$$\begin{aligned} |\lambda^{(m)} - \lambda^{(m-1)}| &< TOL \qquad \text{eða} \\ \|\mathbf{x}^{(m)} - \mathbf{x}^{(m-1)}\| &< TOL \qquad \text{eða} \\ \|A\mathbf{x}^{(m)} - \lambda^{(m)}\mathbf{x}^{(m)}\| &< TOL. \end{aligned}$$

9.2.4 Samhverf fylki

Munum að ef A er samhverft, þá hefur A eiginvigragrunn og eiginvigra sem svara til ólíkra eigingilda eru hornréttir.

Í þessu tilfelli er einfaldara að smíða reiknirit svona:

$$\begin{aligned} \mathbf{y}^{(m)} &= A\mathbf{x}^{(m-1)} \\ \lambda^{(m)} &= \mathbf{x}^{(m-1)^T}\mathbf{y}^{(m)} \\ \mathbf{x}^{(m)} &= \frac{\mathbf{y}^{(m)}}{\sqrt{(\mathbf{y}^{(m)})^T\mathbf{y}^{(m)}}} \end{aligned}$$

Samleitnin verður sú sama: $\lambda^{(m)}$ stefnir á stærsta eigingildið og $\mathbf{x}^{(m)}$ stefnir á tilsvarandi eiginvigur.

9.2.5 Setning um eigingildi og eiginvigra

Látum sem fyrr A vera $n \times n$ fylki, $\lambda_1, \ldots, \lambda_n$ vera eigingildi og $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vera tilsvarandi eiginvigra.

1. Látum $p(x) = a_0 + a_1x + \cdots + a_mx^m$ vera margliðu og skilgreinum $n \times n$ fylkið B með því að stinga A inn í p,

   $$B = p(A) = a_0I + a_1A + \cdots + a_mA^m$$

   Þá eru tölurnar $p(\lambda_1), \ldots, p(\lambda_n)$ eigingildi fylkisins $B = p(A)$ með tilsvarandi eiginvigrum $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

2. Ef A er andhverfanlegt þá eru $1/\lambda_1, \ldots, 1/\lambda_n$ eigingildi A^{-1} með tilsvarandi eiginvigrum $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

9.2.6 Dæmi um veldaaðferð

Sjá vikublað 14.

9.3 Andhverf veldaaðferð

Af síðustu setningu leiðir að fylkið $B = (A - qI)^{-1}$ hefur eigingildin

$$\mu_1 = \frac{1}{\lambda_1 - q}, \ \mu_2 = \frac{1}{\lambda_2 - q}, \ \cdots \ \mu_n = \frac{1}{\lambda_n - q}.$$

Hugsum okkur nú að við viljum finna nálgunargildi fyrir eigingildið λ_k og að við vitum út frá setningu Gerschgorins skífunum nokkurn veginn hvar það er staðsett.

Ef við erum með q nógu nálægt λ_k, þá verður μ_k stærsta eigingildi fylkisins $B = (A - qI)^{-1}$

Þá getum við beitt veldaaðferðinni til þess að búa til runu $\mu^{(m)} \to \mu_k$ og við fáum að

$$\lambda^{(m)} = \frac{1}{\mu^{(m)}} + q \to \lambda_k.$$

Ef veldaaðferðinni er beitt á fylkið $B = (A - qI)^{-1}$ þá þurfum við að reikna út $\mathbf{y}^{(m)} = (A - qI)^{-1}\mathbf{x}^{(m-1)}$ í hverju skrefi.

Þetta er gert þannig að fyrst framkvæmum við LU-þáttun á fylkinu $LU = (A - qI)$ og framkvæmum síðan for- og endurinnsetningu til þess að leysa $LU\mathbf{y}^{(m)} = x^{(m-1)}$.

En tölulegar aðferðir fyrir LU-þáttun skoðuðum við í *kafla 8.2*.

9.3.1 Reiknirit til þess að nálga eigingildi og eiginvigra

Takmarkið er að finna nálgun á eigingildinu λ_k.

1. Finnum $q \in \mathbb{R}$ sem liggur næst eigingildinu λ_k af öllum eigingildum A
2. Þáttum $LU = A - qI$.
3. Við veljum $\mathbf{x}^{(0)}$ með einhverjum hætti og leysum síðan $\mathbf{y}^{(m)}$ út úr jöfnunni

$$LU\mathbf{y}^{(m)} = \mathbf{x}^{(m-1)}.$$

4. Skilgreinum $\mathbf{x}^{(m)} = \mathbf{y}^{(m)}/y^{(m)}_{p_m}$ þar sem p_m er númerið á því hniti í $\mathbf{y}^{(m)}$ sem hefur stærst tölugildi, sem þýðir að það hnit uppfyllir

$$|y^{(m)}_{p_m}| = \|\mathbf{y}^{(m)}\|_\infty = \max_{1\le j\le n} |y^{(m)}_j|.$$

   Ef mörg númer uppfylla þetta skilyrði, þá tökum við bara p_m sem lægsta gildið á j þar sem jafnaðarmerki gildir.

Niðurstaðan verður að

$$\lambda^{(m)} = \frac{1}{y^{(m)}_{p_{m-1}}} + q \to \lambda_k$$

og $\mathbf{x}^{(m)}$ stefnir á tilsvarandi eiginvigur.

9.3.2 Dæmi um öfuga veldaaðferð

Sjá vikublað 14.

KAFLI 10

Monte Carlo hermanir

Chaos is found in greatest abundance wherever order is being sought. It always defeats order, because it is better organized. – Terry Pratchett, Interesting Times: The Play

10.1 Inngangur

Hér ætlum við að skoða hvernig hægt er að nota slembitölur til þess að heilda tölulega og leysa verkefni tengd hermun.

Við munum ekki skoða hvernig slembitölur eru fundnar í tölvum og hversu slembnar þær í raun og veru eru. Það er töluvert efni um þetta á netinu og í bókum.

- https://en.wikipedia.org/wiki/Pseudorandom_number_generator
- https://www.random.org/randomness/
- D. Knuth. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*.

Í Matlab og Octave gefa skipanirnar *rand*, *randn* og *randi* slembitölur með mismunandi eiginleika. *rand* gefur jafndreifðar slembitölur á bilinu $[0, 1]$, *randn* gefur normaldreifðar rauntölur og *randi* gefur heiltölur á ákveðnu bili.

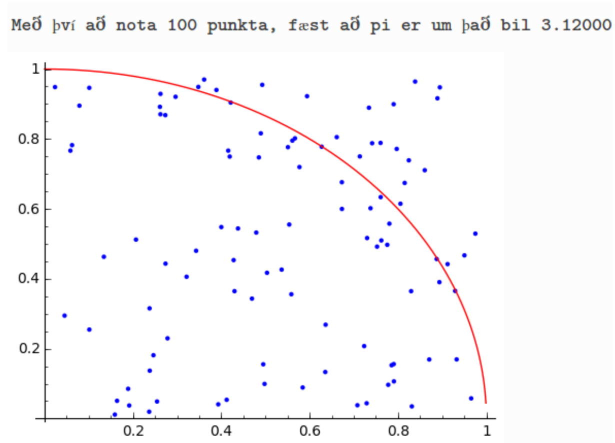
10.1.1 Nálgun á pi

Byrjum á að skoða einfalt verkefni. Við ætlum að nota flatarmál hringskífu til að nálga pi. Við vitum að flatarmál hringskífu með geisla r er πr^2. Setjum $r = 1$ og punkt af handahófi í rétthyrningnum $[-1, 1] \times [-1, 1]$, þá eru líkurnar á að hann lendi á hringskífunni $p = \pi/4$. Þannig að ef við getum fundið líkurnar p þá finnum við gildið á π. Líkurnar er hægt að herma með því að velja marga punkta af handahófi og athuga hversu hátt hlutfall þeirra lendir á skífunni. Af samhverfuástæðum er nóg fyrir okkur að skoða fyrsta fjórðung þannig að við veljum punkta $(x_1, y_1), \ldots, (x_n, y_n) \in [0, 1] \times [0, 1]$ af handahófi og skilgreinum S_n sem fjölda þeirra punkta sem lenda innan hringskífunnar og þá mun

$$\lim_{n\to\infty} \frac{S_n}{n} = \frac{\pi}{4}.$$

10.1.2 Forrit fyrir nálgun á pi

Forritum þetta í Sage sem er forritunarmál fyrir vísindalega útreikninga sem byggir á Python.



Byggt á. https://github.com/BC-Design/sage/blob/master/monte-carlo.html .

Athugasemd: Prófið að breyta gildinu á n í forritinu og sjáið hvort ekki er hægt að bæta nálgunina.

Aðvörun: Þar sem aðferðin er slembin þá fæst ekki alltaf sama svarið þegar forritið er keyrt fyrir sama gildið á n.

10.2 Heildi í einni breytistærð

10.2.1 Nálgun á heildi

Það er lítill vandi að nálga heildi með þessari aðferð. *Meðalgildi* falls f á bilinu $[a, b]$ er skilgreint sem

$$\overline{f} = \frac{1}{b-a}\int_a^b f(x)\,dx,$$

þannig að ef við getum fundið nálgun á meðalgildinu þá getum við nálgað heildið. Ef við veljum af handahófi punkta $x_1, x_2, \ldots, x_n \in [a, b]$ þá er eðlilegt að ætla að meðaltal fallgildanna í þessum punktum stefni á meðalgildi fallsins, það er

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f(x_i) = \overline{f}.$$

10.2.2 Nálgun á heildi: Dæmi

Prófum að nálga heildið

$$\int_0^2 x^2 - \frac{1}{2}x^3 + \frac{1}{10}x\,dx$$

með því að velja af handahófi 1000 punkta á bilinu $[0, 2]$ og nálga meðalgildið með meðaltali fallgildanna. Athugið að rétt svar er $13/15 \approx 0.86667$.

10.3 Margföld heildi og rúmmál

10.3.1 Margföld heildi

Aðferðin hér að ofan er ekki mjög góð til þess að nálga heildi, skekkjan er af stærðargráðunni $1/\sqrt{n}$ samanborið við $1/n^2$ í samsettu trapisureglunni. Helstu kostir þess að nota slembni til að nálga heildi koma fram þegar við þurfum að reikna heildi í mörgum breytistærðum. Þá þurfa hefðbundnar aðferðir eins og samsetta trapisureglan n^d punkta, þar sem d er fjöldi breytistærða. En helsti kostur Monte Carlo heildunar er að það er auðvelt að heilda yfir flókin svæði. Hefðbundnar aðferðir þurfa stikun á svæðinu til þess að ákvarða mörkin á heildunum en Monte Carlo þarf bara að geta ákvarðað hvort tiltekinn punktur er innan svæðisins eða ekki.

10.3.2 Margföld heildi: Dæmi

Reiknum rúmmál hlutarins $S \subset \mathbb{R}^3$ sem samanstendur af öllum punktum í $[0,1]^3$ sem uppfylla eftirfarandi ójöfnur

$$\begin{aligned} x^2 + \sin(y) &\leq z \\ x - z + \exp(y) &\leq 1 \end{aligned}$$

Rúmmálið, V, fæst með því að reikna eftirfarandi heildi

$$V = \int\int\int_S 1\, dx\, dy\, dz.$$

10.4 Hermun

Skoðum að lokum einfalt dæmi sem er ekki auðvelt að leysa beint, en er auðvelt að herma með því að nota sömu hugmyndir og hér að ofan.

10.4.1 Nál Buffons

Nál af einingarlengd er hent af handahófi á blað með tveimur samsíða línum og lengdin á milli línanna er 1. Gefið að miðja nálarinnar lendi á milli línanna, hverjar eru líkurnar á að nálin öll lendi á milli línanna?

KAFLI 11

Viðauki

The trouble with having an open mind, of course, is that people will insist on coming along and trying to put things in it. – Terry Pratchett

- Námskeiðið á Uglu
- Námskeiðið á Piazza

11.1 Kennsluáætlun

Why bother with a cunning plan when a simple one will do? – Terry Pratchett, Thud!

Dags.	Efni	Kaflar
08.01.16	1. Inngangur	1.1-1.7
13.01.16	2. Núllstöðvar	2.1-2.3
14.01.16	**Heimadæmi 1**	
15.01.16	2. Núllstöðvar	2.4-2.6
20.01.16	3. Brúun	3.1-3.2
21.01.16	**Heimadæmi 2**	
22.01.16	3. Brúun	3.3
27.01.16	3. Brúun	3.4
28.01.16	**Heimadæmi 3**	
29.01.16	3. Brúun	3.5-3.6
	Verkefni I kynnt	
03.02.16	3. Brúun	3.7
04.02.16	**Heimadæmi 4**	
05.02.16	3. Brúun	3.8-3.9
10.02.16	4. Töluleg diffrun	4.1-4.2
12.02.16	Samantekt	
	Verkefni I skilað	
17.02.16	4. Töluleg diffrun	4.3
18.02.16	**Heimadæmi 5**	
19.02.16	5. Töluleg heildun	5.1
24.02.16	5. Töluleg heildun	5.2
25.02.16	**Heimadæmi 6**	
26.02.16	5. Töluleg heildun 5.3-5.4	
	Verkefni II kynnt	
02.03.16	6. Upphafsgildisverkefni	6.1-6.2
03.03.16	**Heimadæmi 7**	
04.03.16	6. Upphafsgildisverkefni	6.3-6.4
09.03.16	6. Upphafsgildisverkefni	6.5-6.6
11.03.16	Hagnýtingar	Ítarefni
	Verkefni II skilað	
16.03.16	7. Jaðargildisverkefni	7.1-7.2
17.03.16	**Heimadæmi 8**	
18.03.16	7. Jaðargildisverkefni	7.3
23.03.16	*Páskafrí*	
25.03.16	*Páskafrí*	
30.03.16	8. Jöfnuhneppi	8.1-8.3
31.03.16	**Heimadæmi 9**	
01.04.16	8. Jöfnuhneppi	8.4-8.5
06.04.16	8. Jöfnuhneppi	8.6
07.04.16	**Heimadæmi 10**	
08.04.16	8. Jöfnuhneppi	8.7
13.04.16	9. Eigingildisverkefni	9.1-9.3
15.04.16	10. Monte Carlo hermanir	10.1-10.4
20.04.16	Prófundirbúningur	

11.2 Skipulag námskeiðsins

Always be wary of any helpful item that weighs less than its operating manual. – Terry Pratchett, Jingo

11.2.1 Kennsla

Fyrirlestrar verða á miðvikudögum klukkan 8:20-9:50 í HT-105, Háskólatorgi, og á föstudögum klukkan 10:00-11:30 í HT-105, Háskólatorgi. Kennari er Benedikt Steinar Magnússon <bsm@hi.is>.

Aðstoðarkennarar eru Halla Björg Sigurþórsdóttir, Jón Áskell Þorbjarnarson, Jónas Grétar Jónasson, Páll Ásgeir Björnsson og Pétur Rafn Bryde.

11.2.2 Kennsluefni

Kennslubókin eru þessar nótur, http://notendur.hi.is/bsm/stae405/. sem einnir er hægt að nálgast á pdf-formi. Auk þess set ég forritaskrár og annað ítarefni á Uglu eftir því sem við á.

Þeir sem vilja ítarlegri kennslubækur bendi ég á

- *A Friendly Introduction to Numerical Analysis* eftir Brian Bradie.
- *Numerical Mathematics and Computing* eftir Ward Cheney og David Kincaid
- *Introduction to applied numerical analysis* eftir R. W. Hamming.

> "The purpose of computing is insight, not numbers" -R. W. Hamming

11.2.3 Matlab/Octave

Við munum forrita töluvert í námskeiðinu. Til þess notum við annað hvort *Matlab*, http://www.mathworks.com/products/matlab/ eða *Octave*, http://www.gnu.org/software/octave/.

Á heimasíðu Kristjáns Jónassonar, http://notendur.hi.is/jonasson/matlab/ finnið þið leiðbeiningar um uppsetningu á Matlab. En fyrir þá sem kjósa frjálsan hugbúnað eru leiðbeiningar fyrir Octave hér http://www.gnu.org/software/octave/download.html.

Octave er að mestu leyti sambærilegt við Matlab, bæði ritháttur og svo styður það einnig *m*-skrárnar úr Matlab. Sjá nánar Differences_between_Octave_and_MATLAB.

Þið hafið fullkomið val um það hvort þið notið Matlab eða Octave við að leysa verkefni námskeiðsins.

Ítarefni fyrir Matlab/Octave:

- Inngangur að Matlab/Octave fyrir línulega algebru, https://notendur.hi.is/~bsm/linalg/.
- Kennslubók Kristjáns Jónassonar um Matlab fæst í Bóksölu Stúdenta.
- http://en.wikibooks.org/wiki/Matlab.
- Hjálpin í Matlab og Octave er einnig mjög gagnleg.

11.2.4 Heimadæmi og dæmatímar

Dæmatímarnir í námskeiðunu verða notaðir bæði til að reikna dæmi á töflu og sem stoðtímar fyrir heimadæmin. Alls verða lögð fyrir 10 heimadæmi. Þeim á að skila á fimmtudögum fyrir klukkan 16:00 í hólf viðkomandi dæmatímakennara. Heimadæmin er að finna á vikublöðum sem verða sett viku fyrr í möppuna *Vikublöð* á Uglu.

TIL AÐ ÖÐLAST PRÓFTÖKURÉTT ÞARF AÐ SKILA AÐ MINNSTA KOSTI 7 AF 10 HEIMADÆMUM.

Undanþágur frá þessari reglu fást eingöngu fyrir atbeina Náms- og starfsráðgjafar Háskólans.

Merkt verður við heimadæmin á Uglu undir *Verkefni og hlutapróf* og eru nemendur beðnir um að fylgjast með skráningunni þar og ganga úr skuggu um að allt sé rétt skráð.

11.2.5 Verkefni

Á misserinu verða tvö viðamikil forritunarverkefni. Í kennsluáætlun stendur í hvaða fyrirlestrum þau verða kynnt og hvenær á að skila þeim. Verkefnin eigið þið að leysa í hóp, tvö eða þrjú saman. Allir í hópnum eiga að vera virkir og taka þátt í að leysa alla liði verkefnisins. Forritað er í Matlab eða Octave.

Í vikunum sem skila á verkefunum þá munum við nota dæmatímana sem stoðtíma fyrir verkefnin. Skila á verkefnunum á föstudögum kl. 16:00, fyrra verkefninu 12. febrúar og seinna verkefninu 11. mars.

Matlab/Octave-forritin eiga að vera hluti af úrlausn og skal þeim skilað ásamt skýrslu í gegnum Uglu.

Ef við finnum sömu forritin í fleiri en einni úrlausn, þá lítum við á það sem svindl og lækkum einkunn hjá öllum sem skráðir eru fyrir þeim lausnum og eftir atvikum tilkynnum deildarforseta og setjum í farveg innan sviðsins (sbr. 51. gr. rgl. 569/2009 HÍ).

11.2.6 Lokapróf

Prófið verður 3 tímar og skiptist í fræðilegar krossaspurningar og skrifleg dæmi. Formúlublöð sem fylgja prófinu eru einu skriflegu hjálpargögnin sem leyfð verða. Þið **eigið** að taka með ykkur reiknivélar. Prófað verður bæði úr efni fyrirlestranna og úr dæmareikningi. Nauðsynlegt er að ná prófinu með einkunn 5. Verkefnaeinkunn gildir 30% af lokaeinkunn.

11.2.7 Námsmat

Til þess að standast námskeiðið þarf eftirfarandi:

- Skila að minnsta kosti 7 af 10 heimadæmum.
- Skila báðum verkefnunum.
- Ná lokaprófinu með einkunn 5.
- Lokaeinkunn (lokapróf 70%, verkefnaeinkunn 30%) þarf að vera að minnsta kosti 5.

Þau ykkar sem hafa prófrétt frá síðasta ári haldið verkefnaeinkunn. En þið þurfið að tilkynna mér það með tölvupósti sem fyrst.

11.3 Frágangur heimadæma

Let grammar, punctuation, and spelling into your life! Even the most energetic and wonderful mess has to be turned into sentences. – Terry Pratchett

- Skrifið upp **dæmið** og lausnina snyrtilega
- Vísið í setningar og niðurstöður sem þið notið
- Notið ekki rökfræðitákn eins og $\Leftarrow$, $\Rightarrow$, $\Leftrightarrow$, $\wedge$, $\vee$
- Textinn á að vera samfelldur og læsilegur (lesið hann sjálf yfir)

- Skýrt svar/niðurstaða

If you trust in yourself. . .and believe in your dreams. . .and follow your star. . . you'll still get beaten by people who spent their time working hard and learning things and weren't so lazy. – Terry Pratchett, The Wee Free Men

11.4 Gagnlegir tenglar

She got on with her education. In her opinion, school kept on trying to interfere with it. – Terry Pratchett, Soul Music

- Orðasafn Íslenska stærðfræðafélagsins http://stæ.is/os
- http://mathworld.wolfram.com/topics/NumericalMethods.html
- http://en.wikipedia.org
- Octave Programming Tutorial
- Octave Quick Reference (pdf)
- Getting Started with Matlab (pdf)
- Matlab Academy
- Matlab Tutorials and Learning Resources

Þ