Funktion approksimointi yksinkertaisempien funktioiden avulla

Pyry Ekholm

LuK-tutkielma
Elokuu 2024

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Sisällys

1 Johdanto

Monet matemaattiset ongelmat luonnontieteiden ja tekniikan aloilla sisältävät jonkin funktion f approksimoinnin toisella yksinkertaisemmalla funktiolla g. Tällaisia ongelmia ratkottaessa halutaan, että approksimaatiofunktio g kuvaa funktiota f mahdollisimman tarkasti. Yksi tapa approksimaation tarkkuuden arviointiin on tarkastella funktion ja sen approksimaatiofunktion väliin jäävää pinta-alaa. Mitä pienempi on funktioiden väliin jäävä pinta-ala, sitä tarkempi on approksimaatio. Matemaattisesti välillä $[a, b]$ kahden käyrän väliin jäävä pinta-ala lasketaan integraalilla $A = \int_a^b |f(x) - g(x)|\ dx$.

Integraalin muodossa ilmaistuna approksimaatio on tarkka silloin, kun integraalin arvo on lähellä nollaa. Koska integraalit, joiden integrandissa on itseisarvolauseke, ovat usein haastavasti laskettavia, korvataan integrandissa oleva erotuksen itseisarvo erotuksen neliöllä, jolloin integraali muuttuu muotoon $A = \int_a^b [f(x) - g(x)]^2\ dx$. Tämän muutoksen tekemiseen perustuu myös niin kutsuttu pienimmän neliön approksimaatio, johon tässä tutkielmassa perehdytään tarkemmin.

Tutkielman rakenne seuraa pitkälti Larsonin ja Falvon kirjaa [1], mutta tekstin tueksi on haettu tuloksia myös muista lähteistä. Tutkielman toisessa luvussa käydään läpi pienimmän neliön approksimaation määritelmä, jonka lisäksi esitellään Hessen matriisi ja Sylvesterin kriteeri, joiden avulla voidaan etsiä kahden tai useamman muuttujan funktioiden ääriarvokohdat. Nämä ovat pienimmän neliön approksimoinnin kannalta oleellisia tuloksia, sillä approksimaatiota tehdessä täytyy kyetä etsimään funktion minimikohta.

Tutkielman kolmannessa luvussa esitellään toinen tapa pienimmän neliön approksimaation määrittämiseen. Approksimaatio voidaan suorittaa myös vektorimuodossa, kun apuna käytetään integraalin avulla määriteltävää sisätuloa. Tämä vaatii sen, että kyetään muodostamaan sisätuloavaruuden ortonormaalikanta, jota varten käytetään Gramin-Schimidtin ortonormalisointiprosessia. Tämän lisäksi kolmannessa luvussa perehdytään pienimmän neliön approksimaation sovellukseen Fourier'n approksimaatioon. Tutkielman tavoitteena on sekä esitellä edellä mainittujen menetelmien teoreettista pohjaa että havainnollistaa menetelmien käyttöä esimerkkien avulla.

2 Pienimmän neliön approksimaatio

Tässä luvussa käsitellään pienimmän neliön approksimoinnin perusteita teorian ja esimerkkien avulla. Ensin esitellään pienimmän neliön approksimoinnin määritelmä, jonka jälkeen käydään läpi kahden tai useamman muuttujan funktion ääriarvokohtien määrittäminen. Ääriarvokohtia tai tarkemmin sanottuna funktion minimikohtaa tarvitaan approksimoinnin aikana. Luvun lopussa käytetään esiteltyjä menetelmiä ja määritetään pienimmän neliön approksimaatiofunktiot kahdessa eri tilanteessa.

2.1 Pienimmän neliön approksimaation määritelmä

Perehdytään ensin pienimmän neliön approksimaation määritelmään, joka kertoo, milloin valitun funktiojoukon funktio g on funktion f pienimmän neliön approksimaatiofunktio.

Määritelmä 1. Olkoon funktio f jatkuva välillä $[a, b]$ ja olkoon W kaikkien välillä $[a, b]$ jatkuvien funktioiden sisätuloavaruuden $C[a, b]$ aliavaruus. Funktio g on funktion f pienimmän neliön approksimaatio aliavaruudessa W, jos integraalin

$$I = \int_a^b [f(x) - g(x)]^2 \, dx$$

arvo on pienempi tai yhtäsuuri kuin kaikilla muilla aliavaruuden W funktioilla.

Määritelmän 1 mukaan siis tulee tutkia, milloin integraali I saa pienimmän arvonsa. Voidaan tulkita, että I on funktio funktion g valinnan (yleensä sopivien vakioiden) suhteen ja määrittää niille arvot, joilla funktio I saa pienimmän arvonsa. Toisin sanoen määritetään funktion I lokaalit ääriarvokohdat, mihin tarvitaan lemman 1 tulosta.

Lemma 1. *([2]). Jos funktio f on osittaisderivaattoineen jatkuva joukossa A ja sillä on lokaali ääriarvo pisteessä $\mathbf{a} \in A^\circ$, niin $\nabla f(\mathbf{a}) = \mathbf{0}$.*

Lemman 1 tuloksen mukaan pisteet, joissa funktion gradientti häviää, ovat funktion ääriarvokohtakandidaatteja. Ääriarvokohdan luonne eli se, onko kyseisessä kohdassa lokaali minimi, lokaali maksimi vai satulapiste, saadaan selville Hessen matriisin avulla. Satulapisteellä tarkoitetaan pistettä, joka ei ole lokaali ääriarvokohta.

2.1.1 Hessen matriisi ja Sylvesterin kriteeri

Yhden muuttujan funktion tapauksessa derivaatan nollakohdan ääriarvoluonne selviää tutkimalla funktion toista derivaattaa. Seuraavaksi esiteltävien Hessen matriisin ja Sylvesterin kriteerin avulla voidaan selvittää ääriarvokohtakandidaatin luonne kahden tai useamman muuttujan funktioille.

Määritelmä 2. Funktion f $n \times n$ -kokoinen Hessen matriisi pisteessä $\mathbf{a}$ on

$$H(f, \mathbf{a}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{a}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{a}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{a}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{a}) \end{pmatrix}.$$

Lähteen [2] mukaan kaikkien funktioiden f toisen kertaluvun osittaisderivaatat yhtyvät, eli kaikille muuttujille i ja j sekä kaikille pisteille $\mathbf{a}$ pätee, että $\partial_{ij} f(\mathbf{a}) = \partial_{ji} f(\mathbf{a})$. Tämän perusteella voidaan todeta, että Hessen matriisi on symmetrinen, jonka myötä Sylvesterin kriteeriä voidaan käyttää työkaluna ääriarvokohdan luonnetta tarkastaessa.

Lause 1. *([2]). (Sylvesterin kriteeri) Olkoon A symmetrinen $n \times n$ -matriisi, ja Q sitä vastaava neliömuoto. Olkoon kaikilla $i = 1, 2, \ldots, n$ luku Δ_i matriisin vasemman $i \times i$ ylänurkan determinantti. Tällöin Q on positiividefiniitti silloin ja vain silloin kun kaikki luvut Δ_i ovat positiivisia. Vastaavasti Q on negatiividefiniitti silloin ja vain silloin kun kaikki luvut $(-1)^i \Delta_i$ ovat positiivisia.*

Todistus. Sivuutetaan. Todistus nähtävissä esimerkiksi kirjassa [3]. □

Lähteen [2] mukaan Sylvesterin kriteeristä seuraa, että neliömuoto on positiividefiniitti, jos ja vain jos sen alideterminanttien jono on $+,+,+,\ldots$. Vastaavasti muoto on negatiividefiniitti, jos ja vain jos sen jono on $-,+,-,+,\ldots$, ja muulloin muoto on indefiniitti.

2.1.2 Ääriarvokohdan luonteen määrittäminen

Jokaiselle vähintään kahden muuttujan funktiolle voidaan muodostaa Hessen matriisi, jonka definiittisyyttä tarkastellaan Sylvesterin kriteerillä. Seuraava lause kertoo, mitä ääriarvokohdasta voidaan päätellä Sylvesterin kriteerin tuloksen perusteella.

Lause 2. *([2]) Oletetaan, että funktion f, jossa on n kappaletta muuttujia, kaikki toisen kertaluvun osittaisderivaatat ovat jatkuvia joukossa A ja että $\mathbf{a} \in A^\circ$ on funktion f lokaali ääriarvokohta. Tällöin*

- *Jos Hessen matriisi $H(f,\mathbf{a})$ on positiividefiniitti, niin piste $\mathbf{a}$ on lokaali minimikohta.*
- *Jos Hessen matriisi $H(f,\mathbf{a})$ on negatiividefiniitti, niin piste $\mathbf{a}$ on lokaali maksimikohta.*
- *Jos Hessen matriisi $H(f,\mathbf{a})$ on indefiniitti, niin piste $\mathbf{a}$ on satulapiste.*

Todistus. Tarkka todistus sivuutetaan, mutta se löytyy lähteestä [2]. Todetaan kuitenkin, että Hessen matriisin arvot ovat suunnattuja toisia derivaattoja, josta seuraa, että jos Hessen matriisin arvot ovat kaikki positiivisia, niin kyseessä on lokaali minimikohta. Vastaavasti jos kaikki arvot ovat negatiivisia, niin kyseessä on lokaali maksimi. □

Nämä ovat tarvittavat menetelmät, jotta voidaan minimoida määritelmässä 1 esiintyvä integraali $I = \int_0^1 [f(x) - g(x)]^2 \, dx$ ja varmistua siitä, että funktio g on funktion f pienimmän neliön approksimaatiofunktio avaruudessa W.

2.2 Esimerkkejä pienimmän neliön approksimoinnista

Tutkitaan pienimmän neliön approksimaation käyttöä kahden esimerkin avulla. Esimerkit esitellään kirjassa [1], mutta niihin on lisätty välivaiheita. Tarkoituksena on selvittää, millaiset approksimaatiofunktiot saadaan funktiolle $f(x) = e^x$, kun haetaan approksimaatioksi ensin suora ja sitten toisen asteen polynomi.

Esimerkki 1. Etsitään lineaarinen, eli muotoa $g(x) = a_0 + a_1 x$ oleva pienimmän neliön approksimaatio funktiolle $f(x) = e^x$, kun $x \in [0,1]$.

Määritelmän 1 mukaan funktio g on funktion f pienimmän neliön approksimaatio silloin kun vakiot a_0 ja a_1 ovat sellaiset, että integraali $I = \int_0^1 [f(x) - g(x)]^2 \, dx$

saa pienimmän arvonsa. Etsitään siis tällaiset vakioiden a_0 ja a_1 arvot. Aloitetaan laskemalla integraali, jonka arvoksi saadaan

$$\begin{aligned}
I &= \int_0^1 (e^x - a_0 - a_1 x)^2 \, dx \\
&= \left[\frac{1}{2} e^{2x} - 2a_0 e^x - 2a_1 e^x (x-1) + a_0^2 x + a_0 a_1 x^2 + a_1^2 \frac{x^3}{3} \right]_0^1 \\
&= \frac{1}{2}(e^2 - 1) - 2a_0(e-1) - 2a_1 + a_0^2 + a_0 a_1 + \frac{1}{3} a_1^2.
\end{aligned}$$

Tulkitaan nyt, että I on funktio muuttujien a_0 ja a_1 suhteen ja etsitään sen minimikohta. Lemman 1 mukaan funktiolla on lokaali ääriarvo pisteessä $\mathbf{a}$, jos $\nabla f(\mathbf{a}) = 0$. Funktion gradientti on nolla silloin kun funktion kaikki osittaisderivaatat ovat nollia. Funktion I osittaisderivaatoiksi saadaan

$$\frac{\partial I}{\partial a_0} = -2e + 2 + 2a_0 + a_1$$

ja

$$\frac{\partial I}{\partial a_1} = -2 + a_0 + \frac{2}{3} a_1.$$

Kun molempien osittaisderivaattojen arvoksi asetetaan nolla, saadaan yhtälöpari

$$\begin{cases} 2a_0 + a_1 = 2e - 2 \\ 3a_0 + 2a_1 = 6 \end{cases}$$

Yhtälöparilla on vain yksi ratkaisu, joten funktion I ainoa mahdollinen ääriarvokohta on piste $\mathbf{a} = (4e - 10, -6e + 18)$. Tutkitaan Hessen matriisia ja Sylvesterin kriteeriä käyttäen, onko piste $\mathbf{a}$ funktion I ääriarvokohta. Hessen matriisin muodostamiseen tarvittavat osittaisderivaatat ovat

$$\begin{aligned}
\frac{\partial^2 I}{\partial a_0^2} &= \frac{\partial \partial I}{\partial a_0 \partial a_0} = \frac{\partial(-2e + 2 + 2a_0 + a_1)}{\partial a_0} = 2 \\
\frac{\partial^2 I}{\partial a_0 \partial a_1} &= \frac{\partial^2 I}{\partial a_1 \partial a_0} = \frac{\partial(-2e + 2 + 2a_0 + a_1)}{\partial a_1} = 1
\end{aligned}$$

ja

$$\frac{\partial^2 I}{\partial a_1^2} = \frac{\partial(-2 + a_0 + \frac{2}{3} a_1)}{\partial a_1} = \frac{2}{3}.$$

Näin ollen funktion I Hessen matriisi on

$$H(I, \mathbf{a}) = \begin{pmatrix} 2 & 1 \\ 1 & \frac{2}{3} \end{pmatrix}.$$

Selvitetään Sylvesterin kriteeriä varten tarvittavat luvut Δ_1 ja Δ_2 laskemalla Hessen matriisin vasemman $i \times i$ ylänurkan determinantit, kun $i = 1, 2$. Luvuiksi Δ_1 ja Δ_2 saadaan $\Delta_1 = 2$ ja

$$\Delta_2 = \begin{vmatrix} 2 & 1 \\ 1 & \frac{2}{3} \end{vmatrix} = 2 \cdot \frac{2}{3} - 1 \cdot 1 = \frac{4}{3} - 1 = \frac{1}{3}.$$

Luvut Δ_1 ja Δ_2 ovat positiivisia, joten Sylvesterin kriteerin mukaan Hessen matriisi on positiividefiniitti. Näin ollen lauseen 2 tuloksen mukaan piste $\mathbf{a} = (4e-10, 18-6e)$ on funktion I minimikohta, joten funktion $f(x) = e^x$ lineaarinen pienimmän neliön approksimaatio välillä $x \in [0, 1]$ on funktio

$$g(x) = (4e - 10) + (18 - 6e)x \approx 0,873 + 1,690x.$$

Lähteen [1] mukaan approksimaation tulos on riippuvainen käytetystä approksimaatiomenetelmästä, esimerkiksi Taylorin polynomeja käyttämällä saadaan ratkaisuksi eri funktio kuin esimerkissä 1. Lisäksi samalla approksimaatiomenetelmällä voidaan saada erilaisia ratkaisuja riippuen siitä, että minkälaista funktiota haetaan. Esimerkissä 1 haettiin ratkaisuksi pienimmän neliön approksimaatiolla lineaarista funktiota eli ensimmäisen asteen polynomia. Tutkitaan, miten ratkaisu muuttuu, kun haetaan samalle funktiolle $f = e^x$ approksimaatiofunktioksi toisen asteen polynomi. Funktion f ja sen approksimaatioiden kuvaajat esitetään kuvassa 1.

Esimerkki 2. Etsitään se toisen asteen polynomi, joka on funktion $f(x) = e^x$ pienimmän neliön approksimaatio, kun $x \in [0, 1]$.

Määritelmän 1 mukaan toiseen asteen polynomi $g(x) = a_0 + a_1x + a_2x^2$ on funktion f pienimmän neliön approksimaatio silloin kun vakiot a_0, a_1 ja a_2 ovat sellaiset, että integraali $I = \int_0^1 [f(x) - g(x)]^2\, dx$ saa pienimmän arvonsa. Etsitään tällaiset vakioiden a_0, a_1 ja a_2 arvot. Lasketaan integraali I, jonka arvoksi saadaan

$$\begin{aligned}
I &= \int_0^1 (e^x - a_0 - a_1x - a_2x^2)^2\, dx \\
&= \frac{e^2}{2} - \frac{1}{2} + 2a_0 - 2a_0e + 4a_2 - 2a_2e + a_0^2 + a_0a_1 + \frac{2a_0a_2}{3} + \frac{a_1a_2}{2} + \frac{a_1^2}{3} + \frac{a_2^2}{5} - 2a_1.
\end{aligned}$$

Tulkitaan, että I on funktio muuttujien a_0, a_1 ja a_2 suhteen ja etsitään sen minimikohta. Lemman 1 mukaan funktiolla on lokaali ääriarvo pisteessä $\mathbf{a}$, jos funktion gradientti on nolla, mikä tapahtuu silloin kun funktion kaikki osittaisderivaatat ovat nollia. Funktion I osittaisderivaatoiksi saadaan

$$\begin{aligned}
\frac{\partial I}{\partial a_0} &= 2 - 2e + 2a_0 + a_1 + \frac{2}{3}a_2, \\
\frac{\partial I}{\partial a_1} &= a_0 + \frac{1}{2}a_2 + \frac{2}{3}a_1 - 2
\end{aligned}$$

ja

$$\frac{\partial I}{\partial a_2} = 4 - 2e + \frac{2}{3}a_0 + \frac{1}{2}a_1 + \frac{2}{5}a_2.$$

Kun jokaisen osittaisderivaatan arvoksi asetetaan nolla, saadaan yhtälöryhmä

$$\begin{cases}
6a_0 + 3a_1 + 2a_2 = 6(e - 1) \\
6a_0 + 4a_1 + 3a_2 = 12 \\
20a_0 + 15a_1 + 12a_2 = 60(e - 2).
\end{cases}$$

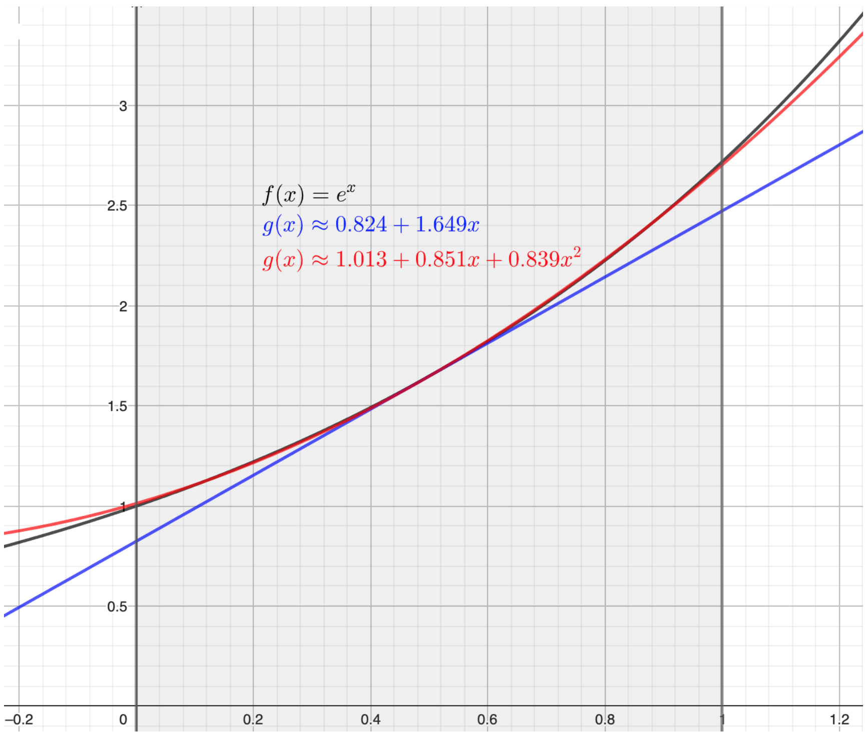
Yhtälöryhmällä on vain yksi ratkaisu, joten funktion I ainoa ääriarvokohtakandidaatti on piste $\mathbf{a} = (39e - 105, -216e + 588, 210e - 570)$. Tarkastetaan Hessen matriisin ja Sylvesterin kriteerin avulla, onko piste $\mathbf{a}$ funktion I ääriarvokohta. Funktion I Hessen matriisi on

$$H(I, \mathbf{a}) = \begin{pmatrix} 2 & 1 & \frac{2}{3} \\ 1 & \frac{2}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} & \frac{2}{5} \end{pmatrix}.$$

Selvitetään Sylvesterin kriteerin tulkinnassa tarvittavat luvut Δ_1, Δ_2 ja Δ_3 laskemalla Hessen matriisin vasemman $i \times i$ ylänurkan determinantit, kun $i = 1, 2, 3$. Luvuiksi Δ_1, Δ_2 ja Δ_3 saadaan $\Delta_1 = 2$, $\Delta_2 = \frac{1}{3}$ ja $\Delta_3 = \frac{1}{270}$.

Koska luvut Δ_1, Δ_2 ja Δ_3 ovat positiivisia, niin Sylvesterin kriteerin mukaan funktion I Hessen matriisi on positiividefiniitti. Tällöin lauseen 2 tuloksen mukaan piste $\mathbf{a} = (39e - 105, -216e + 588, 210e - 570)$ on funktion I minimikohta. Funktion $f(x) = e^x$ pienimmän neliön approksimaatiofunktio välillä $x \in [0, 1]$ on siis

$$g(x) = (39e - 105) + (588 - 216e)x + (210e - 570)x^2 \approx 1.013 + 0.851x + 0.839x^2.$$



Kuva 1: Funktion $f(x) = e^x$ ja sen approksimaatioiden kuvaajat yli välin $x \in [0, 1]$.

3 Sovelluksia

Tässä luvussa esitellään ensin toinen tapa pienimmän neliön approksimaation etsimiseen, kun approksimaatio suoritetaan vektorimuodossa sopivan integraalin määrittämän sisätulon avulla. Tämän jälkeen tutustutaan pienimmän neliön approksimaation sovellukseen Fourier'n approksimaatioon, joka myös perustuu approksimaation suorittamiseen vektorimuodossa. Molempia aiheita käsitellään sekä teorian että esimerkkien avulla.

3.1 Pienimmän neliön approksimaatio sisätulon avulla

Pienimmän neliön approksimaation määritelmän mukainen integraali I voidaan ilmaista myös vektorimuodossa. Tätä varten käytetään tunnettua (kts. [4]) integraalin antamaa sisätuloa funktioavaruudessa $C[a, b]$

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

jolloin integraali I saadaan muotoon

$$I = \int_a^b [f(x) - g(x)]^2\ dx = (f - g, f - g) = ||f - g||^2.$$

Tämä tarkoittaa, että approksimaatiofunktio g on se funktio, jolla normi $||f - g||$ saa pienimmän arvonsa. Toisin sanoen funktion f pienimmän neliön approksimaatio on valitusta funktiojoukosta se funktio g, joka on lähimpänä funktiota f sisätulon (f, g) suhteen. Seuraava lause tarjoaa tavan funktion g määrittämiseen.

Lause 3. *([1]) Olkoon funktio f jatkuva välillä $[a, b]$ ja olkoon W avaruuden $C[a, b]$ aliavaruus. Funktion f pienimmän neliön approksimaatio avaruudessa W on*

$$g = (f, \mathbf{w}_1)\mathbf{w}_1 + (f, \mathbf{w}_2)\mathbf{w}_2 + \cdots + (f, \mathbf{w}_n)\mathbf{w}_n,$$

missä $B = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ on aliavaruuden W ortonormaalikanta .

Todistus. Osoitetaan, että funktio g on funktion f pienimmän neliön approksimaatio osoittamalla, että $||f - g|| \leq ||f - \mathbf{w}||$ kaikille aliavaruuden W vektoreille $\mathbf{w}$. Ilmaistaan erotus $f - g$ muodossa

$$f - g = f - (f, \mathbf{w}_1)\mathbf{w}_1 - (f, \mathbf{w}_2)\mathbf{w}_2 - \cdots - (f, \mathbf{w}_n)\mathbf{w}_n.$$

Nähdään, että erotus $f - g$ on ortogonaalinen jokaista vektoria $\mathbf{w}_i$ kohti, joten se on ortogonaalinen jokaiselle avaruuden W vektorille. Etenkin erotus $f - g$ on ortogonaalinen erotukseen $g - \mathbf{w}$ nähden. Nyt vektorien summaan voidaan käyttää Pythagoraan lausetta $f - \mathbf{w} = (f - g) + (g - \mathbf{w})$, josta saadaan

$$||f - \mathbf{w}||^2 = ||f - g||^2 + ||g - \mathbf{w}||^2.$$

Täten $||f - g||^2 \leq ||f - \mathbf{w}||^2$ ja lauseen tulos $||f - g|| \leq ||f - \mathbf{w}||$ seuraa tästä. □

Lauseen 3 tuloksen mukaan pienimmän neliön approksimaation etsiminen vaatii, että sisätuloavaruuden W ortonormaalikanta tunnetaan. Kanta on ortonormaali, jos kannan kaikki vektorit $\mathbf{v}_i$ ovat keskenään ortogonaaliset ja jokaisen vektorin pituus on yksi. Lähteen [4] mukaan vektorit ovat ortogonaaliset silloin kun ne ovat kohtisuorassa toisiaan vastaan, eli silloin kun vektoreiden sisätulo $(\mathbf{v}_i, \mathbf{v}_j)$ on nolla. Tutustutaan seuraavaksi menetelmään, jolla vektoriavaruuden jokin kanta voidaan muuttaa ortonormaaliksi.

Lause 4. *([1]) (Gramin-Schmidtin ortonormalisointiprosessi)*
1. Olkoon joukko $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ *sisätuloavaruuden* V *kanta.*
2. Olkoon joukko $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$*, missä* $\mathbf{w}_i$ *on*

$$\begin{aligned}
\mathbf{w}_1 &= \mathbf{v}_1, \\
\mathbf{w}_2 &= \mathbf{v}_2 - \frac{(\mathbf{v}_2, \mathbf{w}_1)}{(\mathbf{w}_1, \mathbf{w}_1)}\mathbf{w}_1, \\
\mathbf{w}_3 &= \mathbf{v}_3 - \frac{(\mathbf{v}_3, \mathbf{w}_1)}{(\mathbf{w}_1, \mathbf{w}_1)}\mathbf{w}_1 - \frac{(\mathbf{v}_3, \mathbf{w}_2)}{(\mathbf{w}_2, \mathbf{w}_2)}\mathbf{w}_2, \\
&\vdots \\
\mathbf{w}_n &= \mathbf{v}_n - \frac{(\mathbf{v}_n, \mathbf{w}_1)}{(\mathbf{w}_1, \mathbf{w}_1)}\mathbf{w}_1 - \frac{(\mathbf{v}_n, \mathbf{w}_2)}{(\mathbf{w}_2, \mathbf{w}_2)}\mathbf{w}_2 - \cdots - \frac{(\mathbf{v}_n, \mathbf{w}_{n-1})}{(\mathbf{w}_{n-1}, \mathbf{w}_{n-1})}\mathbf{w}_{n-1}.
\end{aligned}$$

Tällöin B' *on avaruuden* V *ortogonaalinen kanta.*
3. Olkoon $\mathbf{u}_i = \frac{\mathbf{w}_i}{||\mathbf{w}_i||}$*. Tällöin joukko* $B'' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ *on avaruuden* V *ortonormaalikanta.*

Todistus. Osoitetaan induktiolla luvun h suhteen, että kun $h = 1, \ldots, n$, niin vektorit $\mathbf{w}_1, \ldots, \mathbf{w}_h$ ovat ortogonaaliset. Tapaus $h = n$ antaa lauseen kohdan 2 ja kohta 3 seuraa tästä.

Kun $h = 1$, niin tilanne on triviaali. Olkoon $h > 1$ ja tehdään induktio-oletus, että vektorit $\mathbf{w}_1, \ldots, \mathbf{w}_{h-1}$ ovat ortogonaaliset. Osoitetaan, että tällöin myös vektorit $\mathbf{w}_1, \ldots, \mathbf{w}_h$ ovat ortogonaaliset. Induktio-oletuksen nojalla riittää osoittaa, että sisätulot $(\mathbf{w}_h, \mathbf{w}_1) = \cdots = (\mathbf{w}_h, \mathbf{w}_{h-1}) = 0$. Kun $1 \leq j \leq h - 1$, niin

$$(\mathbf{w}_h, \mathbf{w}_j) = (\mathbf{v}_h, \mathbf{w}_j) - \sum_{i=1}^{h-1} \frac{(\mathbf{v}_h, \mathbf{w}_i)}{(\mathbf{w}_i, \mathbf{w}_i)}(\mathbf{w}_i, \mathbf{w}_j).$$

Induktio-oletuksen mukaan silloin kun $i \neq j$, niin summassa sisätulo $(\mathbf{w}_i, \mathbf{w}_j) = 0$ ja näin ollen

$$(\mathbf{w}_h, \mathbf{w}_j) = (\mathbf{v}_h, \mathbf{w}_j) - \frac{(\mathbf{v}_h, \mathbf{w}_j)}{(\mathbf{w}_j, \mathbf{w}_j)}(\mathbf{w}_j, \mathbf{w}_j) = 0,$$

joten B' on sisätuloavaruuden V ortogonaalinen kanta. □

Gramin-Schmidtin ortonormalisointiprosessin avulla avaruudelle voidaan muodostaa sen ortonormaalinen kanta. Kun tällainen kanta on löydetty, voidaan lauseen 3 tulosta käyttää jatkuvan funktion pienimmän neliön approksimaation etsimiseen. Seuraavassa esimerkissä käytetään mainittuja tuloksia ja etsitään pienimmän neliön approksimaatio funktiolle $f(x) = \sin x$ toisen asteen funktioiden aliavaruudessa.

Esimerkki 3. Etsitään toisen asteen funktioiden aliavaruuden W suhteen pienimmän neliön approksimaatio funktiolle $f(x) = \sin x$, kun $0 \leq x \leq \pi$.

Jotta tehtävä voidaan ratkaista lauseen 3 tuloksen avulla, täytyy ensin muodostaa avaruuden W kannan $\{1, x, x^2\}$ ortonormaalikanta, joka tehdään käyttämällä

Gramin-Schmidtin ortonormalisointiprosessia. Prosessin tulokseksi saadaan

$$\begin{aligned} B &= \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\} \\ &= \left\{ \frac{1}{\sqrt{\pi}}, \frac{\sqrt{3}}{\pi\sqrt{\pi}}(2x-\pi), \frac{\sqrt{5}}{\pi^2\sqrt{\pi}}(6x^2-6\pi x+\pi^2) \right\}. \end{aligned}$$

Lauseen 3 tuloksen mukaan pienimmän neliön approksimaatiofunktio on muotoa

$$g(x) = (f, \mathbf{w}_1)\mathbf{w}_1 + (f, \mathbf{w}_2)\mathbf{w}_2 + (f, \mathbf{w}_3)\mathbf{w}_3.$$

Funktion g summattaviksi tekijöiksi saadaan

$$\begin{aligned} (f, \mathbf{w}_1)\mathbf{w}_1 &= \int_0^\pi \sin\, x \cdot \frac{1}{\sqrt{\pi}}\ dx \cdot \frac{1}{\sqrt{\pi}} = \frac{1}{\pi}\int_0^\pi \sin\, x\ dx = \frac{2}{\pi}, \\ (f, \mathbf{w}_2)\mathbf{w}_2 &= \int_0^\pi \sin\, x \cdot \frac{\sqrt{3}}{\pi\sqrt{\pi}}(2x-\pi)\ dx \cdot \frac{\sqrt{3}}{\pi\sqrt{\pi}}(2x-\pi) \\ &= \frac{3}{\pi^3}(2x-\pi)\int_0^\pi \sin\, x \cdot (2x-\pi)\ dx \\ &= \frac{3}{\pi^3}(2x-\pi)\Big[2\ \sin\, x + (\pi - 2x)\ \cos\, x\Big]_0^\pi = 0 \end{aligned}$$

ja

$$\begin{aligned} (f, \mathbf{w}_3)\mathbf{w}_3 &= \int_0^\pi \sin\, x \cdot \frac{\sqrt{5}}{\pi^2\sqrt{\pi}}(6x^2-6\pi x+\pi^2)\ dx \cdot \frac{\sqrt{5}}{\pi^2\sqrt{\pi}}(6x^2-6\pi x+\pi^2) \\ &= \frac{5}{\pi^5}(6x^2-6\pi x+\pi^2)\int_0^\pi \sin\, x \cdot (6x^2-6\pi x+\pi^2)\ dx \\ &= \frac{10}{\pi^5}(6x^2-6\pi x+\pi^2)(\pi^2-12). \end{aligned}$$

Funktion $f(x) = \sin\, x$ pienimmän neliön approksimaatiofunktio välillä $x \in [0, \pi]$ on

$$g(x) = \frac{2}{\pi} + \frac{10}{\pi^5}(6x^2-6\pi x+\pi^2)(\pi^2-12) \approx -0,4177x^2 + 1,3122x - 0,0505.$$

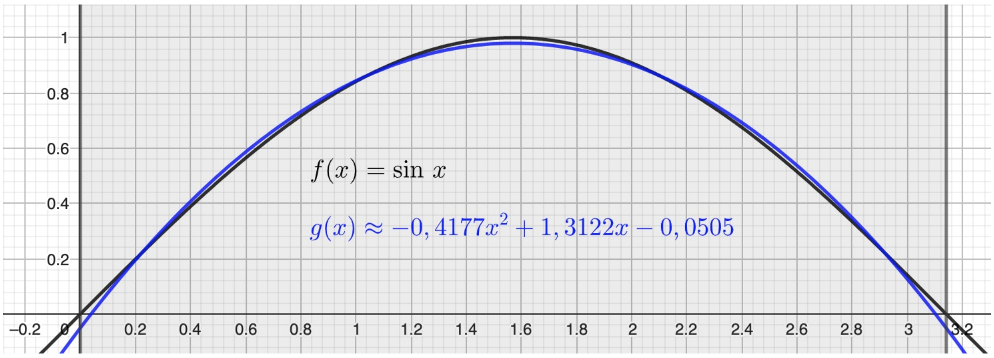
Funktio f ja sen approksimaatiofunktio g esitetään kuvassa 2.

3.2 Fourier'n approksimaatio

Esitellään seuraavaksi pienimmän neliön approksimaation sovellus Fourier'n approksimaatio, jossa approksimaatiofunktiot ovat muotoa

$$g(x) = \frac{a_0}{2} + a_1\ \cos\, x + \cdots + a_n\ \cos\, nx + b_1\ \sin\, x + \cdots + b_n\ \sin\, nx.$$

Approksimaatiofunktiot kuuluvat avaruuden $C[0, 2\pi]$ aliavaruuteen W, joka on kannan $S = \{1, \cos\, x, \ldots, \cos\, nx, \sin\, x, \ldots, \sin\, nx\}$ muodostama. Approksimaatiofunktiota kutsutaan funktion f kertaluvun n Fourier'n approksimaatioksi yli välin $[0, 2\pi]$.



Kuva 2: Funktion $f(x) = \sin\, x$ ja sen approksimaatiofunktion kuvaajat yli välin $x \in [0, \pi]$.

Huomataan, että kanta S muodostuu $2n + 1$ kappaleesta vektoreita, jotka ovat lähteen [1] mukaan ortogonaalisia sisätuloavaruuden $C[0, 2\pi]$ suhteen. Kanta S voidaan siis ortonormalisoida Gramin-Schmidtin ortonormalisointiprosessilla, jonka tulokseksi saadaan

$$S' = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos\, x, \ldots, \frac{1}{\sqrt{\pi}} \cos\, nx, \frac{1}{\sqrt{\pi}} \sin\, x, \ldots, \frac{1}{\sqrt{\pi}} \sin\, nx \right\}.$$

Koska kanta S' on ortonormaali, voidaan Fourier'n approksimaatiofunktio g esittää lauseen 3 mukaisessa muodossa $g(x) = (f, \mathbf{w}_0)\mathbf{w}_0 + (f, \mathbf{w}_1)\mathbf{w}_1 + \cdots + (f, \mathbf{w}_{2n})\mathbf{w}_{2n}$, jolloin Fourier'n kertoimet $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ voidaan muodostaa integraaleista

$$a_0 = (f, \mathbf{w}_0)\frac{2}{\sqrt{2\pi}} = \frac{2}{\sqrt{2\pi}} \int_0^{2\pi} f(x)\frac{1}{\sqrt{2\pi}} dx = \frac{1}{\pi} \int_0^{2\pi} f(x)\, dx,$$

$$a_j = (f, \mathbf{w}_j)\frac{1}{\sqrt{\pi}} = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x)\frac{1}{\sqrt{\pi}} \cos\, jx\, dx = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos\, jx\, dx$$

ja

$$b_j = (f, \mathbf{w}_{n+j})\frac{1}{\sqrt{\pi}} = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(x)\frac{1}{\sqrt{\pi}} \sin\, jx\, dx = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin\, jx\, dx.$$

Nyt on määritelty Fourier'n approksimaation muodostamiseen vaadittavat tekijät. Fourier'n approksimaatio esitellään kokonaisuudessaan määritelmässä 3.

Määritelmä 3. Välillä $[0, 2\pi]$ jatkuvan funktion f pienimmän neliön approksimaatio vektoriavaruudessa, joka on kannan $\{1, \cos\, x, \ldots, \cos\, nx, \sin\, x, \ldots, \sin\, nx\}$ muodostama, on muotoa

$$g(x) = \frac{a_0}{2} + a_1 \cos\, x + \cdots + a_n \cos\, nx + b_1 \sin\, x + \cdots + b_n \sin\, nx,$$

missä Fourier'n kertoimet $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ ovat

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x)\, dx,$$

$$a_j = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos\, jx\, dx,\ j = 1, 2, \ldots, n$$

ja

$$b_j = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin jx \, dx, \ j = 1, 2, \ldots, n.$$

Tarkastellaan seuraavaksi Fourier'n approksimaation käyttöä esimerkin 4 avulla.

Esimerkki 4. Etsitään kolmannen kertaluvun Fourier'n approksimaatiofunktio funktiolle $f(x) = x$, kun $0 \leq x \leq 2\pi$.

Määritelmän 3 mukaan Fourier'n approksimaatiofunktio on muotoa

$$g(x) = \frac{a_0}{2} + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x,$$

missä

$$\begin{aligned} a_0 &= \frac{1}{\pi} \int_0^{2\pi} x \, dx = \frac{1}{\pi} 2\pi^2 = 2\pi, \\ a_j &= \frac{1}{\pi} \int_0^{2\pi} x \cos jx \, dx = \left[\frac{1}{\pi j^2} \cos jx + \frac{x}{\pi j} \sin jx \right]_0^{2\pi} = 0 \end{aligned}$$

ja

$$b_j = \frac{1}{\pi} \int_0^{2\pi} x \sin jx \, dx = \left[\frac{1}{\pi j^2} \sin jx - \frac{x}{\pi j} \cos jx \right]_0^{2\pi} = -\frac{2}{j}.$$

Näin ollen funktion $f(x) = x$ kolmannen kertaluvun Fourier'n approksimaatiofunktio välillä $x \in [0, 2\pi]$ on

$$g(x) = \frac{2\pi}{2} - \frac{2}{1} \sin x - \frac{2}{2} \sin 2x - \frac{2}{3} \sin 3x = \pi - 2 \sin x - \sin 2x - \frac{2}{3} \sin 3x.$$

Esimerkissä 4 etsittiin kolmannen kertaluvun Fourier'n approksimaatiota. Kun tarkastellaan approksimaatiofunktion Fourier'n kertoimia yleisesti, huomataan, että $a_0 = 2\pi$, $a_1 = a_2 = \ldots = a_n = 0$ ja $b_1 = -\frac{2}{1}, b_2 = -\frac{2}{2}, \ldots, b_n = -\frac{2}{n}$. Kertoimien avulla voidaan muodostaa funktion $f(x) = x$ yleinen Fourier'n approksimaatio

$$g(x) = \pi - \left(2\sin x + \sin 2x + \frac{2}{3} \sin 3x + \ldots + \frac{2}{n} \sin nx \right).$$

Fourier'n approksimaatiolle pätee, että kertaluvun n kasvaessa Fourier'n approksimaatio tarkentuu. Tarkemmin sanottuna, kun kertaluku $n \to \infty$, niin approksimaation virhe $||f - g|| \to 0$ kaikilla muuttujan x arvoilla välillä $(0, 2\pi)$.

Viitteet

[1] Ron Larson & David C. Falvo: *Elementary Linear Algebra*. Houghton Mifflin Harcourt Publishing Company, Boston, 2009.

[2] Jyrki Lahtonen: *Vektorilaskenta*. Turun yliopisto, 2023.

[3] Roger A. Horn & Charles R. Johnson: *Matrix Analysis*. Cambridge University Press, 1985.

[4] Markku Koppinen: *Lineaarialgebra*. Turun yliopisto, 2006.