

Luku 4

Käyrät

Hyvin usein esiintyy tilanteita, että tiettyä funktiota haluttaisiin approksimoida yksinkertaisemmalla funktiolla, tai että on annettu jokin pistejoukko jota haluttaisiin kuvata jollain yhtenäisellä käyrällä (tai pinnalla). Koska polynomit ovat yksinkertaisin mielenkiintoinen funktioluokka on varsin luonnollista, että näihin molempiin ongelmiin voisi yrittää löytää ratkaisua polynomien avulla.

4.1 Approksimointi

Lähdetään liikkeelle approksimointitehtävästä. On annettu jokin $f : [a, b] \rightarrow \mathbb{R}$; miten löydettäisiin polynomi joka olisi mahdollisimman lähellä funktiota f kyseisellä välillä? Eräs idea on käyttää Taylorin polynomeja: olkoon $p \in [a, b]$ ja lasketaan

$$T_n(x) = f(p) + f'(p)(x - p) + \cdots + \frac{f^{(n)}(p)}{n!} (x - p)^n$$

Ideana on, että kun n kasvaa saataisiin yhä parempi approksimaatio funktiolle f . Mutta miten tätä approksimaation hyvyttä olisi mielekästä mitata? Usein on luonnollista käyttää maksiminormia.

Määritelmä 4.1. *Olkoon $f \in C(I)$. Funktion f maksiminormi on*

$$\|f\|_\infty = \max_{x \in I} |f(x)|$$

Konkreettinen arvio virheestä saadaan seuraavasti.

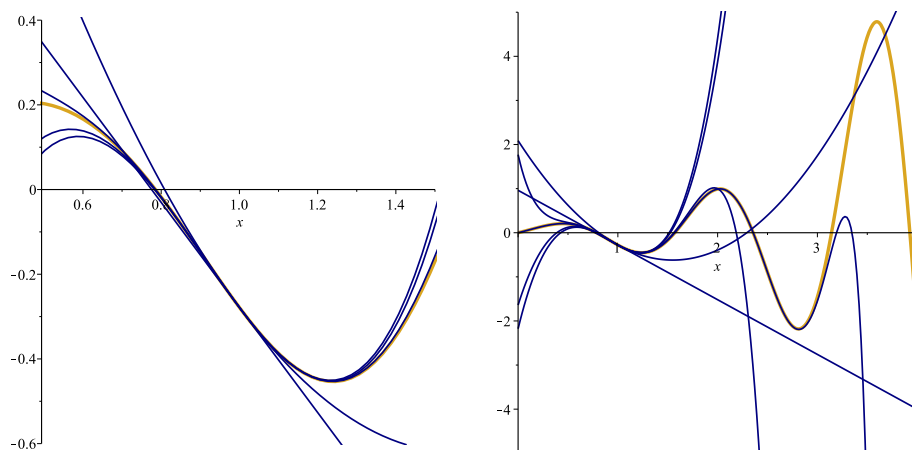
Lemma 4.1. *Olkoon $f \in C^{n+1}(I)$ missä $I = [a, b]$. Tällöin on olemassa M siten, että*

$$|f(x) - T_n| \leq \frac{M(b-a)^{n+1}}{(n+1)!} \quad (4.1)$$

Todistus. Taylorin lauseen A.1 perusteella on olemassa $\beta \in I$ siten, että

$$f(x) - T_n = \frac{f^{(n+1)}(\beta)(x-p)^{n+1}}{(n+1)!}$$

Koska $f \in C^{n+1}(I)$, niin valitsemalla nyt $M = \|f^{(n+1)}\|_\infty < \infty$ saadaan väite. ✧



Kuva 4.1: Vasemmalla approksimaatiot $T_1, \dots, T_4$ välillä $[0.5, 1.5]$ ja oikealla välillä $[0, 4]$. Oikealla on myös T_{20} .

Valitettavasti vain ”harvoin” $\|f - T_n\|_\infty \rightarrow 0$ kun $n \rightarrow \infty$.

Lemma 4.2. *Olkoon f analyyttinen pisteen p ympäristössä. Tällöin on olemassa ε siten, että $\|f - T_n\|_\infty \rightarrow 0$ kun $n \rightarrow \infty$ välillä $[p - \varepsilon, p + \varepsilon]$.*

Siis funktion pitää olla analyyttinen, ja silti pitää mahdollisesti rajoittua pienelle välille. Mutta itse asiassa Taylorin polynomit eivät käytännössä sovellu approksimointiin edes analyyttisille funktioille.

Esimerkki 4.1. Tarkastellaan funktiota $f(x) = e^{x-2} \sin(4x)$ välillä $[0, 4]$ ja valitaan $p = 1$. Kuvassa 4.1 on funktio ja Taylorin polynomit $T_1, \dots, T_4$. Approksimointitehtävissä usein haluttaisiin, että virhe on suurin piirtein saman suuruinen koko välillä. Kuten nähdään, niin Taylorin polynomit approksimoivat hyvin pisteen p lähellä, mutta varsin huonosti kaukana p :stä. ★

Jos halutaan approksimoida jatkuvia funktioita, tai funktioita joilla on vain ”vähän” jatkuvia derivaattoja, niin Taylorin polynomin muodostaminen on joko mahdotonta, tai sitten epämielekästä. Jos esimerkiksi $f(x) = |x|$ ja valitaan $p = 1$, niin selvästi $T_n(x) = x$ kaikilla $n \geq 1$. Approksimaatio ei voi mitenkään ”tietää”, että derivaatalla on epäjatkuvuuskohta origossa, joten approksimaatio negatiivisilla arvoilla ei parane kun n kasvaa.

Vaikka siis Taylorin polynomit ja tätä vastaava Taylorin lause A.1 ovat yleisesti ottaen tärkeitä, niin tähän nimenomaiseen ongelmaan ne soveltuvat huonosti.

Jos sitten oletetaan, että f on ”vain” jatkuva, niin voiko sitä ylipäättään approksimoida hyvin polynomeilla? Positiivinen vastaus seuraa kuuluisasta Weierstrassin approksimaatiolauseesta.

Lause 4.1. *Olkoon $f \in C(I)$ missä $I = [a, b]$ on rajoitettu väli. Tällöin kaikilla $\varepsilon > 0$ on olemassa polynomi g siten, että*

$$\|f - g\|_\infty < \varepsilon$$

Toinen tapa muotoilla sama asia on, että on olemassa jono polynomeja g_n siten, että $\|f - g_n\|_\infty \rightarrow 0$ kun $n \rightarrow \infty$. Luonnollisesti tällöin polynomien g_n asteet lähestyvät ääretöntä kun $n \rightarrow \infty$.

Weierstrassin lauseelle on monta hyvinkin erityyppistä todistusta. Katsotaan seuraavassa erään todistuksen idea koska se on konstruktivinen. Yleisyyttä rajoittamatta voidaan tarkastella väliä $I = [0, 1]$ koska sopivalla muuttujan vaihdolla aina voidaan palauttaa tehtävä tähän.

Olkoon $\mathbb{P}_n$ korkeintaan astetta n olevien polynomien avaruus. Siis jokainen $g \in \mathbb{P}_n$ voidaan esittää muodossa

$$g = c_0 + c_1x + \cdots + c_nx^n$$

joten sanotaan, että polynomit $\{1, x, \dots, x^n\}$ muodostavat avaruuden $\mathbb{P}_n$ *luonnollisen kannan*. Polynomit voidaan kuitenkin myös ilmaista *Bernsteinin polynomien* avulla (katso Lemma A.3). On siis olemassa jotkin toiset vakiot a_k siten, että

$$g = \sum_{k=0}^n a_k B_k^n \quad \text{missä} \quad B_k^n = \frac{n!}{k!(n-k)!} x^k (1-x)^{n-k}$$

Seuraavassa todistuksessa käytetään seuraavia Bernsteinin polynomien ominaisuuksia

Lemma 4.3.

$$\begin{aligned} \sum_{k=0}^n B_k^n &= 1 \\ \sum_{k=0}^n \frac{k}{n} B_k^n &= x \\ \sum_{k=0}^n \frac{k^2}{n^2} B_k^n &= x^2 + \frac{x(1-x)}{n} \end{aligned}$$

Todistus. Ensimmäinen kaava on todistettu lemmassa A.3. Helposti voidaan tarkistaa, että $\frac{k}{n} B_k^n = x B_{k-1}^{n-1}$, joten

$$\sum_{k=0}^n \frac{k}{n} B_k^n = x \sum_{k=1}^n B_{k-1}^{n-1} = x$$

Edelleen

$$\frac{k^2}{n^2} B_k^n = \frac{kx}{n} B_{k-1}^{n-1} = \frac{x}{n} B_{k-1}^{n-1} + \frac{(n-1)x^2}{n} B_{k-2}^{n-2}$$

Tästä seuraa

$$\sum_{k=0}^n \frac{k^2}{n^2} B_k^n = x^2 + \frac{x(1-x)}{n}$$

◇

Todistus (Lause 4.1). Lause pätee kaikilla $f \in C(I)$, mutta oletetaan seuraavassa, että $f \in C^2(I)$. Muodostetaan Bernsteinin polynomien avulla

$$f_n(x) = \sum_{k=0}^n f(k/n) B_k^n(x) \Rightarrow$$

$$f(x) - f_n(x) = \sum_{k=0}^n (f(x) - f(k/n)) B_k^n(x)$$

Taylorin Lauseen A.1 perusteella

$$f(k/n) = f(x) + f'(x)(k/n - x) + \frac{f''(\beta_k)}{2} (k/n - x)^2$$

Siispä

$$f(x) - f_n(x) = \sum_{k=0}^n f'(x)(x - k/n) B_k^n(x) - \frac{1}{2} \sum_{k=0}^n f''(\beta_k)(x - k/n)^2 B_k^n(x)$$

Lemman 4.3 perusteella

$$\sum_{k=0}^n f'(x)(x - k/n) B_k^n(x) = 0$$

Siispä

$$|f(x) - f_n(x)| = \frac{1}{2} \left| \sum_{k=0}^n f''(\beta_k)(x - k/n)^2 B_k^n(x) \right| \leq \frac{M}{2} \sum_{k=0}^n (x - k/n)^2 B_k^n(x)$$

missä $M = \|f''(x)\|_\infty$. Lemman 4.3 perusteella

$$\sum_{k=0}^n (x^2 - 2kx/n + k^2/n^2) B_k^n(x) = x^2 - 2x^2 + x^2 + \frac{x(1-x)}{n} = \frac{x(1-x)}{n}$$

joten lopulta

$$|f(x) - f_n(x)| \leq \frac{Mx(1-x)}{2n} \leq \frac{M}{8n}$$

◇

Tällä approksimaatiolla on seuraava mielenkiintoinen ja tärkeä ominaisuus.

Lemma 4.4.

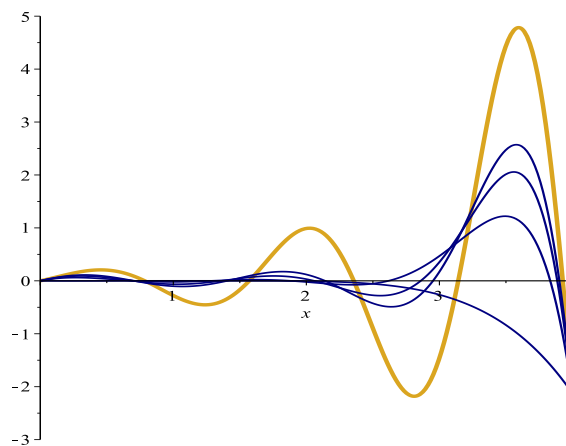
$$\|f_n\|_\infty \leq \|f\|_\infty \quad \text{kaikilla } n$$

Todistus.

$$|f_n(x)| = \left| \sum_{k=0}^n f(k/n) B_k^n \right| \leq \sum_{k=0}^n |f(k/n) B_k^n| \leq \sum_{k=0}^n |f(k/n)| B_k^n$$

$$\leq \|f\|_\infty \sum_{k=0}^n B_k^n = \|f\|_\infty$$

◇



Kuva 4.2: Bernsteinin polynomien avulla saadut approksimaatiot esimerkin 4.1 tapauksessa. Polynomien asteluvut ovat 5, 10, 15 ja 20.

Valitettavasti Bernsteinin polynomitkaan eivät ole kovin hyviä approksimoinnin kannalta koska suppeneminen on hyvin hidasta kuten kuvasta 4.2 nähdään. Jälleen voidaan kuitenkin todeta, että Bernsteinin polynomit ovat erittäin hyödyllisiä, mutta toisessa yhteydessä. Palataan tähän Bézier-käyrien yhteydessä.

On siis nähty, että polynomeilla voidaan mielivaltaisen hyvin approksimoida jatkuvia funktioita. Ongelman voisi kuitenkin muotoilla hieman toisin.

Ongelma 4.1. Olkoon annettu $f \in C(I)$. Etsi $g \in \mathbb{P}_n$ siten, että $\|f - g\|_\infty$ on mahdollisimman pieni. *

Koska Bernsteinin polynomien avulla saatu konstruktio suppenee niin hitaasti niin voisi kuvitella, että jollain toisella polynomilla saataisiin huomattavasti parempi approksimaatio. Palataan tähän ongelmaan kun ollaan tarkasteltu interpolointia.

Approksimointivirhettä on joskus luontevaa mitata toisellakin tavalla.

Määritelmä 4.2. Olkoon $f, g : [a, b] \rightarrow \mathbb{R}$. Funktioitten f ja g sisätulo on

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx$$

f ja g ovat ortogonaalisia, jos $\langle f, g \rangle = 0$. f :n normi on

$$\|f\|_2 = \sqrt{\langle f, f \rangle} = \left(\int_a^b f(x)^2 dx \right)^{1/2}$$

Myös tämän normin suhteen saadaan mielivaltaisen hyvä approksimaatio.

Lemma 4.5. Olkoon $f \in C(I)$ missä $I = [a, b]$. Tällöin on olemassa jono polynomeja g_n siten, että $\|f - g_n\|_2 \rightarrow 0$ kun $n \rightarrow \infty$.

Todistus. Weierstrassin lauseen 4.1 perusteella on olemassa jono polynomeja g_n siten, että $\|f - g_n\|_\infty \rightarrow 0$ kun $n \rightarrow \infty$. Tällöin

$$\|f - g_n\|_2^2 = \int_a^b (f - g_n)^2 dx \leq \int_a^b \|f - g_n\|_\infty^2 dx = (b - a) \|f - g_n\|_\infty^2$$

✧

Entäpä vastaava optimointitehtävä:

Ongelma 4.2. Olkoon annettu $f \in C(I)$. Etsi $g \in \mathbb{P}_n$ siten, että $\|f - g\|_2$ on mahdollisimman pieni. *

Nyt osoittautuu, että tälle ongelmalle saadaan ortogonaalisuuden avulla ratkaisu varsin helposti. Rajoitutaan jälleen seuraavassa yleisyyttä rajoittamatta välille $I = [0, 1]$. Nyt nähdään että avaruuteen $\mathbb{P}_n$ on hyödyllistä ottaa jälleen uusi kanta. Sanotaan, että kanta $\{g_0, \dots, g_n\}$ on *ortogonaalinen*, jos

$$\langle g_i, g_j \rangle = 0 \quad \text{kaikilla} \quad i \neq j$$

Ortogonaalisia kantoja on luonnollisesti useita, mutta tässä tapauksessa erityisen käteviä ovat *Legendren polynomit* $\{L_0, \dots, L_n\}$ jotka ovat siis ortogonaalisia ja lisäksi polynomin L_k aste on k (katso määritelmä A.3). Ortogonaalisuus on tärkeää seuraavan ominaisuuden takia. Olkoon $f \in \mathbb{P}_n$ ja haluttaisiin löytää $c_0, \dots, c_n$ siten, että

$$f = c_0 L_0 + \dots + c_n L_n$$

Kerrotaan puolittain polynomilla L_k ja otetaan sisätulo puolittain:

$$\langle f, L_k \rangle = \left\langle \sum_{j=0}^n c_j L_j, L_k \right\rangle = \sum_{j=0}^n c_j \langle L_j, L_k \rangle = c_k \|L_k\|_2^2 = \frac{c_k}{2k+1}$$

Siispä $c_k = (2k+1) \langle f, L_k \rangle$. Mutta samoin saadaan

Lemma 4.6. *Ongelman 4.2 yksikäsitteinen ratkaisu on*

$$g = \sum_{k=0}^n (2k+1) \langle f, L_k \rangle L_k$$

Lisäksi virheelle pätee

$$\|f - g\|_2^2 = \|f\|_2^2 - \|g\|_2^2 = \|f\|_2^2 - \sum_{k=0}^n (2k+1) (\langle f, L_k \rangle)^2$$

Todistus. Olkoon $c_k = (2k + 1)\langle f, L_k \rangle$. Todistus perustuu seuraavaan g :n ominaisuuteen:

$$\langle f - g, L_k \rangle = 0 \quad \text{kaikilla} \quad k = 0, \dots, n \quad (4.2)$$

Tämä pätee koska

$$\begin{aligned} \langle g, L_k \rangle &= \left\langle \sum_{j=0}^n c_j L_j, L_k \right\rangle \\ &= \sum_{j=0}^n c_j \langle L_j, L_k \rangle = (2k + 1) \langle f, L_k \rangle \langle L_k, L_k \rangle = \langle f, L_k \rangle \end{aligned}$$

Koska $\{L_0, \dots, L_n\}$ on avaruuden $\mathbb{P}_n$ kanta, niin ehdosta (4.2) seuraa, että $\langle f - g, q \rangle = 0$ kaikilla $q \in \mathbb{P}_n$. Tästä saadaan

$$\begin{aligned} \|f - g + q\|_2^2 &= \langle f - g + q, f - g + q \rangle \\ &= \|f - g\|_2^2 + 2\langle f - g, q \rangle + \|q\|_2^2 = \|f - g\|_2^2 + \|q\|_2^2 \geq \|f - g\|_2^2 \end{aligned}$$

Siis $\|f - g\|_2 \leq \|f - g + q\|_2$ kaikilla $q \in \mathbb{P}_n$ ja yhtäsuuruus saadaan vain jos $q = 0$.

Virhearvio saadaan soveltamalla Pythagoraan lausetta!

$$\begin{aligned} \|f\|_2^2 &= \|f - g + g\|_2^2 = \langle f - g + g, f - g + g \rangle \\ &= \|f - g\|_2^2 + 2\langle f - g, g \rangle + \|g\|_2^2 = \|f - g\|_2^2 + \|g\|_2^2 \end{aligned}$$

Edelleen

$$\begin{aligned} \|g\|_2^2 &= \left\langle \sum_{j=0}^n c_j L_j, \sum_{k=0}^n c_k L_k \right\rangle = \sum_{j=0}^n \sum_{k=0}^n c_k c_j \langle L_j, L_k \rangle \\ &= \sum_{k=0}^n c_k^2 \langle L_k, L_k \rangle = \sum_{k=0}^n \frac{c_k^2}{2k + 1} \end{aligned}$$

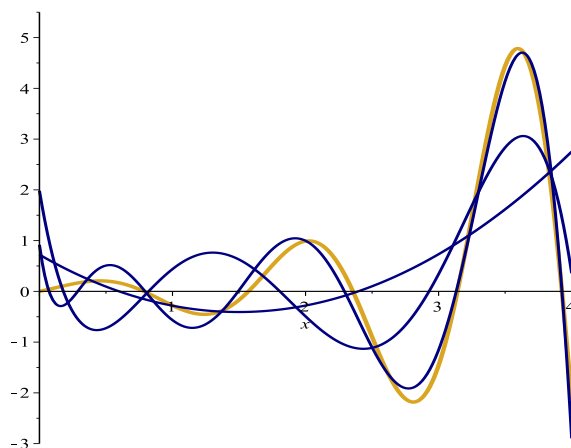
✧

Geometrisesti ajateltuna ongelman 4.2 ratkaisu saatiin siis projisoimalla annettu $f \in C(I)$ ortogonaalisesti avaruudelle $\mathbb{P}_n$. Konstruktio on täsmälleen sama kuin tapauksessa jossa jokin avaruuden $\mathbb{R}^n$ piste projisoitaisiin jollekin aliavaruudelle $V \subset \mathbb{R}^n$.

Kuvassa 4.3 on muutamia Legendren polynomien avulla saatuja approksimaatioita. Huomataan, että suppeneminen on varsin nopeaa, jo 8. asteen approksimaatio on varsin lähellä alkuperäistä funktiota.

Kertoimien $c_k = (2k + 1)\langle f, L_k \rangle$ laskeminen oli helppoa kannan ortogonaalisuuden takia, samalla tavalla kuin Fourier-sarjojen yhteydessä Fourier-kertoimien laskeminen on helppoa sini- ja kosinifunktioitten ortogonaalisuuden perusteella. Kertoimia c_k voikin siis sanoa funktion f (yleistetyiksi) Fourier-kertoimiksi kannan $\{L_0, \dots, L_n\}$ suhteen.

Huomattakoon, että edellinen lemma on mielekäs huomattavasti laajemmalle funktiojoukolle kuin jatkuvat funktiot. Funktion f jatkuvuuttahan ei tarvittu todistuksessa mitenkään. Oleellista oli vain, että Fourier-kertoimet $c_k = (2k + 1)\langle f, L_k \rangle$ on hyvin määritelty. Tätä asiaa käsitellään tarkemmin kurssilla FOURIER-ANALYYSI.



Kuva 4.3: Legendren polynomien avulla saadut approksimaatiot esimerkin 4.1 tapauksessa. Polynomien asteluvut ovat 2, 5 ja 8.

Käytännön toteutuksen kannalta ongelmaksi jää miten laskea integraali $\langle f, L_k \rangle$. Palataan tähän kysymykseen myöhemmin numeerisen integroinnin yhteydessä. Osottautuu, että Legendren polynomit ovat hyödyllisiä myös numeerisen integroinnin kannalta.

4.2 Interpolointi

Interpoloinnilla voidaan tarkoittaa kahta varsin erilaista asiaa:

Ongelma 4.3. On annettu jokin funktio f välillä $[a, b]$ ja halutaan jokin funktio g siten, että $f(x_j) = g(x_j)$ missä $a \leq x_0 < x_1 < \dots < x_n \leq b$. Tällöin sanotaan, että g *interpoloi* funktiota f pisteissä x_j . *

Ongelma 4.4. On annettu pisteet (x_j, y_j) , $0 \leq j \leq n$ ja halutaan jokin funktio g siten, että $y_j = g(x_j)$ kaikilla j . Tällöin sanotaan, että g *interpoloi* pistejoukkoa (x_j, y_j) . *

Päällisin puolin näyttäisi, että ongelma 4.3 on sama kuin ongelma 4.4, kun valitaan $y_j = f(x_j)$. Kuitenkin näkökulma on hyvin erilainen. Ensimmäisessä tapauksessa voidaan tarkastella interpolointivirhettä $\|f - g\|_\infty$ tai $\|f - g\|_2$ ja pyrkiä valitsemaan g siten, että virhe olisi mahdollisimman pieni, ottaen mahdollisesti huomioon funktion f erityispiirteet.

Jälkimmäisessä tapauksessa ei ole mitään funktiota, eikä siis voida puhua virheestäkään, vaan funktion g sopivuutta tai mielekkyyttä pitää arvioida muulla tavalla. Lähdetään liikkeelle ongelmasta 4.3 koska se on lähellä yllä käsiteltyä approksimointinäkökulmaa. Tarkastellaan ensin polynomi-interpolointia.

Ongelma 4.5. Olkoon $I = [a, b]$ ja $f \in C(I)$. Etsi $g \in \mathbb{P}_n$ siten, että $f(x_j) = g(x_j)$ missä $a \leq x_0 < x_1 < \dots < x_n \leq b$. *

Koska $\dim(\mathbb{P}_n) = n + 1$ ja on $n + 1$ interpolointiehtoa niin voisi kuvitella, että ongelmalla on yksikäsitteinen ratkaisu.

Jos on annettu vain 2 pistettä niin selvästi ratkaisu on näitten pisteitten kautta kulkeva suora. Jos $g(x) = c_0 + c_1x$ niin saadaan

$$\begin{cases} g(x_0) = c_0 + c_1x_0 = f(x_0) \\ g(x_1) = c_0 + c_1x_1 = f(x_1) \end{cases}$$

Huomaa, että tämä on lineaarinen yhtälöryhmä kertoimille c_j . Jos $c = (c_0, c_1)$ ja $v = (f(x_0), f(x_1))$, niin voidaan kirjoittaa

$$Ac = v \quad \text{missä} \quad A = \begin{pmatrix} 1 & x_0 \\ 1 & x_1 \end{pmatrix}$$

Ratkaisu on olemassa koska $\det(A) = x_1 - x_0$ ja oletuksen mukaan $x_0 \neq x_1$. Ratkaissamalla yhtälö saadaan

$$g(x) = \frac{x_1f(x_0) - x_0f(x_1)}{x_1 - x_0} + x \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f(x_0) \frac{x - x_1}{x_0 - x_1} + f(x_1) \frac{x - x_0}{x_1 - x_0}$$

Merkitään nyt

$$\mathcal{L}_0^1 = \frac{x - x_1}{x_0 - x_1} \quad \text{ja} \quad \mathcal{L}_1^1 = \frac{x - x_0}{x_1 - x_0}$$

Ratkaisu saatiin siis näitten polynomien lineaarikombinaationa, ja lisäksi tietysti $\{\mathcal{L}_0^1, \mathcal{L}_1^1\}$ on avaruuden $\mathbb{P}_1$ kanta. Tässä vaiheessa viimeistään voi tehdä huoletta sen johtopäätöksen, että uudenlainen interpolointi- tai approksimointitehtävä aina vaatii uuden kannan avaruudelle $\mathbb{P}_n$!

Määritelmä 4.3. *Olko annettu pisteet $x_0 < x_1 < \dots < x_n$. Tätä pistejoukkoa vastaavat Lagrangen polynomit ovat*

$$\mathcal{L}_k^n(x) = \prod_{j \neq k} \frac{x - x_j}{x_k - x_j} \quad , \quad 0 \leq k \leq n$$

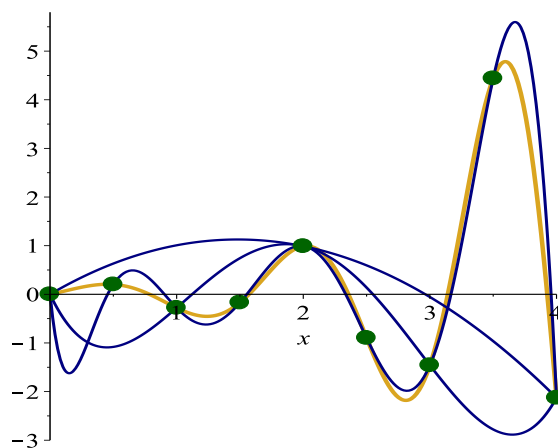
Selvästi pätee

$$\mathcal{L}_k^n(x_j) = \begin{cases} 1 & , j = k \\ 0 & , j \neq k \end{cases}$$

joten polynomit $\mathcal{L}_k^n$ ovat lineaarisesti riippumattomia, ja muodostavat siten avaruuden $\mathbb{P}_n$ kannan. Lagrangen polynomien avulla interpolointitehtävän ratkaisu voidaan kirjoittaa suoraan.

Lemma 4.7. *Ongelman 4.5 yksikäsitteinen ratkaisu saadaan kaavalla*

$$g(x) = \sum_{k=0}^n f(x_k) \mathcal{L}_k^n(x)$$



Kuva 4.4: Funktio ja interpoloivat polynomit joitten aste on 2, 4 ja 8. 8. asteen polynomi antaa jo varsin hyvän approksimaation paitsi välin päissä.

Todistus. On helppo tarkistaa, että g on ratkaisu. Yksikäsitteisyys jätetään harjoitustehtäväksi. $\diamond$

Esimerkki 4.2. Kuvassa 4.4 on muodostettu interpolointipolynomit funktiolle $f(x) = e^{x-2} \sin(4x)$. Interpolointipisteet on valittu tasavälein. Kuten nähdään niin jo varsin matala-asteisella polynomilla saadaan parempi approksimaatio kuin Bernsteinin polynomeilla. $\star$

Miten sitten interpolointivirhettä voitaisiin arvioida? Itse asiassa varsin hyvän arvion saa suhteellisen helposti kun vain ”arvaa oikein” miten lähteä liikkeelle. Olkoon jälleen $a = x_0 < x_1 < \dots < x_n = b$ ja merkitään $h_j = x_j - x_{j-1}$, $h = \max_j h_j$ ja $I_j = [x_{j-1}, x_j]$. Osoittautuu, että virheen arvioinnissa tarvitaan tietoa apufunktion $\Phi(x) = (x - x_0) \dots (x - x_n)$ suuruudesta.

Lemma 4.8.

$$\|\Phi\|_\infty \leq \frac{n! h^{n+1}}{4}$$

Todistus. Olkoon $q(x) = (x - x_{j-1})(x - x_j)$. Tällöin heti nähdään, että

$$|q(x)| \leq h^2/4 \quad \text{kun} \quad x \in I_j$$

Jos siis edelleen $x \in I_j$, niin

$$|\Phi(x)| \leq \frac{h^2}{4} (x - x_0) \dots (x - x_{j-2})(x_{j+1} - x) \dots (x_n - x)$$

Mutta nyt $x - x_{j-k} \leq kh$ ja vastaavasti $x_{j+k} - x \leq (k+1)h$ joten

$$(x - x_0) \dots (x - x_{j-2}) \leq j! h^{j-1} \quad \text{ja} \quad (x_{j+1} - x) \dots (x_n - x) \leq (n-j)! h^{n-j}$$

Tästä edelleen

$$|\Phi(x)| \leq \frac{h^2}{4} j! h^{j-1} (n-j)! h^{n-j} = \frac{h^{n+1}}{4} j! (n-j)! \leq \frac{n! h^{n+1}}{4}$$

◇

Tämän avulla saadaankin sitten

Lause 4.2. *Olkoon $f \in C^{n+1}(I)$ ja g Lemmassa 4.7 muodostettu interpolointipolynomi. Tällöin on olemassa $\beta \in I$ siten, että*

$$f(x) - g(x) = \frac{f^{(n+1)}(\beta)}{(n+1)!} \Phi(x)$$

Tästä saadaan virheen ylärajalle arvio

$$\|f(x) - g(x)\|_\infty \leq \frac{\|f^{(n+1)}\|_\infty \|\Phi\|_\infty}{(n+1)!} \leq \frac{\|f^{(n+1)}\|_\infty h^{n+1}}{4(n+1)}$$

Todistus. Määritellään funktio

$$u(s) = f(s) - g(s) - \frac{f(x) - g(x)}{\Phi(x)} \Phi(s)$$

Tässä siis kiinnitetään jokin $x \neq x_j$ ja ajatellaan, että s on ”varsinainen” muuttuja. Nyt siis havaitaan, että $u(x) = u(x_j) = 0$ kaikilla j , joten Lauseen A.2 perusteella on olemassa jokin β siten, että $u^{(n+1)}(\beta) = 0$. Sitten vain lasketaan:

$$\begin{aligned} u^{(n+1)}(\beta) &= f^{(n+1)}(\beta) - g^{(n+1)}(\beta) - \frac{f(x) - g(x)}{\Phi(x)} \Phi^{(n+1)}(\beta) \\ &= f^{(n+1)}(\beta) - \frac{f(x) - g(x)}{\Phi(x)} (n+1)! = 0 \quad \Rightarrow \\ f(x) - g(x) &= \frac{f^{(n+1)}(\beta)}{(n+1)!} \Phi(x) \end{aligned}$$

Tästä saadaan virheen ylärajalle arvio kun käytetään edellisen lemmän arviota funktiolle Φ . ◇

Huomaa, että jos interpolointi on *tasavälistä*, siis $h = h_j = x_j - x_{j-1}$ kaikilla j , niin $h = (b - a)/n$ jolloin virhearvioksi saadaan

$$\|f - g\|_\infty \leq \frac{M (b - a)^{n+1}}{4(n+1)n^{n+1}} \quad (4.3)$$

Esimerkki 4.3. Olkoon $f(x) = 1/x$ ja tarkastellaan väliä $[1, 3]$ ja interpolointipisteitä $x_0 = 1$, $x_1 = 2$ ja $x_2 = 3$. Nyt siis

$$\mathcal{L}_0^2 = \frac{1}{2}(x-2)(x-3) \quad , \quad \mathcal{L}_1^2 = -(x-1)(x-3) \quad , \quad \mathcal{L}_2^2 = \frac{1}{2}(x-1)(x-2)$$

$$g = f(1)\mathcal{L}_0^2 + f(2)\mathcal{L}_1^2 + f(3)\mathcal{L}_2^2 = \frac{1}{6}x^2 - x + \frac{11}{6}$$

Olkoon virhe $\mathcal{E} = f - g$. Maksimivirhe saadaan derivaatan nollakohdasta:

$$\mathcal{E}' = -1/x^2 - x/3 + 1 = 0 \quad \Rightarrow \quad x^3 - 3x^2 + 3 = 0$$

Kaksi nollakohdista on halutulla välillä: $z_1 \approx 1.35$ ja $z_2 \approx 2.53$. Tästä saadaan sitten

$$\|f - g\|_\infty = \max\{|\mathcal{E}(z_1)|, |\mathcal{E}(z_2)|\} \leq 0.05$$

Toisaalta Lauseen 4.2 mukaan

$$\|f - g\|_\infty \leq \frac{\|f'''\|_\infty \|\Phi\|_\infty}{6}$$

Nyt $f''' = -6/x^4$ joten $\|f'''\|_\infty = 6$. Edelleen $\Phi = (x-1)(x-2)(x-3) = x^3 - 6x^2 + 11x - 6$. Tämän maksimit ovat

$$|\Phi(2 \pm \frac{1}{\sqrt{3}})| = \frac{2}{3\sqrt{3}} \approx 0.4$$

joten saadaan arvio $\|f - g\|_\infty \leq 0.4$. ★

Numeerisen laskennan esihistoriallisella kaudella, siis ennen tietokoneaikaa, Lause 4.2 oli varsin tärkeä. Käsien laskiessa logaritmitaulukot olivat erittäin hyödyllisiä, ja interpoloinnin avulla voitiin sitten taulukkoa ”täydentää”, jos haluttiin laskea logaritmi sellaiselle luvulle joka oli taulukon lukujen välissä.

Esimerkki 4.4. Koska $\ln(1/x) = -\ln(x)$, niin luonnollisesti riittää laskea logaritmit arvoille $x > 1$. Logaritmitaulukossa on seuraavaa dataa:

x	$\ln(x)$
3.4758	1.245824668
3.4759	1.245853438
3.4760	1.245882207
3.4761	1.245910976

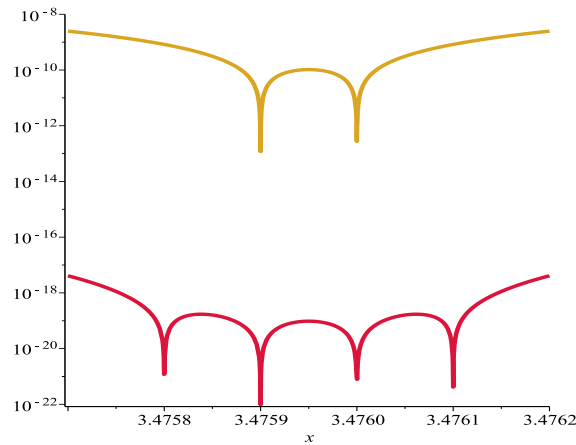
Haluttaisiin laskea logaritmi pistessä $\bar{x} = 3.47594$. Lasketaan kahden datapisteen lineaarinen interpoloiva polynomi g_1 ja kaikkien neljän pisteen avulla kolmannen asteen polynomi g_3 . Tästä saadaan

$$|\ln(\bar{x}) - g_1(\bar{x})| \approx 9.9 \cdot 10^{-11} \quad \text{ja} \quad |\ln(\bar{x}) - g_3(\bar{x})| \approx 9.2 \cdot 10^{-20}$$

Kolmannen asteen polynomilla saadaan siis huomattavasti tarkempi tulos. Lasketaan vielä Lauseen 4.2 antamat virhearviot. Nyt $f(x) = \ln(x)$ joten $f''(x) = -x^{-2}$ ja $f^{(4)}(x) = -6x^{-4}$. Tästä saadaan

$$|f(x) - g_1(x)| \leq \frac{0.0001^2}{8 \cdot 3.4759^2} \approx 1.0 \cdot 10^{-10} \quad \text{ja} \quad |f(x) - g_3(x)| \leq \frac{6 \cdot 0.0001^4}{16 \cdot 3.4758^4} \approx 2.6 \cdot 10^{-19}$$

Nähdään siis, että virhearviot ovat hyvin lähellä todellista virhettä. Kuvassa 4.5 on piirretty virheet tarkasteluvälillä. ★



Kuva 4.5: Esimerkin 4.4 interpolointivirheet sekä lineaarisen että kolmannen asteen interpoloinnin tapauksessa.

Palataan sitten ongelmaan 4.1. Millä polynomilla saataisiin paras approksimaatio maksiminormin mielessä? Itse asiassa tämä on varsin hankala kysymys, jonka ratkaisu löytyy kirjasta [12]. Tarkastellaan kuitenkin seuraavaa ongelmaa joka on ”melkein” sama.

Ongelma 4.6. Etsi interpolointipisteet $a \leq x_0 < \dots < x_n \leq b$ siten, että $\|\Phi\|_\infty$ on mahdollisimman pieni, missä $\Phi = (x - x_0) \dots (x - x_n)$. *

Haluttaisiin siis löytää pisteet x_j siten, että saataisiin Lemmaa 4.8 pienempi yläraja Φ :n normille. Osoittautuu, että vastaus tähän löytyy *Tshebyshevin polynomien* avulla. Lemman A.2 mukaan Tshebyshevin polynomille C_n pätee

$$|C_n(x)| \leq 1 \quad \text{kaikilla} \quad -1 \leq x \leq 1 \quad \text{ja}$$

$$C_n(\cos(\theta_{n-k})) = 0 \quad \text{missä} \quad \theta_k = \frac{(2k+1)\pi}{2n} \quad \text{ja} \quad k = 0, \dots, n-1$$

Määritellään sitten bijektio $\varphi : [-1, 1] \rightarrow [a, b]$

$$\varphi(x) = \frac{1}{2}(b-a)x + \frac{1}{2}(b+a) \quad \leftrightarrow \quad \varphi^{-1}(x) = \frac{2x-a-b}{b-a}$$

Olkoon edelleen $\tilde{C}_n = C_n \circ \varphi^{-1}$ jolloin siis

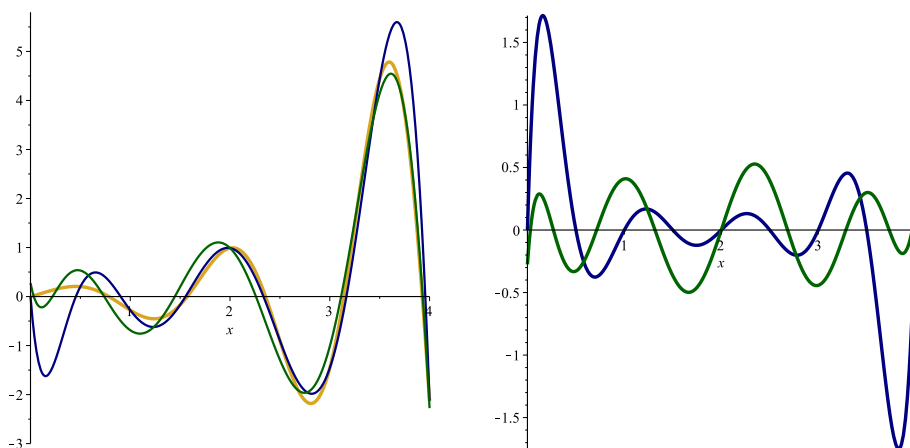
$$|\tilde{C}_n(x)| \leq 1 \quad \text{kaikilla} \quad a \leq x \leq b \quad \text{ja}$$

$$\tilde{C}_n(x_k) = 0 \quad \text{missä} \quad x_k = \varphi(\cos(\theta_{n-k})) \quad , \quad k = 0, \dots, n-1$$

Pisteitä x_k sanotaan välin $[a, b]$ *Tshebyshevin pisteiksi*. Nyt voidaan osoittaa

Lemma 4.9. *Ongelman 4.6 ratkaisu saadaan valitsemalla*

$$\Phi = \frac{2(b-a)^{n+1}}{4^{n+1}} \tilde{C}_{n+1} \quad \text{jolloin} \quad \|\Phi\|_\infty = \frac{2(b-a)^{n+1}}{4^{n+1}}$$



Kuva 4.6: Tshebyshevin polynomin nollakohtien avulla saatu approksimaatio esimerkin 4.1 tapauksessa verrattuna tasaväliseen interpolointiin. Polynomien asteluvut on 8. Oikealla vastaavat virheet. Tshebyshevin approksimaatio on huomattavasti parempi välin päissä.

Todistus. Huomaa, että $\Phi = x^{n+1} + \dots$ ja $C_{n+1} = 2^n x^{n+1} + \dots$, joten polynomin $\tilde{C}_{n+1}$ eteen tarvitaan ylläoleva kerroin. Tästä luonnollisesti heti saadaan Φ :n normi. Edelleen voidaan osoittaa, että millään muulla polynomilla ei saada pienempää normia. Tämä todistus löytyy kirjasta [4]. $\diamond$

Tästä saadaan sitten

Seuraus 4.1. *Olkoon $f \in C(I)$ ja olkoon g Tshebyshevin pisteitä käyttävä interpolointipolynomi. Tällöin interpolointivirheelle saadaan arvio*

$$|f(x) - g(x)| \leq \frac{2\|f^{(n+1)}\|_{\infty}(b-a)^{n+1}}{(n+1)!4^{n+1}} \quad (4.4)$$

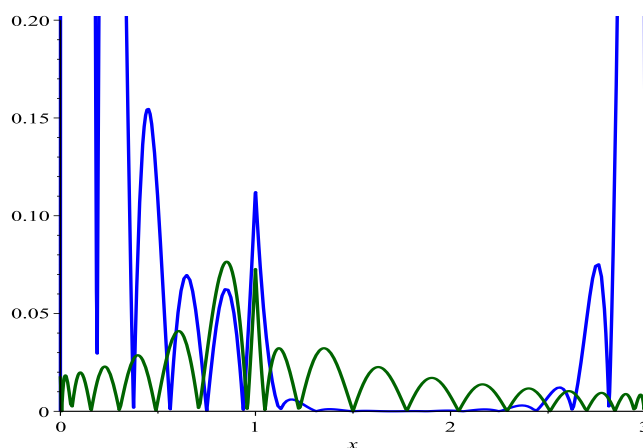
Todistus. Asetetaan $\Phi = \frac{2(b-a)^{n+1}}{4^{n+1}} \tilde{C}_{n+1}$ kuten edellisessä Lemmassa jolloin

$$|f(x) - g(x)| = \left| \frac{f^{(n+1)}(\beta)}{(n+1)!} \Phi(x) \right| \leq \frac{2\|f^{(n+1)}\|_{\infty}(b-a)^{n+1}}{(n+1)!4^{n+1}}$$

$\diamond$

Vertaamalla tasavälistä arviota (4.3) Tshebyshevin pisteillä saatuun arvioon (4.4) nähdään, että parannus ei ole kovinkaan merkittävä. Kuvassa 4.6 on verrattu tasavälistä interpolointia Tshebyshevin pisteitten avulla interpolointiin. Kuten nähdään niin ero on huomattava vain välin päissä.

Kaikissa näissä virhearvioissa on se ”vika”, että täytyy olettaa että annettu funktio f on monta kertaa jatkuvasti derivoituva. Käytännössä suppeneminen onkin huomattavasti hitaampaa jos f :llä on vain ”vähän” jatkuvia derivaattoja.



Kuva 4.7: Esimerkin 4.5 virhe sekä tasavälisessä (sininen) että Tshebyshev-interpoloinnissa (vihreä). Molemmissa tapauksissa virheellä on selvä piikki kohdassa $x = 1$.

Esimerkki 4.5. Tarkastellaan funktiota

$$f(x) = \begin{cases} x^3 - 2x + 2 & , 0 \leq x \leq 1 \\ -x^4/2 + 7x - 11/2 & , 1 \leq x \leq 3 \end{cases}$$

Funktio on jatkuva, mutta derivaatta ei ole jatkuva kohdassa $x = 1$. Tämän takia approksimaatio on huono pisteen $x = 1$ lähellä kuten kuvasta 4.7 nähdään. ☆

Edellä kuvattu interpolointi toimii hyvin jos pisteitten lukumäärä ei ole kovin iso. Kuitenkin polynomin asteluvun kasvaessa tulee ongelmia.

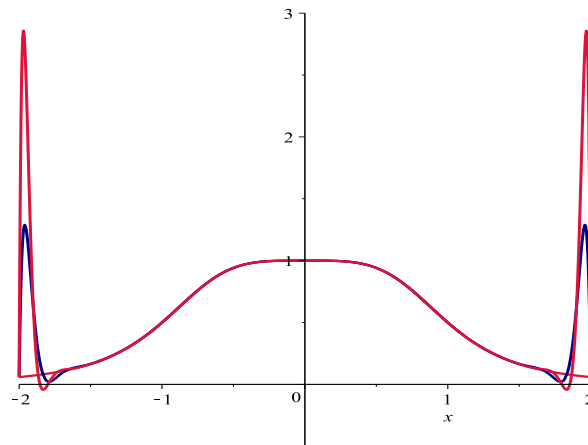
Esimerkki 4.6. *Rungen ilmiö.* Olkoon $f(x) = 1/(1 + x^4)$ ja halutaan interpoloida tasavälisesti välillä $[-2, 2]$. Kun interpolointipolynomin astetta kasvatetaan, tietyn rajan jälkeen tarkkuus ei enää parane vaan polynomi rupeaa ”värähtelemään” yhä voimakkaammin kuten kuvassa nähdään 4.8. Voidaan osoittaa, että värähtelyn amplitudi lähestyy ääretöntä. Siis jos g_n on astetta n oleva interpoloituva polynomi, niin $\|g_n\|_\infty \rightarrow \infty$ kun $n \rightarrow \infty$. ☆

Polynomi-interpolaatio ei siis ole kovin toimiva silloin jos interpolointipisteitten määrä on iso. Sama johtopäätös pätee ongelmaan 4.4: jos annettu datapistejoukko on iso niin polynomi-interpolaatio aiheuttaa Rungen ilmiön tyypisiä vaikeuksia. Huomaa, että tässä tapauksessa ongelmaa ei voi kiertää valitsemalla vaikkapa Tshebyshevin pisteitä interpolointiin koska datapisteet on annettu, eikä niitä voi mennä muuttamaan.

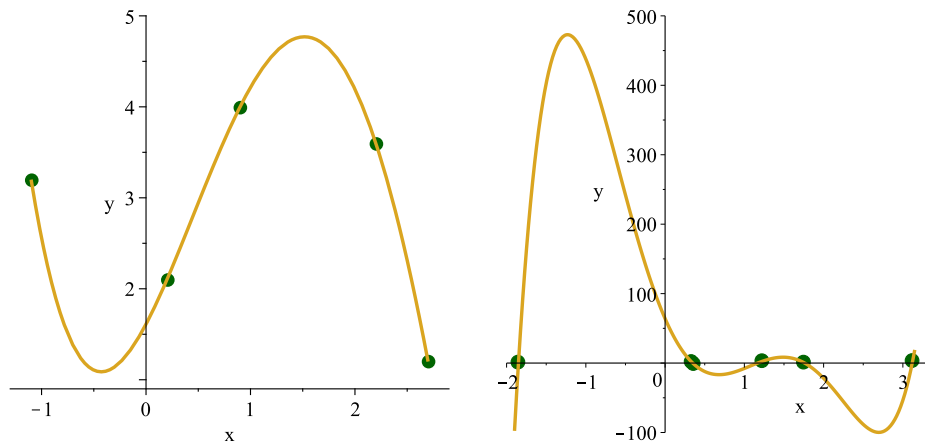
Miten sitten pitäisi menetellä isojen datamäärien kanssa? Koska polynomit ovat käteviä ei niistä tietenkään luovuta, vaan käytetään niitä eri tavalla.

4.3 Splinit

Kun interpolointipisteet on annettu pistejoukkona eivätkä ne tule mistään tunnetusta funktiosta, data on otettava sellaisenaan, eikä voida ruveta etsimään optimaalisia in-



Kuva 4.8: Esimerkin 4.6 funktio ja sen 24 ja 30 asteiset interpoloivat polynomit.



Kuva 4.9: Kaksi pistejoukkoa ja niitä vastaavat interpoloivat polynomit. Huomaa, että oikealla on välillä $[0, 1]$ kaksi interpolointipistettä hyvin lähellä toisiaan.

terpolointipisteitä. Tällöin polynomi-interpolointi johtaa helposti Rungen ilmiön tyyppisiin vaikeuksiin.

Esimerkki 4.7. Kuvassa 4.9 on annettu pistejoukot ja niitä vastaavat interpoloivat polynomit 2 tapauksessa. Vasemman kuvan lopputulos vaikuttaa järkevältä. Sen sijaan oikealla datan y -arvot ovat välillä $[-1.5, 3]$ kun taas vastaavan interpoloivan polynomin arvot ovat välillä $[-100, 500]$. Selvästi tämä käyrä ei mielekkäästi kuvaa annettua dataa.

★

Edellisen esimerkin tyyppiset vaikeudet voi välttää *splinien* avulla [19]. Ideana on, että jaetaan tarkasteltava väli osaväleihin ja vaaditaan, että interpoloiva funktio on (matala-asteinen) polynomi jokaisella osavälillä, ja vähintään jatkuva kaikkialla. Splinejä sanotaankin joskus *palapolynomeiksi*.

Olkoon $a = x_0 < x_1 < \dots < x_n = b$, $I = [a, b]$, $I_j = [x_{j-1}, x_j]$. Olkoon $f : I \rightarrow \mathbb{R}$ jokin funktio ja olkoon f_j funktion f rajoittuma välille I_j . Siis $f_j : I_j \rightarrow \mathbb{R}$ ja $f_j(x) = f(x)$ kaikilla $x \in I_j$. Nyt voidaan asettaa

Määritelmä 4.4. *Funktio $\mathcal{S} : I \rightarrow \mathbb{R}$ on splini, jos*

(i) $\mathcal{S} \in C(I)$ ja

(ii) $\mathcal{S}_j$ on polynomi kaikilla j

Pisteitä x_j sanotaan splinin solmupisteiksi.

Esimerkki 4.8. Olkoon annettu pisteet $(0, 1)$, $(1, 3)$ ja $(2, -1)$. Tämän perusteella voidaan piirtää murtoviiva näitten pisteitten kautta:

$$g(x) = \begin{cases} 2x + 1 & , \quad 0 \leq x \leq 1 \\ -4x + 7 & , \quad 1 \leq x \leq 2 \end{cases}$$

Selvästi siis g on jatkuva ja määräytyy yksikäsitteisesti annetun datan perusteella. Sovelluksissa haluttaisiin usein kuitenkin sileämpi käyrä; siis että vähintään ensimmäinen derivaatta olisi jatkuva. Lisätään polynomin astelukua: siis olkoon u funktio joka on toisen asteen polynomi kummallakin osavälillä. Merkitään $I_1 = [0, 1]$ ja $I_2 = [1, 2]$ ja olkoon u_j funktion u rajoittuma välille I_j . Nyt haluttaisiin, että u toteuttaisi ehdot

$$\begin{aligned} u_1(0) &= 1 & , & & u_1(1) &= u_2(1) = 3 \\ u'_1(1) &= u'_2(1) & , & & u_2(2) &= -1 \end{aligned}$$

Saadaan siis 5 ehtoa, mutta toisaalta on 6 parametria:

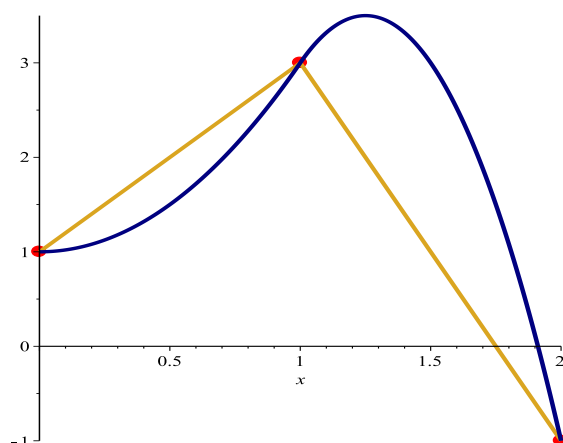
$$\begin{aligned} u_1 &= a_0 + a_1x + a_2x^2 & \text{ja} & & u_2 &= c_0 + c_1x + c_2x^2 & \Rightarrow \\ \begin{cases} a_0 = 1 \\ a_0 + a_1 + a_2 = 3 \\ c_0 + c_1 + c_2 = 3 \\ a_1 + 2a_2 = c_1 + 2c_2 \\ c_0 + 2c_1 + 4c_2 = -1 \end{cases} & \Rightarrow & \begin{cases} a_0 = 1 \\ a_1 = c_2 + 8 \\ a_2 = -c_2 - 6 \\ c_0 = 2c_2 + 7 \\ c_1 = -3c_2 - 4 \end{cases} \end{aligned}$$

Jäi siis yksi vapaa parametri. Jos valitaan esimerkiksi $c_2 = -8$, niin saadaan kuva 4.10.

★

Määritelmä 4.4 on tavallaan tarpeettoman yleinen koska käytännössä lähes aina vaaditaan, että polynomien $\mathcal{S}_j$ asteluku on sama kaikilla j . Määritelmää voikin sitten tarkentaa: edellisessä esimerkissä muodostettu murtoviiva g on *lineaarinen splini*, koska (i) $g \in C(I)$ ja (ii) g on ensimmäisen asteen polynomi rajoitettuna osaväleille. Kaikkein suosituimmiksi ovat osoittautuneet *kuutiosplinit*.

Määritelmä 4.5. *Funktio $\mathcal{S} : I \rightarrow \mathbb{R}$ on kuutiosplini, jos*



Kuva 4.10: Lineaarinen ja neliöllinen splini esimerkin 4.8 tapauksessa.

(i) $\mathcal{S} \in C^2(I)$ ja

(ii) $\mathcal{S}_j$ on kolmannen asteen polynomi kaikilla j

Selvästi ei ole järkevää vaatia kolmannen derivaatan jatkuvuutta koska $\mathcal{S}_j'''$ on vakio kolmannen asteen polynomille. Siis jos kuutiospliniille pätee $\mathcal{S} \in C^3(I)$, niin tällöin itse asiassa $\mathcal{S} \in \mathbb{P}_3$. Interpolointiongelmia kuutiosplineille voidaan muotoilla seuraavasti.

Ongelma 4.7. Olkoon annettu pisteet $p^k = (x_k, y_k)$, $k = 0, \dots, n$, missä $a = x_0 < \dots < x_n = b$. Etsi kuutiosplini $\mathcal{S}$ siten, että $y_k = \mathcal{S}(x_k)$ kaikilla k . *

Seuraavan argumentin perusteella pitäisi olla mahdollista muodostaa tällainen interpoloiva splini. Kolmannen asteen polynomilla on 4 vapaata parametria, joten kokonaisuudessa on $4n$ parametria koska osavälejä on n kappaletta. Sisäpisteitä on $n-1$ joten jatkuvuusehdoista saadaan yhteensä $3(n-1)$ ehtoa. Lisäksi on $n+1$ interpolointiehtoa. Koska

$$4n - 3(n-1) - n - 1 = 2$$

niin voisi olettaa, että kuutiosplini on olemassa, ja lisäksi jää kaksi vapaasti valittavaa parametria. Tämähän ei vielä ole todistus koska periaatteessa ehdothan voisivat olla ristiriitaisia, jolloin ratkaisua ei olisikaan. Osoitetaan nyt, että edellinen päättely on oikea, ja katsotaan miten kuutiosplini voidaan käytännössä laskea.

Suoraviivaisesti edeten saataisiin siis yleisessä tapauksessa $4n-2$ yhtälöä $4n$ tuntemattomalle. Osoittautuu että kun polynomit parametrisoidaan ”sopivasti” päästään paljon helpommalla. Olkoon siis $\mathcal{S}_j$ kuutiosplinin $\mathcal{S}$ rajoittuma välille I_j ja merkitään

$$z_j = \mathcal{S}''(x_j) = \mathcal{S}_j''(x_j) = \mathcal{S}_{j+1}''(x_j)$$

Tavoitteena on osoittaa, että tällaiset luvut z_j ovat olemassa ja näyttää miten ne voidaan laskea. Jos luvut z_j ovat olemassa, niin $\mathcal{S}''$ on lineaarinen splini datajoukolle (x_j, z_j) , siis jatkuva funktio ja voidaan kirjoittaa

$$\mathcal{S}_j''(x) = \frac{1}{h_j} \left(z_{j-1}(x_j - x) + z_j(x - x_{j-1}) \right)$$

missä $h_j = x_j - x_{j-1}$. Ideana on nyt saada lauseke $\mathcal{S}$:lle integroimalla tätä kaksi kertaa. Ensimmäinen integraali antaa

$$\mathcal{S}'_j(x) = \int \mathcal{S}''_j(x) dx = \frac{1}{2h_j} \left(-z_{j-1}(x_j - x)^2 + z_j(x - x_{j-1})^2 + \tilde{a}_j \right)$$

missä $\tilde{a}_j$ on integroimisvakio ja integroimalla uudestaan saadaan

$$\mathcal{S}_j(x) = \frac{1}{6h_j} \left(z_{j-1}(x_j - x)^3 + z_j(x - x_{j-1})^3 + \tilde{a}_j x + \tilde{b}_j \right)$$

Selvästi lineaarinen termi $\tilde{a}_j x + \tilde{b}_j$ voidaan kirjoittaa muodossa $a_j(x_j - x) + b_j(x - x_{j-1})$ joten etsitään kuutiospliniä muodossa

$$\mathcal{S}_j(x) = \frac{1}{6h_j} \left(z_{j-1}(x_j - x)^3 + z_j(x - x_{j-1})^3 + a_j(x_j - x) + b_j(x - x_{j-1}) \right)$$

Pitäisi siis edelleen laskea parametrit a_j , b_j ja z_j . Interpolointiehdosta $\mathcal{S}(x_j) = y_j$ seuraa

$$\begin{cases} \mathcal{S}_j(x_j) = \frac{1}{6}(h_j^2 z_j + b_j) = y_j \\ \mathcal{S}_j(x_{j-1}) = \frac{1}{6}(h_j^2 z_{j-1} + a_j) = y_{j-1} \end{cases} \Rightarrow \begin{cases} a_j = 6y_{j-1} - h_j^2 z_{j-1} \\ b_j = 6y_j - h_j^2 z_j \end{cases} \quad (4.5)$$

Ensimmäisen derivaatan jatkuvuudesta saadaan ehto

$$\mathcal{S}'_j(x_j) = \mathcal{S}'_{j+1}(x_j) \Leftrightarrow \frac{1}{6h_j} (3h_j^2 z_j + b_j - a_j) = \frac{1}{6h_{j+1}} (-3h_{j+1}^2 z_j + b_{j+1} - a_{j+1})$$

Tästä sieventämällä saadaan lopulta

$$\begin{aligned} h_j z_{j-1} + 2(h_j + h_{j+1})z_j + h_{j+1}z_{j+1} &= \tilde{v}_j, \quad 1 \leq j < n \quad \text{missä} \\ \tilde{v}_j &= 6(y_{j+1} - y_j)/h_{j+1} - 6(y_j - y_{j-1})/h_j \end{aligned} \quad (4.6)$$

Tästä saadaankin

Lause 4.3. *Ongelmalla 4.7 on aina ratkaisu. Ratkaisu on yksikäsitteinen, jos kiinnitetään $\mathcal{S}''(x_0)$ ja $\mathcal{S}''(x_n)$.*

Todistus. Valitaan siis z_0 ja z_n mielivaltaisesti ja osoitetaan, että tällöin saadaan yksikäsitteinen ratkaisu. Olkoon $z = (z_1, \dots, z_{n-1})$ ja olkoon $v \in \mathbb{R}^{n-1}$ vektori, jolle pätee

$$\begin{aligned} v_1 &= \tilde{v}_1 - h_1 z_0, \quad v_{n-1} = \tilde{v}_{n-1} - h_n z_n \quad \text{ja} \\ v_j &= \tilde{v}_j \quad \text{kun} \quad 1 < j < n-1 \end{aligned}$$

Määritellään edelleen matriisi

$$A = \begin{pmatrix} 2(h_1 + h_2) & h_2 & 0 & \dots & 0 \\ h_2 & 2(h_2 + h_3) & h_3 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & \dots & 0 & h_{n-1} & 2(h_{n-1} + h_n) \end{pmatrix} \quad (4.7)$$

Näillä merkinnöillä yhtälösystemi (4.6) voidaan kirjoittaa muodossa $Az = v$. Ratkaisun z olemassaolo seuraa Lemmasta 4.10. Tämän jälkeen parametrit a_j ja b_j voidaan laskea yhtälöistä (4.5). $\diamond$

Lemma 4.10. *Yhtälöllä $Az = v$ on yksikäsitteinen ratkaisu.*

Todistus. Nyt ratkaisun olemassaolon voisi todistaa monella tavalla. Käytetään tässä kiintopisteiteraatiota! Olkoon $D = \text{diag}(a_{11}, \dots, a_{nn})$ ja $\hat{A} = A - D$. Tällöin

$$\begin{aligned} Az = v &\Leftrightarrow Dz = -\hat{A}z + v \Leftrightarrow z = -D^{-1}\hat{A}z + D^{-1}v \Rightarrow \\ z^{k+1} &= -D^{-1}\hat{A}z^k + D^{-1}v = Tz^k + D^{-1}v \end{aligned}$$

Selvästi nyt

$$\begin{aligned} \sum_{j=1}^n |t_{ij}| &= \frac{h_i}{h_i + h_{i+1}} + \frac{h_{i+1}}{h_i + h_{i+1}} = 1/2 \quad , \quad 1 < i < n \\ \sum_{j=1}^n |t_{1j}| &= \frac{h_2}{h_1 + h_2} \leq 1/2 \quad , \quad \sum_{j=1}^n |t_{nj}| = \frac{h_{n-1}}{h_{n-1} + h_n} \leq 1/2 \end{aligned}$$

Siispä $\|T\|_\infty = 1/2$, joten T on kontraktio ja siis yhtälöllä $Az = v$ on yksikäsitteinen ratkaisu. $\diamond$

Edellä siis kiinnitettiin ratkaisu antamalla ”reunaehdot” $\mathcal{S}''(x_0)$ ja $\mathcal{S}''(x_n)$. Muunkinlaisia reunaehtoja voidaan käyttää mutta koska tämä useimmiten ei oleellisesti vaikuta lopputulokseen niin ei nyt puututa siihen.

Esimerkki 4.9. Käytetään esimerkin 4.8 dataa. Tällöin $h_1 = h_2 = 1$ ja yhtälöistä (4.6) tulee tässä tapauksessa vain yksi yhtälö:

$$z_0 + 4z_1 + z_2 = -36$$

Valitaan vaikkapa $z_0 = z_2 = -1$, joten $z_1 = -17/2$. Sitten kaavojen (4.5) perusteella

$$\begin{aligned} a_1 &= 6 - z_0 = 7 \quad , & b_1 &= 18 - z_1 = 53/2 \\ a_2 &= 18 - z_1 = 53/2 \quad , & b_2 &= -6 - z_2 = -5 \end{aligned}$$

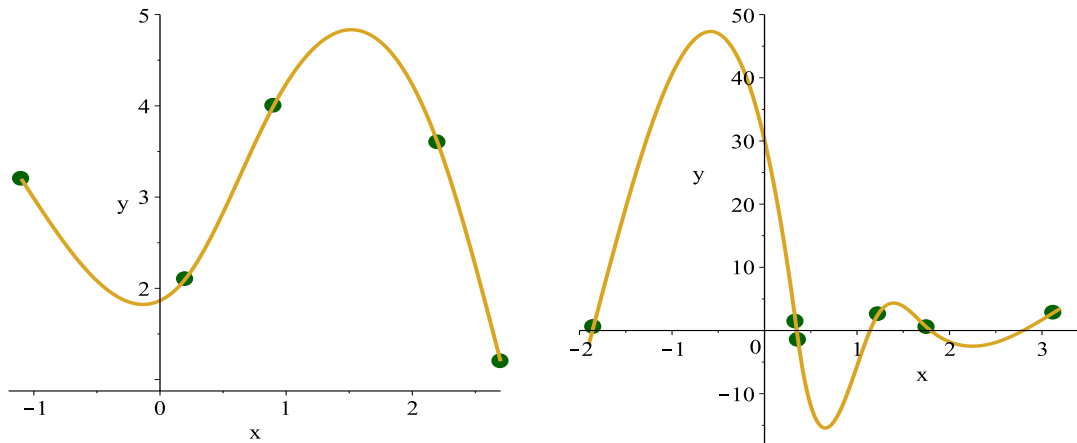
Ratkaisu voidaan siis kirjoittaa muodossa

$$\begin{aligned} \mathcal{S}_1(x) &= \frac{1}{12} \left(-2(1-x)^3 - 17x^3 + 14(1-x) + 53x \right) \\ \mathcal{S}_2(x) &= \frac{1}{12} \left(-17(2-x)^3 - 2(x-1)^3 + 53(2-x) - 10(x-1) \right) \end{aligned}$$

★

Kun siis muodostetaan kuutiosplini, niin pitää ratkaista lineaarinen systeemi (4.7). Tämä on itse asiassa varsin helppoa koska A on *tridiagonaalimatriisi*, toisin sanoen $a_{ij} = 0$, jos $|i-j| > 1$. Tällaisille matriiseille on algoritmeja jotka ovat paljon nopeampia kuin yleisille matriiseille tarkoitetut algoritmit.

Esimerkki 4.10. Käytetään esimerkin 4.7 dataa, ja muodostetaan vastaavat kuutiosplinit. Kuten kuvasta 4.11 nähdään lopputulos on huomattavasti parempi kuin polynomi-interpoloinnin tapauksessa. $\star$



Kuva 4.11: Esimerkin 4.7 dataa vastaavat kuutiosplinit.

Koska kuutiosplini on jokaisella osavälillä 3. asteen polynomi niin on selvää, että sillä voi olla korkeintaan $3n$ nollakohtaa. Itse asiassa nollakohtia on huomattavasti vähemmän.

Lemma 4.11. *Kuutiosplinillä on korkeintaan $n + 2$ nollakohtaa.*

Todistus. Olkoon kuutiosplinillä $\mathcal{S}$ r nollakohtaa. Tällöin Rollen lauseen perusteella derivaatalla $\mathcal{S}'$ on ainakin $r - 1$ nollakohtaa. Mutta koska $\mathcal{S}' \in C^1(I)$, niin myös siihen voi soveltaa Rollen lausetta, joten $\mathcal{S}''$:lla on ainakin $r - 2$ nollakohtaa. Mutta $\mathcal{S}''$ on lineaarinen splini, joten sillä voi olla korkeintaan n nollakohtaa. Siis $r - 2 \leq n$. $\diamond$

Kuutiosplineillä on seuraava mielenkiintoinen ominaisuus.

Lause 4.4. *Olkoon $f \in C^2(I)$ missä $I = [a, b]$ ja $\mathcal{S}$ kuutiosplini joka interpoloi funktiota f solmupisteissä tällä välillä ja oletetaan lisäksi, että $f'(a) = \mathcal{S}'(a)$ ja $f'(b) = \mathcal{S}'(b)$. Tällöin*

$$\|\mathcal{S}''\|_2 \leq \|f''\|_2$$

Todistus. Tämä saadaan osittaisintegroinnilla. Olkoon $a = x_0 < x_1 < \dots < x_n = b$. Lasketaan ensin

$$\int_a^b f'' \mathcal{S}'' dx = \left|_a^b f' \mathcal{S}'' - \int_a^b f' \mathcal{S}''' dx = \left|_a^b \mathcal{S}' \mathcal{S}'' - \int_a^b f' \mathcal{S}''' dx \right.$$

Edelleen

$$\int_{x_k}^{x_{k+1}} f' \mathcal{S}''' dx = \left|_{x_k}^{x_{k+1}} f \mathcal{S}''' - \int_{x_k}^{x_{k+1}} f \mathcal{S}'''' dx = \left|_{x_k}^{x_{k+1}} \mathcal{S} \mathcal{S}''' \right.$$

Siispä

$$\begin{aligned}\int_a^b f'' S'' dx &= \left|_a^b f' S'' - \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f' S''' dx \right|_a^b = \left|_a^b S' S'' - \sum_{k=0}^{n-1} \left|_{x_k}^{x_{k+1}} S S''' \right|_{x_k} \right|_a^b \\ &= \left|_a^b S' S'' - \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} S' S''' dx \right|_a^b = \int_a^b (S'')^2 dx\end{aligned}$$

Sisätulon avulla ilmaistuna

$$\langle f'' - S'', S'' \rangle = 0$$

Mutta tällöin saadaan Pythagoraan lauseen avulla täsmälleen samoin kuin Lemmassa 4.6

$$\|f''\|_2^2 = \|f'' - S'' + S''\|_2^2 = \|f'' - S''\|_2^2 + \|S''\|_2^2 \geq \|S''\|_2^2$$

◇

Tämä tulos johtaakin sitten splini-sanankuuperäiseen merkitykseen.

Määritelmä 4.6. *Funktion f graafin kaarevuus on*

$$\kappa(x) = \frac{f''(x)}{(1 + f'(x)^2)^{3/2}}$$

Siis jos f' on ”pieni”, niin kaarevuus on lähes verrannollinen toiseen derivaattaan, joten lauseen voi tulkita siten, että kuutiosplini pyrkii löytämään approksimaation jonka kaarevuus on ainakin lähellä minimiä. Jos sitten tarkastellaan ohutta ja taipuisaa keppiä niin voidaan osoittaa, että kepin jännityksen potentiaalienergia on verrannollinen kaarevuuden neliöön, siis

$$E \sim \|\kappa\|_2^2 = \int_a^b \kappa^2 dx$$

Kun tällainen keppi pakotetaan kulkemaan tiettyjen pisteitten kautta se pyrkii minimoimaan potentiaalienergian, ja siis asettumaan siten, että kaarevuus on mahdollisimman pieni. Tällaisia taipuisia viivottimia, joita siis kutsuttiin splineiksi (engl. spline) käytettiin aikoinaan, ja ehkäpä osittain vieläkin, ainakin laiva- ja sittemmin lentokone-teollisuudessa. Kuvasta 4.12 pitäisi selvitä idea.

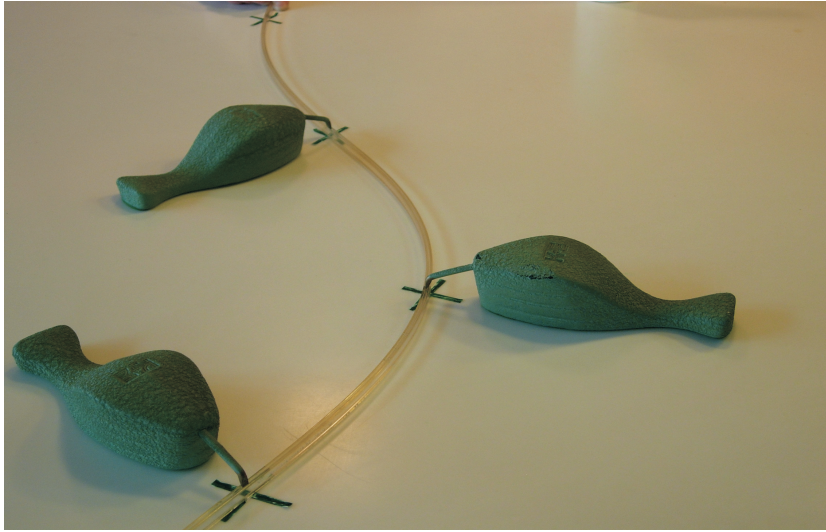
4.4 Yleisen tasokäyrän interpolointi

Usein esiintyy tilanteita joissa käyriä ei voi esittää jonkin funktion graafina muodossa $y = f(x)$. Tämäkin tilanne on helppo käsitellä.

Määritelmä 4.7. *Olkoon $I \subset \mathbb{R}$ jokin reaaliakselin väli.*

(i) *Käyrä on kuvaus $c : I \rightarrow \mathbb{R}^2$.*

(ii) *Käyrän tangentti pisteessä $p = c(t_0)$ on $c'(t_0)$.*



Kuva 4.12: ”Alkuperäinen” splini.

(iii) Käyrä on säännöllinen, jos $c'(t) \neq 0$ kaikilla t .

Toisin sanoen säännöllisille käyrille tangentti on aina hyvin määritelty. Käyristä puhuttaessa usein aiheuttaa sekaannusta, että usein käyrää ajatellaan kuvauksen c kuvajoukkona. Siis esimerkiksi yksikköympyrä $S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ on tason yksiulotteinen osajoukko joka on käyrän $c(t) = (\cos(t), \sin(t))$ kuvajoukko.

Olko sitten annettu pisteet $p^k = (x_k, y_k)$, $k = 0, \dots, n$ ja haluttaisiin muodostaa käyrä $c : [0, 1] \rightarrow \mathbb{R}^2$ joka kulki kaikkien pisteitten kautta. Tämä voidaan varsin helposti toteuttaa sekä polynomien että splinien avulla.

Valitaan ensin jotkin t_k siten, että

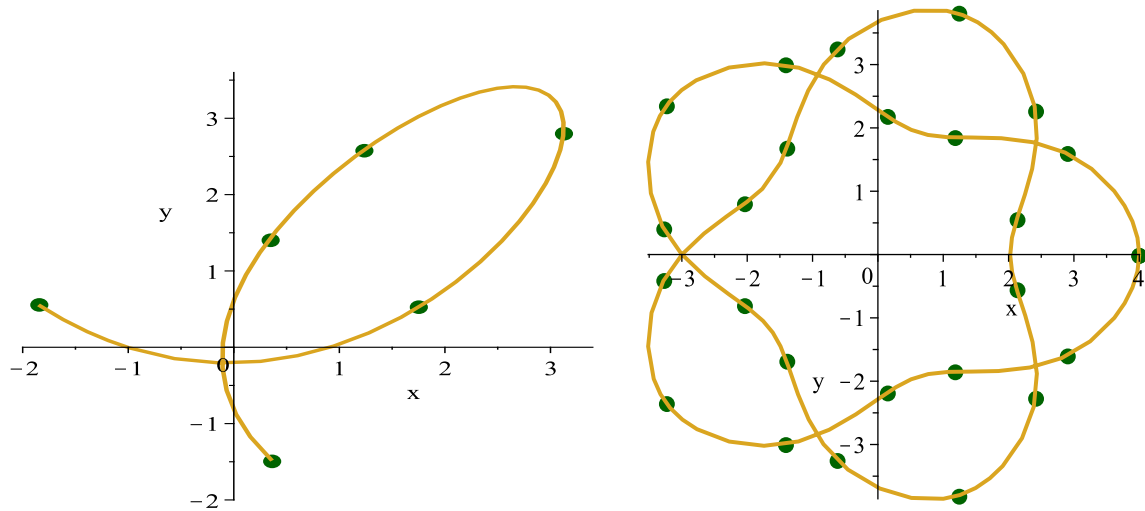
$$0 = t_0 < t_1 < \dots < t_n = 1$$

Muodostetaan sitten näitten avulla Lagrangen polynomit $\mathcal{L}_k^n$ tavalliseen tapaan (Määritelmä 4.3). Tämän jälkeen asetetaan

$$c(t) = \sum_{k=0}^n p^k \mathcal{L}_k^n(t)$$

Tämä antaa interpoloivan käyrän. Kuitenkin jos pisteitä on paljon, niin tämä konstruktio on huono Rungen ilmiön takia. Käytännössä siis haluttaisiin interpoloiva splini. Tämäkin voidaan muodostaa varsin suoraviivaisesti. Valitaan solmupisteet t_k kuten edellä. Sitten

- (i) muodostetaan (kuutio)splini $\mathcal{S}_x$ datapisteille $\{(t_k, x_k)\}$
- (ii) muodostetaan (kuutio)splini $\mathcal{S}_y$ datapisteille $\{(t_k, y_k)\}$
- (iii) asetetaan $c(t) = (\mathcal{S}_x(t), \mathcal{S}_y(t))$



Kuva 4.13: Esimerkin 4.11 pisteet ja interpoloivat käyrät.

Huomaa, että ei ole mitään ainoaa oikeata tapaa valita pisteitä t_k , ja erilaiset valinnat voivat vaikuttaa lopulliseen käyrään huomattavastikin. Luonnollisesti tuntuisi järkevältä, että jos kahden peräkkäisen pisteen etäisyys on ”iso”, niin vastaavasti tätä pitäisi vastata ”iso”parametriväli. Mutta asia ei ole näin yksinkertainen koska paljon riippuu siitä miten pisteet sijaitsevat toisiinsa nähden.

Esimerkki 4.11. Kuvan 4.13 tapauksessa vasemmalla on annettu 6 pistettä. Tässä tuntuisi ehkä luonnollisemmalta etsiä parametrisoitua käyrää. Kuvassa on valittu parametrin t arvot

$$t = (0, 0.2, 0.4, 0.7, 0.8, 1)$$

Kokeilemalla huomaa, että jos valitsee parametrin t_i eri tavalla, saattaa saada varsin erinäköisiä käyriä. Kuvassa oikealla on valittu t_k tasavälein. Koska käyrä on umpinainen, niin nyt voidaan splini $c : [0, 1] \rightarrow \mathbb{R}^2$ tulkita jaksolliseksi käyräksi. Reunaehdoiksi on siis mielekästä valita $c'(0) = c'(1)$ ja $c''(0) = c''(1)$. ★

4.5 Bézier-käyrät

Palataan sitten lopuksi vielä Bernsteinin polynomeihin. Jos oli annettu jatkuva funktio $f : [0, 1] \rightarrow \mathbb{R}$, niin tätä vastaava Bernsteinin approksimaatio saatiin muodostamalla

$$f_n(x) = \sum_{k=0}^n f(k/n) B_k^n(x)$$

Merkitään nyt $x_k = k/n$ ja $y_k = f(k/n)$. Tällöin siis Bernsteinin approksimaatio vastaa käyrää

$$c(t) = (t, f_n(t))$$

Mutta nyt Lemman 4.3 mukaan

$$t = \sum_{k=0}^n \frac{k}{n} B_k^n(t) = \sum_{k=0}^n x_k B_k^n(t)$$

joten

$$c(t) = \sum_{k=0}^n \begin{pmatrix} x_k \\ y_k \end{pmatrix} B_k^n(t)$$

Olkoon sitten annettu mielivaltaiset $p^k = (x_k, y_k)$, $k = 0, \dots, n$.

Määritelmä 4.8. *Pisteitten p^k määrittämä Bézier-käyrä on*

$$c(t) = \sum_{k=0}^n p^k B_k^n(t)$$

Toisaalta pisteitä p^k sanotaan käyrän c kontrollipisteiksi (tai Bézier-pisteiksi).

Huomaa, että $c(0) = p^0$ ja $c(1) = p^n$, mutta c ei interpoloi muita pisteitä p^k .

Lause 4.5. *Olkoon $\text{conv}(S)$ joukon S konvekksi peite (Määritelmä A.7). Tällöin*

$$c(t) \in \text{conv}(S) \quad \text{kaikilla} \quad 0 \leq t \leq 1$$

Todistus. Koska $B_k^n(t) \geq 0$ ja $\sum_{k=0}^n B_k^n(t) = 1$, niin $c(t)$ on konvekssi kombinaatio pisteistä $p^k = (x_k, y_k)$ joten lause seuraa Lemmasta A.4. $\diamond$

Siis Bézier-käyrä pysyy kontrollipisteitten määrittämässä konveksissa peitteessä, kuten kuvasta 4.14 nähdään. Bernsteinin polynomit ovat hyviä Bézier-käyrien teoreettisiin tarkasteluihin. Laskennalliselta kannalta seuraava De Casteljaun algoritmi on kuitenkin kätevämpi.

Algoritmi 7 DE CASTELJAUN ALGORITMI

Input: kontrollipisteet $p^0, p^1, \dots, p^n$

Output: kontrollipisteitten määrittämä Bézier-käyrä $c = c_0^n$

for $j = 0$ to n **do**

$$c_j^0(t) := p^j$$

end for

for $\ell = 1$ to n **do**

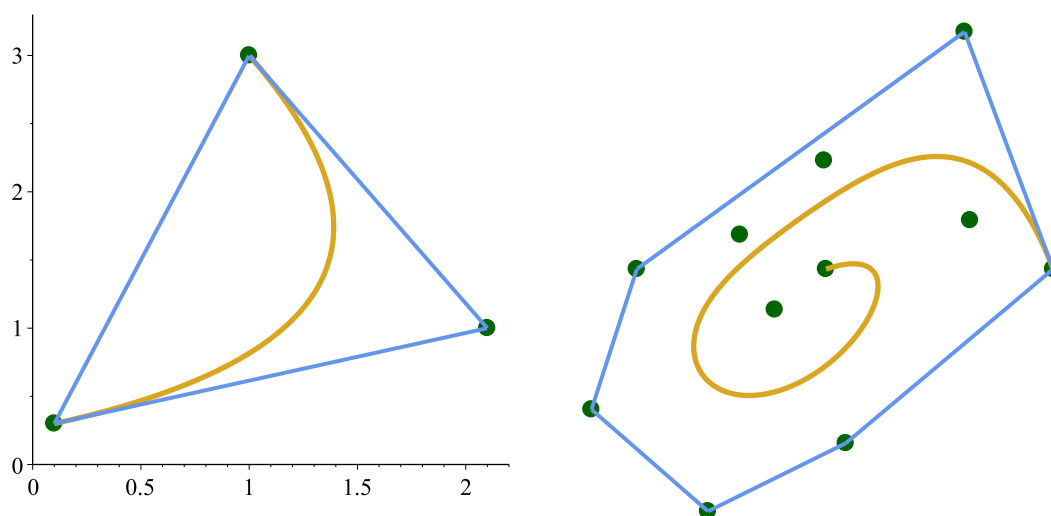
for $j = 1$ to $n - \ell$ **do**

$$c_j^\ell(t) := (1 - t)c_j^{\ell-1}(t) + tc_{j+1}^{\ell-1}(t)$$

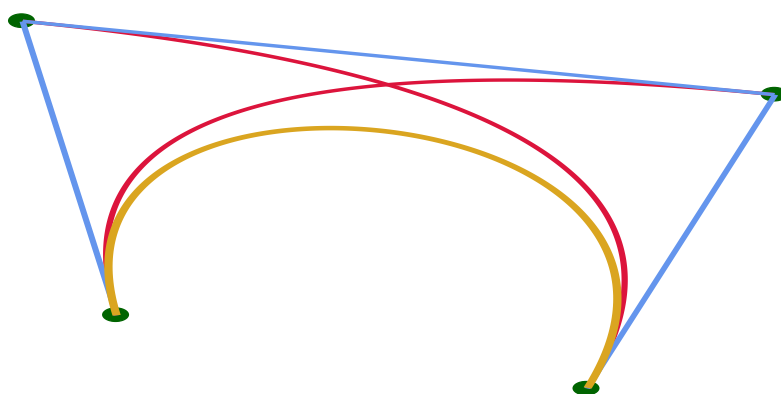
end for

end for

Algoritmissa siis otetaan konvekseja kombinaatioita alempiasteisista käyristä, jolloin pysytään konveksissa peitteessä. Kuva 4.15 selventää asiaa. Huomaa algoritmin askel $c_j^\ell(t) := (1 - t)c_j^{\ell-1}(t) + tc_{j+1}^{\ell-1}(t)$ ja vertaa sitä Lemman A.3 kohtaan 4.



Kuva 4.14: Vasemmalla kolmen pisteen määrittämä Bézier-käyrä, oikealla 11 pistettä. Sinisellä on piirretty pisteitten määrittämän konveksin peitteen reuna.



Kuva 4.15: De Casteljaun algoritmin eteneminen. Ensimmäisessä vaiheessa muodostetaan siniset suorat, sitten punaiset toisen asteen käyrät, ja lopuksi kaikkien pisteitten määrittämä Bézier-käyrä.



Kuva 4.16: Erityyppisiä jousia.

Bézier-käyrät ovat osoittautuneet käteviksi tietokoneavusteisessa suunnittelussa. Esimerkiksi kuvassa 4.16 on erityyppisiä jousia, ja haluttaisiin löytää sopiva parametrisointi käyrälle joka määrittää jousen muodon. Luonnollisesti tässä voisi käyttää splinejä. Mutta jos sitten haluttaisiin kokeilla eri muotoja ja/tai muuttaa alkuperäistä muotoa, niin Bézier-käyrät ovat kätevämpiä. Erityisesti silloin kun halutaan suunnitella jokin täysin uusi kappale tai muoto, jolloin datapisteitä ei alussa ole ollenkaan annettu, niin Bézier-käyrät (ja niitten yleistykset) ovat parempia kuin klassiset splinit. Samoja tekniikoita käytetään nykyään myös animaatioelokuvien tekemisessä. Bézier-käyriä on perusteellisesti tarkasteltu kirjassa [7].

Bézier ja de Casteljau keksivät toisistaan riippumatta Bézier-käyrät 60-luvun alussa. Bézier oli Renault'n ja de Casteljau Citroënin insinööri ja molempien ongelmana oli parantaa ja nopeuttaa auton muodon suunnittelua. Heidän ratkaisunsa näihin ongelmiin johti erilaisten CAD (computer aided design) ja CAM (computer aided manufacturing) järjestelmien toteuttamiseen.

4.6 Harjoitustehtäviä

1. Näytä, että Lemmassa 4.7 saatu ratkaisu on yksikäsitteinen.
2. Olkoon annettu pisteet $p^0 = (-1, 0)$, $p^1 = (1, 4)$ ja $p^2 = (2, 1)$. Laske näitä vastaava Bézier-käyrä.