



SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

ORTOGONALA POLYNOM OCH GAUSS-KVADRATUR

av

Valentin Kadushkin

2025 - No K15

ORTOGONALA POLYNOM OCH GAUSS-KVADRATUR

Valentin Kadushkin

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Rikard Bögvald

2025

Abstract

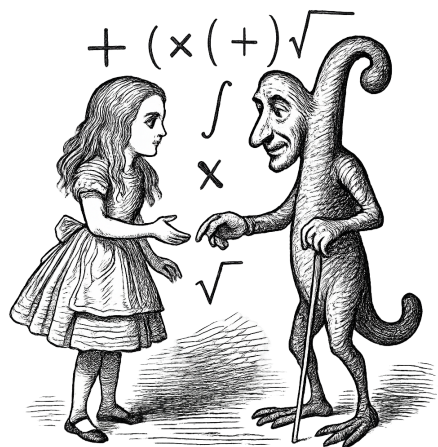
This thesis examines the theory of Gaussian quadrature and its connection to orthogonal polynomials. Gaussian quadrature is an efficient numerical method for approximating integrals especially if the integrand can be expressed as a polynomial or is well approximated by one. The method is based on choosing nodes and weights in such a way that the integral is exact for all polynomials up to degree $2n - 1$. The nodes are zeros of an orthogonal polynomial while the weights are determined based on the orthogonality properties. The thesis first discusses the basic properties of orthogonal polynomials, focusing on classical examples such as Legendre polynomials. This is followed by a formal proof of the exactness of Gaussian quadrature as well as a discussion of the method's applications. The aim is to provide a theoretical understanding of why Gaussian quadrature works and how orthogonal polynomials play a central role in its construction.

Abstrakt

Denna uppsats behandlar teorin bakom Gauss-kvadratur och dess koppling till ortogonala polynom. Gauss kvadratur är en effektiv numerisk metod för att approximera integraler särskilt när integranden kan uttryckas som ett polynom eller väl kan approximeras av ett sådant. Metoden bygger på att välja noder och vikter på ett sådant sätt att integralen blir exakt för alla polynom upp till grad $2n - 1$. Noderna är nollställena till ett ortogonalt polynom medan vikterna bestäms utifrån ortogonalitetsegenskaperna. Uppsatsen inleder med att diskutera grundläggande egenskaper hos ortogonala polynom med fokus på klassiska exempel som Legendrepolynom. Därefter följer ett bevis för Gauss-kvadraturens exakthet samt en diskussion om metodens tillämpningar. Syftet med uppsatsen är att ge en teoretisk förståelse till varför Gauss kvadratur fungerar och hur ortogonala polynom spelar en central roll i metodens konstruktion.

Innehåll

1	Inledning	4
1.1	Bakgrund	4
2	Teoretisk bakgrund	6
2.1	Hilbertrum	6
2.2	Ortogonal funktioner	12
2.3	Ortogonal polynom	14
2.4	Klassiska exempel på ortogonal polynom	18
2.5	Grahm-Schmidt processen	20
3	Gauss-kvadratur	23
3.1	Bakgrund	23
3.2	Gauss-Legendre-kvadratur	28
3.3	Hur hittar man vikter och noder?	30
3.4	Jämförelse med traditionella numeriska metoder	34
4	Slutsats	37



1 Inledning

1.1 Bakgrund

Integraler spelar en av de mest centrala rollerna i matematiken och har många tillämpningar inom bl.a. fysik, ekonomi, medicin, datorgrafik och sannolikhets-teori. Genom integration kan man beskriva hur mycket arbete som krävs för att flytta ett föremål, hur stor mängd läkemedel som sprids i kroppen över tid, eller hur stor en ekonomisk vinst blir över en viss tidsperiod.

Det faktum att integraler knyter samman teori och verklighet gör integration till ett ytterst intressant ämne att utforska. Integration medför en rad utmaningar som matematiker, sedan gamla tider, har försökt lösa genom att upptäcka nya metoder och tekniker. En av de största utmaningarna man stöter på är att analytisk integration inte alltid är möjlig. Anledningen kan vara att det finns funktioner som helt enkelt saknar primitiva funktioner som kan uttryckas i termer av elementära funktioner (polynom, trigonometriskafunktioner etc). Det leder till att analytisk integration i praktiken kan vara omöjlig att utföra. Ett klassiskt exempel på detta är täthetsfunktionen $f(x)$ för normalfördelningen [1, s. 104]:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Med tiden utvecklades datorer och som konsekvens till detta blev numeriska metoder mer och mer praktiska. En av de mest effektiva numeriska metoder för integration som vi har idag kallas *Gauss-kvadratur*. Metoden är formellt definierat av Carl Gustav Jacobi år 1826. Denna metod bygger på teorin om ortogonala polynom.

Syftet med detta arbete är studera ortogonala polynom samt att förstå och undersöka den matematiska teorin bakom Gauss-kvadratur och hur ortogonala polynom ligger till grunden för denna teori.

2 Teoretisk bakgrund

2.1 Hilbertrum

För att kunna tala om ortogonalitet mellan polynom krävs att de ligger i ett inre-produktrum. Låt oss därför påminna läsaren om vektorrum och hur detta koncept kan utvidgas till ett mer generell koncept - nämligen något som kallas för ett *Hilbertrum*.

Formellt är ett vektorrum en icke-tom mängd V av objekt som har en addition och en skalärmultiplikation enligt vissa regler. Objekt i ett vektorrum kallas för vektorer men i själva verket kan det vara matriser, funktioner, polynom, reella tal och naturligtvis vektorer.

Definition 2.1: Vektorrum som algebraisk struktur

Ett *vektorrum* över kroppen $\mathbb{R}$ (eller $\mathbb{C}$) är en mängd V utrustad med två operationer:

- En **addition** $+: V \times V \rightarrow V$ som till varje par $(\vec{v}, \vec{w})$ i V ordnar en ny vektor $\vec{v} + \vec{w} \in V$.
- En **skalärmultiplikation** $\cdot: \mathbb{R} \times V \rightarrow V$ som till varje par $(k, \vec{v})$ där $k \in \mathbb{R}$ och $\vec{v} \in V$ ordnar en ny vektor $k\vec{v} \in V$.

Dessa operationer måste uppfylla de tio axiomen som anges i Definition 2.2.

Definition 2.2: Axiom för ett vektorrum

Låt $\vec{u}, \vec{v}$ och $\vec{w}$ vara element i V och låt k och l vara skalärer. Då är V ett vektorrum om och endast om följande villkor är uppfyllda:

1. V är sluten under addition, det vill säga om $\vec{u}$ och $\vec{v}$ är i V så är $\vec{u} + \vec{v}$ också i V .
2. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
3. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
4. V innehåller ett objekt $\vec{0}$ (kallad *nollvektor*) så att $\vec{u} + \vec{0} = \vec{u}$ för varje $\vec{u} \in V$
5. För varje objekt $\vec{u} \in V$ finns det ett objekt $-\vec{u} \in V$ sådant att

$$\vec{u} + (-\vec{u}) = \vec{0}$$

6. V är sluten under multiplikation, det vill säga om $\vec{u} \in V$ och k är en skalär så är $k\vec{u} \in V$
7. $(k + l)\vec{u} = k\vec{u} + l\vec{u}$
8. $k(\vec{u} + \vec{v}) = k\vec{u} + k\vec{v}$
9. $(kl)\vec{u} = k(l\vec{u})$
10. $1\vec{u} = \vec{u}$

David Hilbert generaliserar idén om Euklidiskt rum (dvs tredimensionellt vektorrum) till oändligt-dimensionella rum. Rent matematiskt är ett Hilbertrum ett *inreproduktum* som är **fullständigt**. Hilbertrum uppstår typiskt som oändligt-

dimensionella *funktionsrum*. Ett element i ett Hilbertrum kan uttryckas med sina koordinater i förhållande till en ortonormal bas i rummet (direkt analogt med koordinaterna $\{x, y, z\}$ i tredimensionellt vektorrum).

Definition 2.3: Hilbertrum

Ett Hilbertrum är ett vektorrum med två extra kriterier:

11. Rummet är utrustat med en inre produkt:

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R} \quad (\text{eller } \mathbb{C})$$

som uppfyller följande egenskaper för alla $u, v, w \in V$ och skalär a :

- $\langle u, v \rangle = \langle v, u \rangle$ (symmetri)
- $\langle u, v \rangle = \overline{\langle v, u \rangle}$ (symmetri i baskroppen $\mathbb{C}$)
- $\langle au + v, w \rangle = a\langle u, w \rangle + \langle v, w \rangle$ (linjäritet i första argumentet)
- $\langle v, v \rangle \geq 0$, med likhet endast om $v = 0$ (positiv definitthet)

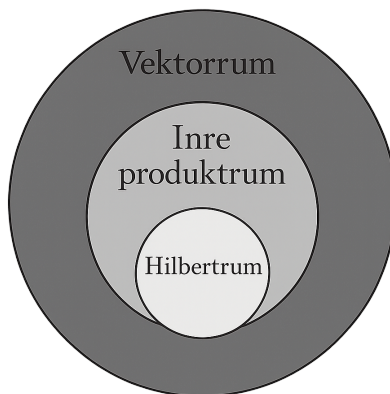
12. Vektorrummet är fullständigt med avseende på normen inducerad av den inre produkten:

$$\|v\| := \sqrt{\langle v, v \rangle}$$

Det betyder att varje Cauchy-följd $\{v_n\}_{n=1}^\infty$ i V konvergerar mot ett element i V .

Talföljden $\{v_n\}_{n=1}^\infty$ kallas en *Cauchy-följd* om

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \text{ så att } \|v_n - v_m\| < \varepsilon \quad \text{för alla } n, m \geq N.$$



Figur 1: Relationen mellan Vektorrum, Inre produktrum samt Hilbertrum.

Det är viktigt att poängtera att kravet på att vektorrummet ska vara ett Hilbertrum innebär att varje Cauchy-följd i rummet måste konvergera mot ett element i rummet. Att en följd är en Cauchy-följd betyder endast att elementen konvergerar, men det betyder inte nödvändigtvis att följdens har ett gränsvärde inom rummet. Det är just detta som skiljer ett allmänt inre produktrum från ett Hilbertrum - att i ett Hilbertrum så konvergerar alla Cauchy-följder.

Sats 2.1: Funktioner på ett intervall bildar ett vektorrum

Låt V vara mängden av alla reella funktioner definierade på ett intervall $[a, b] \subset \mathbb{R}$. Med addition och skalärmultiplikation definierade som

$$(f + g)(x) = f(x) + g(x), \quad (\alpha f)(x) = \alpha f(x)$$

för alla $f, g \in V$, $\alpha \in \mathbb{R}$ och $x \in [a, b]$ är V ett vektorrum över $\mathbb{R}$ (eller $\mathbb{C}$).

Bevis. Vi visar att V uppfyller axiomen för ett vektorrum:

- Slutenhets under addition: För $f, g \in V$ är $(f + g)(x) = f(x) + g(x)$ en reell funktion på $[a, b]$ således $f + g \in V$.

- Associativitet av addition: För $f, g, h \in V$ och varje $x \in [a, b]$ gäller:

$$((f + g) + h)(x) = (f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x)) = (f + (g + h))(x).$$

- Existens av nollelement: Nollfunktionen $k(x)$ definierad genom $k(x) = 0$ för alla x uppfyller

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

- Existens av inverser: För varje $f \in V$ definieras $(-f)(x) = -f(x)$ och

$$(f + (-f))(x) = f(x) + (-f(x)) = 0.$$

- Kommutativitet av addition: För $f, g \in V$ och varje $x \in [a, b]$ gäller

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x).$$

- Slutenhets under multiplikation: För $\alpha \in \mathbb{R}$ och $f \in V$ är $(\alpha f)(x) = \alpha f(x)$ en reell funktion på $[a, b]$ alltså $\alpha f \in V$.

- Distributivitet : För $\alpha \in \mathbb{R}$ och $f, g \in V$ gäller

$$\alpha(f + g)(x) = \alpha(f(x) + g(x)) = \alpha f(x) + \alpha g(x) = (\alpha f + \alpha g)(x).$$

- Distributivitet (skalärer): För $\alpha, \beta \in \mathbb{R}$ och $f \in V$ gäller

$$(\alpha + \beta)f(x) = \alpha f(x) + \beta f(x) = (\alpha f + \beta f)(x).$$

- Associativitet av skalmultiplikation: För $\alpha, \beta \in \mathbb{R}$ och $f \in V$ gäller

$$\alpha(\beta f)(x) = \alpha(\beta f(x)) = (\alpha\beta)f(x) = ((\alpha\beta)f)(x).$$

- Existens av enhetsfaktor: För alla $f \in V$ gäller

$$1 \cdot f(x) = 1 \cdot f(x) = f(x).$$

Alla axiomen för ett vektorrum är uppfyllda och V är således ett vektorrum över $\mathbb{R}$. $\square$

Exempel 2.1

Ett triviale exempel på ett Hilbertrum är det tvådimensionella rummet $\mathbb{R}^2$ med den vanliga inre produkten.

Betrakta vektorrummet

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

Definiera inre produkten för två vektorer $\mathbf{u} = (u_1, u_2)$ och $\mathbf{v} = (v_1, v_2)$ som

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2.$$

Normen som induceras av inre produkten är

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{u_1^2 + u_2^2}.$$

Positiv definit:

$$\langle \mathbf{u}, \mathbf{u} \rangle = u_1^2 + u_2^2 \geq 0$$

med lika med noll endast om $\mathbf{u} = (0, 0)$.

Symmetri:

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + u_2 v_2 = v_1 u_1 + v_2 u_2 = \langle \mathbf{v}, \mathbf{u} \rangle$$

Linjäritet i första argumentet:

För $a, b \in \mathbb{R}$ och $\mathbf{u}, \mathbf{w}, \mathbf{v} \in \mathbb{R}^2$ gäller:

$$\langle a\mathbf{u} + b\mathbf{w}, \mathbf{v} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle + b\langle \mathbf{w}, \mathbf{v} \rangle$$

Eftersom $\mathbb{R}^2$ är ett ändligdimensionellt normerat rum över $\mathbb{R}$ är det automatiskt fullständigt. Alltså är det ett Hilbertrum.

Fullständighet:

Låt $\vec{v}_n = (x_n, y_n)$ vara en Cauchy-följd i $\mathbb{R}^2$. Det betyder att vektorerna till slut ligger hur nära varandra som helst det vill säga:

$$\|\vec{v}_n - \vec{v}_m\| < \varepsilon \quad \text{för alla } n, m \geq N$$

Eftersom normen i $\mathbb{R}^2$ definieras som:

$$\|\vec{v}_n - \vec{v}_m\| = \sqrt{(x_n - x_m)^2 + (y_n - y_m)^2},$$

innebär det att både följderna (x_n) och (y_n) också måste vara Cauchy-följder i $\mathbb{R}$. Men $\mathbb{R}$ är fullständigt så båda dessa följder konvergerar:

$$x_n \rightarrow x, \quad y_n \rightarrow y$$

Därför konvergerar hela vektorföljden:

$$\vec{v}_n = (x_n, y_n) \rightarrow (x, y) \in \mathbb{R}^2$$

Således konvergerar varje Cauchy-följd i $\mathbb{R}^2$.

2.2 Ortogonala funktioner

Ur matematiskt perspektiv betyder *ortogonal* vinkelrät. En mängd av vektorer $\{\vec{v}, \vec{u}, \vec{w}\}$ i rummet $\mathbb{R}^3$ är ortogonala om skalärprodukten av varje par är noll.

$$\vec{u} \cdot \vec{v} = 0, \quad \vec{u} \cdot \vec{w} = 0, \quad \vec{v} \cdot \vec{w} = 0.$$

Man talar således om en ortogonal mängd av vektorer. Om $\{\vec{v}, \vec{u}, \vec{w}\}$ är dessutom enhetsvektorer säger man att en sådan mängd är *ortonormal*. Ortogonala vektorer är linjärt oberoende och spänner därmed upp hela rummet $\mathbb{R}^3$.

På samma sätt som man definierar en inre produkt för vektorer i $\mathbb{R}^n$ kan man definiera det för funktioner. Detta gör att man kan prata om ortogonalitet även i fallet med funktioner.

Sats 2.2: Inre produkt och Hilbertrum

Givet en kontinuerlig viktfunction $w(x) > 0$ på (a, b) definieras en inre produkt mellan kontinuerliga funktioner f och g genom:

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

Med denna inre produkt bildar mängden av polynom ett inre produktrum.

Bevis. Vi visar först att uttrycket

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

definierar en inre produkt.

- Linjäritet i första argumentet: För alla funktioner f, g, h och alla skalär α gäller:

$$\langle \alpha f + g, h \rangle = \int_a^b (\alpha f(x) + g(x))h(x)w(x) dx = \alpha \langle f, h \rangle + \langle g, h \rangle$$

- Symmetri: Eftersom multiplikation av reella tal är kommutativ gäller det att:

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

- Positiv definitthet: Vi har

$$\langle f, f \rangle = \int_a^b f(x)^2 w(x) dx$$

Eftersom $w(x) > 0$ nästan överallt och $f(x)^2 \geq 0$ med likhet endast om $f(x) = 0$ överallt följer det att $\langle f, f \rangle = 0$ om och endast om $f = 0$.

Mer detaljerat: Antag att $f(x_0) \neq 0$, där $x_0 \in [a, b]$. Då finns $\delta > 0$ så att

$$|f(x) - f(x_0)| < \frac{|f(x_0)|}{2} =: \varepsilon \quad \text{om} \quad |x - x_0| < \delta$$

på grund av kontinuiteten. Alltså är

$$|f(x)| \geq |f(x_0)| - \frac{|f(x_0)|}{2} = \frac{|f(x_0)|}{2} \quad \text{på} \quad [x_0 - \delta, x_0 + \delta].$$

Vilket ger att

$$\int_a^b |f(x)|^2 dx \geq \int_{x_0-\delta}^{x_0+\delta} |f(x)|^2 dx \geq \frac{|f(x_0)|^2}{4} \cdot 2\delta > 0.$$

Om funktionen inte är noll överallt så blir inre-produkten större än noll vilket bevisar att inre produkten är positivt definit.

- Mängden av polynom är ett vektorrum: Vi vet från grundläggande algebra att summan av polynom är ett polynom och en konstant gånger ett polynom är också ett polynom.

Således är $\langle \cdot, \cdot \rangle$ en inre produkt.

□

Nu är vi redo att definiera ortogonalitet samt ortonormalitet för reellvärda funktioner:

Definition 2.3: Ortogonalitet

Två reellvärda funktioner $f(x)$ och $g(x)$ sägs vara **ortogonala** på intervallet $[a, b]$ om:

$$\int_a^b f(x)g(x) dx = 0 \quad (1)$$

Definition 2.4: Ortonormalitet

Två reellvärda funktioner $f(x)$ och $g(x)$ sägs vara **ortonormala** på intervallet $[a, b]$ om:

$$\int_a^b f(x)g(x) dx = \delta_{fg} \quad (2)$$

där

$$\delta_{fg} = \begin{cases} 1 & \text{om } f = g \\ 0 & \text{om } f \neq g \end{cases}$$

δ_{fg} är Kroneckers delta.

En oändlig följd av polynom p_n där varje p_n har grad n och mängden är linjärt oberoende bildar en bas för vektorrummet av alla polynom. För att konstruera en ortogonal bas från denna mängd kan man använda Gram-Schmidts ortogonaliseringsprocess med en godtycklig inreprodukt.

Sats 2.3: Bas för polynomrum

Låt $p_n(x) = x^n + \sum_{k=0}^{n-1} a_{n,k}x^k$ där $a_{n,k} \in \mathbb{R}$ för varje $n = 0, 1, \dots, N$. Då är mängden $\{p_0(x), p_1(x), \dots, p_N(x)\}$ en bas för vektorrummet av alla reella polynom av grad högst N .

Bevis. För att visa att $\{p_0(x), p_1(x), \dots, p_N(x)\}$ är en bas visar vi först att de är linjärt oberoende.

Antag att det finns reella tal $c_0, c_1, \dots, c_N$ sådana att

$$\sum_{n=0}^N c_n p_n(x) = 0 \quad \text{för alla } x.$$

Eftersom varje $p_n(x)$ har ledande term x^n med koefficient 1 så bidrar endast $p_N(x)$ med termen x^N . Det innebär att koefficienten framför x^N i summan ovan är c_N och då gäller det att $c_N = 0$.

På samma sätt är nu den ledande termen i summan $c_{N-1}x^{N-1}$ och den kan inte elimineras av några andra polynom eftersom alla $p_n(x)$ för $n < N - 1$ har grad lägre än $N - 1$. Därför måste även $c_{N-1} = 0$. Genom att upprepa detta resonemang för $c_{N-2}, c_{N-3}, \dots, c_0$ får vi att alla koefficienter är noll. Mängden är således linjärt oberoende.

Eftersom de är $N + 1$ stycken och rummet av polynom av grad högst N har dimension $N + 1$ så spänner mängden således upp hela rummet.

Därmed är $\{p_0(x), p_1(x), \dots, p_N(x)\}$ en bas. □

2.3 Ortogonala polynom

Börjar med ett konkret definition:

Definition 2.5: Ortogonala polynom

En mängd ortogonala polynom är en oändlig följd av polynom $p_0(x), p_1(x), p_2(x), \dots$, där $p_n(x)$ är ett polynom av grad n och varje par av polynom i mängden är ortogonala mot varandra, mer formellt:

$$\int_a^b f(x)g(x) dx = 0 \quad \text{om } f(x) \neq g(x) \quad (3)$$

Definition 2.6: Ortonormerade polynom

Mängden av polynom är **ortonormerad** om:

$$\int_a^b f(x)g(x) dx = \delta_{fg} \quad (4)$$

där $\delta_{fg} = \begin{cases} 1 & \text{om } f = g \\ 0 & \text{om } f \neq g \end{cases}$ är Kroneckers delta.

Definition 2.7: Ortogonalitet med viktfunction

En mängd ortogonala polynom är en oändlig följd av polynom $p_0(x), p_1(x), p_2(x), \dots$, där $p_n(x)$ är ett polynom av grad n , och varje par av polynom i mängden är ortogonala mot varandra med avseende på en viktfunction $w(x) > 0$ på intervallet (a, b) , det vill säga:

$$\int_a^b f(x)g(x)w(x) dx = 0 \quad \text{om} \quad f(x) \neq g(x)$$

Ortogonal polynom har följande egenskaper som kan formuleras som satser:

Sats 2.4: Linjär kombination av ortogonala polynom

Varje polynom $f(x)$ av grad n kan skrivas som en linjär kombination av polynomen $p_0(x), p_1(x), \dots, p_n(x)$.

Bevis. Eftersom $\{p_0(x), p_1(x), \dots, p_n(x)\}$ är ortogonal mängd av polynom där varje $p_i(x)$ har grad i är dessa polynom linjärt oberoende (Enligt Sats 2.3). Mängden av polynom av grad högst n bildar ett vektorrum av dimension $n+1$, och $\{p_0, \dots, p_n\}$ är en bas för detta rum. Därför kan varje polynom $f(x)$ av grad n uttryckas som en linjär kombination av dessa polynom. $\square$

Sats 2.5: Ortogonalitet mot polynom av lägre grad

Givet en ortogonal mängd av polynom $\{p_0(x), p_1(x), \dots\}$ är varje polynom $p_k(x)$ ortogonalt mot alla polynom av grad mindre än k .

Bevis. Eftersom mängden är ortogonal gäller:

$$\int_a^b p_k(x)p_j(x)w(x)dx = 0 \quad \text{för alla } j < k$$

Eftersom varje polynom av grad mindre än k kan skrivas som en linjär kombination av $p_0, p_1, \dots, p_{k-1}$, följer att p_k är ortogonalt mot varje sådant polynom. $\square$

Sats 2.6: Rekursionsformel för ortogonala polynom

Varje ortogonal polynomföljd $\{p_0(x), p_1(x), \dots\}$ uppfyller en rekursionsformel som relaterar tre på varandra följande polynom:

$$p_{n+1}(x) = (a_n x + b_n)p_n(x) - c_n p_{n-1}(x),$$

där a_n, b_n och c_n är reella konstanter som beror på n .

Bevis. Eftersom $p_n(x)$ är ett polynom av grad n så är $x p_n(x)$ ett polynom av grad $n + 1$. Därför kan man skriva $x p_n(x)$ som en linjärkombination av alla polynom i basen $\{p_0(x), p_1(x), \dots, p_{n+1}(x)\}$. Men eftersom polynomen är ortogonala gäller att

$$\langle x p_n(x), p_i(x) \rangle = 0 \quad \text{för alla } i < n - 1 \text{ eller } i > n + 1.$$

Därför är det endast $p_{n+1}(x)$, $p_n(x)$ och $p_{n-1}(x)$ som eventuellt ger icke-noll bidrag i linjärkombinationen. Alltså existerar det konstanter $\alpha, \beta, \gamma \in \mathbb{R}$ sådana att:

$$x p_n(x) = \alpha p_{n+1}(x) + \beta p_n(x) + \gamma p_{n-1}(x).$$

Vi kan nu lösa ut $p_{n+1}(x)$ ur detta uttryck:

$$p_{n+1}(x) = \frac{1}{\alpha} (x p_n(x) - \beta p_n(x) - \gamma p_{n-1}(x)).$$

Om vi sätter $a_n = \frac{1}{\alpha}$, $b_n = -\frac{\beta}{\alpha}$ och $c_n = \frac{\gamma}{\alpha}$ får vi rekursionsformen:

$$p_{n+1}(x) = (a_n x + b_n)p_n(x) - c_n p_{n-1}(x)$$

$\square$

Sats 2.7: Reella och distinkta rötter

Varje polynom $p_n(x)$ i en ortogonal polynomföljd har exakt n reella och distinkta rötter, alla strikt inom ortogonalitetsintervallet (a, b) .

Bevis. Antag att $p_n(x)$ har $m < n$ reella rötter $x_1, x_2, \dots, x_m$ på (a, b) . Definiera ett polynom $q(x) = (x - x_1)(x - x_2) \dots (x - x_m)$ som har grad $m < n$. Då har $p_n(x)$ och $q(x)$ samma nollställen. Eftersom $p_n(x)$ är ortogonalt mot alla polynom av lägre grad gäller det att:

$$\int_a^b p_n(x)q(x)w(x) dx = 0$$

Eftersom $p_n(x)$ och $q(x)$ har samma nollställen inom (a, b) är produkten $p_n(x)q(x)$ icke-negativ eller icke-positiv över hela intervallet (då de har samma tecken mellan rötterna).

Eftersom $w(x) > 0$ på (a, b) (påverkar inte teckenväxlingen) är integranden $p_n(x)q(x)w(x)$ icke-negativ eller icke-positiv över hela intervallet och vars integral över $[a, b]$ kan inte bli noll. Detta motsäger dock ortogonaliteten. Således måste $p_n(x)$ ha exakt n reella och distinkta rötter inom (a, b) .

□

2.4 Klassiska exempel på ortogonala polynom

Legendrepolynom

Det finns flera olika familjer av ortogonala polynom. Den kanske mest klassiska familjen av ortogonala polynom är Legendrepolynom. Legendrepolynom definieras på intervallet $[-1, 1]$ och viktfunktionen är $w(x) = 1$. Inre produkt definieras som:

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx.$$

De sex första Legendrepolynomen är:

$$P_0(x) = 1.$$

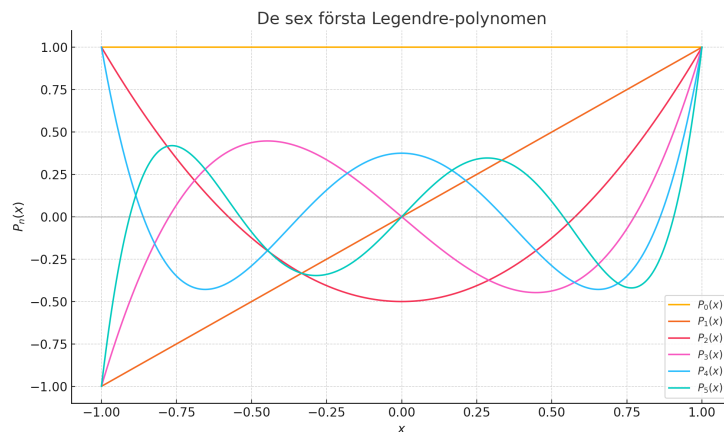
$$P_1(x) = x.$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1).$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x).$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3).$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x).$$



Figur 2: Legendrepolynom $P_0(x) - P_5(x)$

Några viktiga egenskaper hos Legendrepolynom som har formulerats som satser ovan (2.3 – 2.7):

1. Ortogonalitet:

$$\int_{-1}^1 P_n(x)P_m(x) dx = 0 \quad \text{för } n \neq m$$

2. Normalisering:

$$\int_{-1}^1 P_n(x)^2 dx = \frac{2}{2n+1}$$

3. Rekursionsformel:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x)$$

4. Rötter:

- $P_n(x)$ har exakt n reella och unika rötter.
- Alla rötter ligger strikt inom intervallet $(-1, 1)$.

Det finns flera familjer av ortogonala polynom. tabell 1 visar de klassiska ortogonala polynomen och deras intervall ortogonalitetsintervall [5]:

Polynom	Viktfunktion $w(x)$	Ortogonalitetsintervall
Legendre	1	$[-1, 1]$
Chebyshev	$\frac{1}{\sqrt{1-x^2}}$	$[-1, 1]$
Laguerre	e^{-x}	$[0, \infty)$
Hermite	e^{-x^2}	$(-\infty, \infty)$
Jacobi	$(1-x)^\alpha(1+x)^\beta$	$[-1, 1]$

Tabell 1: Exempel på klassiska ortogonala polynom med deras viktfunktioner och ortogonalitetsintervall.

2.5 Gram-Schmidt processen

Gram-Schmidt processen är en teknik som tar en godtycklig bas (inte nödvändigtvis ortogonal eller med längd 1) och transformerar om den till en ortonormal bas.

Definition 2.8: Gram-Schmidts ortogonaliseringsprocess

Antag att $\{v_1, v_2, \dots, v_n\}$ är en linjärt oberoende mängd av vektorer i ett inre-produktrum. Då kan man konstruera en ortonormal bas $\{u_1, u_2, \dots, u_n\}$ genom att använda följande process:

$$\begin{aligned}u_1 &= \frac{v_1}{\|v_1\|} \\w_k &= v_k - \sum_{j=1}^{k-1} \langle v_k, u_j \rangle u_j \quad \text{för } k = 2, 3, \dots, n \\u_k &= \frac{w_k}{\|w_k\|}\end{aligned}$$

Då är mängden $\{u_1, u_2, \dots, u_n\}$ en ortonormal bas som spänner upp samma underrum som den ursprungliga mängden $\{v_1, v_2, \dots, v_n\}$.

Denna ide kan generaliseras till polynom. Man börjar med en uppsättning av linjärt oberoende polynom och använder Gram-Schmidt-processen för att ortogonalisera dem. Den enda skillnaden är att man inte behöver normalisera vilket betyder att man i första och sista steget av Gram-Schmidt-processen sätter $u_1(x) = v_1(x)$ och $u_k(x) = w_k(x)$ respektive. En klassiskt uppsättning är polynomen $1, x, x^2, x^3, \dots$

Sats 2.8: Existens av ortogonala polynom

Låt $[a, b]$ vara ett kompakt intervall och låt $w(x)$ vara en kontinuerlig funktion på $[a, b]$ sådan att $w(x) > 0$ för alla $x \in [a, b]$. Då existerar en sekvens av ortogonala polynom med avseende på vikten $w(x)$ och dessa kan konstrueras med hjälp av Gram-Schmidts ortogonaliseringsprocess. Detta visar att ortogonala polynom existerar i alla sådana situationer.

Bevis. Låt $w(x)$ vara en kontinuerlig funktion på $[a, b]$ sådan att $w(x) > 0$ för alla $x \in [a, b]$. Vi definierar en inre produkt på rummet av kontinuerliga funktioner som:

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

Betrakta en följd av monomer $\{1, x, x^2, x^3, \dots\}$ som är linjärt oberoende och spänner upp rummet av polynom på $[a, b]$. Vi tillämpar Gram-Schmidts or-

ortogonaliseringsprocess för att konstruera en sekvens av ortogonala polynom $\{\phi_0, \phi_1, \phi_2, \dots\}$ enligt:

$$\begin{aligned}\phi_0(x) &= 1, \\ \phi_1(x) &= x - \frac{\langle x, \phi_0 \rangle}{\langle \phi_0, \phi_0 \rangle} \phi_0(x), \\ \phi_2(x) &= x^2 - \frac{\langle x^2, \phi_0 \rangle}{\langle \phi_0, \phi_0 \rangle} \phi_0(x) - \frac{\langle x^2, \phi_1 \rangle}{\langle \phi_1, \phi_1 \rangle} \phi_1(x), \\ &\vdots \\ \phi_n(x) &= x^n - \sum_{k=0}^{n-1} \frac{\langle x^n, \phi_k \rangle}{\langle \phi_k, \phi_k \rangle} \phi_k(x).\end{aligned}$$

Varje $\phi_n(x)$ är ett polynom av grad n och ortogonalt mot alla tidigare $\phi_k(x)$ för $k < n$. Eftersom viktfunktionen $w(x)$ är positiv och kontinuerlig på $[a, b]$ är alla inre produkter väldefinierade och positiva där det krävs vilket garanterar att processen kan fortsätta för alla n . Det existerar således en följd av ortogonala polynom $\{\phi_n(x)\}$ på $[a, b]$ med avseende på viktfunktionen $w(x)$ konstruerade via Gram-Schmidts process. $\square$

Exempel 2.2

Låt oss använda Gram-Schmidt-processen för att ortogonalisera basen

$$v_1(x) = 1, \quad v_2(x) = x, \quad v_3(x) = x^2$$

med avseende på den inre produkten:

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x) dx$$

Steg 1:

$$u_1(x) = v_1(x) = 1$$

Steg 2: Ortogonalisera $v_2(x)$ mot $u_1(x)$:

$$\begin{aligned}\langle v_2(x), u_1(x) \rangle &= \int_{-1}^1 x \cdot 1 dx = 0 \\ u_2(x) &= v_2(x) - \frac{\langle v_2(x), u_1(x) \rangle}{\langle u_1(x), u_1(x) \rangle} u_1(x) = x - 0 = x\end{aligned}$$

Steg 3: Ortogonalisera $v_3(x)$ mot $u_1(x)$ och $u_2(x)$:

$$\langle v_3(x), u_1(x) \rangle = \int_{-1}^1 x^2 \cdot 1 dx = \frac{2}{3}$$

$$\langle v_3(x), u_2(x) \rangle = \int_{-1}^1 x^2 \cdot x \, dx = \int_{-1}^1 x^3 \, dx = 0$$

$$\langle u_1(x), u_1(x) \rangle = \int_{-1}^1 1^2 \, dx = 2$$

$$\langle u_2(x), u_2(x) \rangle = \int_{-1}^1 x^2 \, dx = \frac{2}{3}$$

$$u_3(x) = v_3(x) - \frac{\langle v_3(x), u_1(x) \rangle}{\langle u_1(x), u_1(x) \rangle} u_1(x) - \frac{\langle v_3(x), u_2(x) \rangle}{\langle u_2(x), u_2(x) \rangle} u_2(x)$$

$$u_3(x) = x^2 - \frac{\frac{2}{3}}{2} \cdot 1 - 0 = x^2 - \frac{1}{3}$$

Den ortogonala basen är således:

$$u_1(x) = 1, \quad u_2(x) = x, \quad u_3(x) = x^2 - \frac{1}{3}$$

Observera att detta är de tre första Legendrepolynom upp till en viss faktor.

3 Gauss-kvadratur

3.1 Bakgrund

I matematiken så relaterar ordet *kvadratur* till beräkning av arean av ett matematisk objekt. När man talar om *kvadratur* talar man främst om numerisk integration.

En grundläggande idé bakom numerisk integration är att man approximativt kan beräkna integraler av funktioner genom att man först approximerar funktionen med ett polynom och sedan integrerar detta polynom. Enligt *Stone–Weierstrass satsen* [5] kan varje kontinuerlig funktion på ett slutet intervall $[a, b]$ approximeras godtyckligt bra med ett polynom. Denna princip ligger i grunden för många kvadraturmetoder inklusive Gauss-kvadratur.

Sats 3.1: Stone–Weierstrass sats

För varje funktion $f \in C[a, b]$ (kontinuerlig på intervallet $[a, b]$) och för varje $\varepsilon > 0$ finns ett polynom $p(x)$ sådant att:

$$|f(x) - p(x)| < \varepsilon \quad \text{för alla } x \in [a, b].$$

Detta innebär att varje kontinuerlig funktion på ett slutet intervall kan approximeras godtyckligt väl med ett polynom.

Som nämnt ovan så kan inte alla funktioner integreras. Då är det lämpligt att använda numeriska metoder. Idag finns det flera numeriska metoder som *Trapetsregel* eller *Simpsons regel*. *Gauss kvadratur* är ännu en metod men den sticker ut ur mängden - den är nämligen otrolig noggran. Äldre metoder som *Trapetsregel* eller *Simpsons regel* delar upp integrerings intervallet i små jämna delintervall genom att välja ett visst antal noder. Ju mer delintervall man väljer desto mer exakt blir värdet på integralen. *Gauss kvadratur* fungerar på liknande sätt i termer av att välja delintervall, men metoden väljer noderna på ett smart sätt.

Gauss kvadratur ger exakta resultat för polynom upp till grad $2n - 1$ vilket betyder att om funktionen $f(x)$ som man vill integrera har grad som är mindre eller lika med än $2n - 1$ så kommer metoden producera en exakt lösning. Det visar sig att de mest optimala noderna är nollställena till ortogonala polynom.

Definition 3.1: Gauss-kvadratur

Gauss-kvadratur är en numerisk metod för att approximera en integral av formen:

$$\int_a^b f(x) w(x) dx \approx \sum_{i=1}^n w_i f(x_i) \quad (5)$$

Här är:

- $f(x)$ en kontinuerlig funktion som ska integreras,
- $w(x)$ en positiv viktfunction definierad på intervallet $[a, b]$,
- x_i är abscissorna (noder).
- w_i är tillhörande positiva vikter som väljs så att formeln är exakt för alla polynom av grad upp till $2n - 1$.

Noderna kan i princip väljas hur som helst, men detta påverkar naturligtvis noggrannheten. Målet är ju att välja noderna på ett sådant sätt att noggrannheten maximeras. Det är här teorin om ortogonala polynom kommer in. Säg att vi vill approximera integralen:

$$\int_{-1}^1 p(x) dx \approx \sum_{i=1}^n w_i f(x_i)$$

Där $p(x)$ är ett godtyckligt polynom med $\deg(p(x)) \leq 2n - 1$. Låt vidare $L_n(x)$ vara ett ortogonalt polynom av grad n . Då kan man utföra polynomdivision och skriva $p(x)$ som:

$$p(x) = q(x)L_n(x) + r(x).$$

Där $\deg(q(x)) \leq n - 1$ och $\deg(r(x)) < n$. Då kan integralen skrivas som:

$$\int_{-1}^1 p(x) dx = \int_{-1}^1 q(x)L_n(x) + r(x) dx = \int_{-1}^1 q(x)L_n(x) dx + \int_{-1}^1 r(x) dx.$$

Eftersom $q(x)$ har en lägre grad än $L_n(x)$ kommer integralen av deras produkt att blir noll och då får man:

$$\int_{-1}^1 p(x) dx = \int_{-1}^1 r(x) dx.$$

Om man väljer noderna $x_1, \dots, x_n$ som nollställen till $L_n(x)$ så blir:

$$p(x_i) = q(x_i)L_n(x_i) + r(x_i) = 0 + r(x_i) = r(x_i)$$

Det innebär att i de valda noderna är $p(x_i) = r(x_i)$ vilket ger:

$$\sum_{i=1}^n w_i p(x_i) = \sum_{i=1}^n w_i r(x_i).$$

Eftersom $r(x)$ har grad mindre än n kan vikterna w_i väljas så att summan $\sum w_i r(x_i)$ exakt beräknar $\int_{-1}^1 r(x) dx$ (se sats 3.2). Då får man:

$$\sum_{i=1}^n w_i p(x_i) = \int_{-1}^1 r(x) dx = \int_{-1}^1 p(x) dx.$$

Detta visar att kvadraturformeln är exakt för alla polynom $p(x)$ med grad högst $2n - 1$. Genom att välja noderna som nollställen till det ortogonala polynomet $L_n(x)$ uppnår man alltså maximal noggrannhet för ett givet antal punkter n .

Sats 3.2: Exakthet för Gauss-kvadratur

Gauss-kvadratur med n noder är exakt för alla polynom $f(x)$ av grad upp till $2n - 1$. Om $f(x)$ är ett polynom av grad högst $2n - 1$ gäller att:

$$\int_a^b f(x) w(x) dx = \sum_{i=1}^n w_i f(x_i). \quad (6)$$

där $w(x)$ är en positiv viktfunction och x_i, w_i är noder och vikter valda enligt Gauss-kvadraturregeln.

Bevis. Eftersom Gauss-kvadratur bygger på ortogonala polynom relativt viktfunctionen $w(x)$, väljs noderna $x_1, \dots, x_n$ som rötter till det ortogonala polynomet av grad n . En viktig egenskap hos ortogonala polynom är att varje polynom av grad mindre än $2n$ kan exakt interpoleras genom en linjärkombination av lägre gradiga ortogonala polynom.

Låt $f(x)$ vara ett polynom av grad högst $2n - 1$. Vi kan då skriva $f(x)$ som

$$f(x) = q(x)p_n(x) + r(x),$$

där $p_n(x)$ är det ortogonala polynomet av grad n , $q(x)$ är ett polynom av grad högst $n - 1$, och $r(x)$ är ett polynom av grad högst $n - 1$.

Integrerar vi $f(x)$ över intervallet $[a, b]$ får vi

$$\int_a^b f(x) w(x) dx = \int_a^b q(x)p_n(x) w(x) dx + \int_a^b r(x) w(x) dx.$$

Eftersom $p_n(x)$ är ortogonalt mot alla polynom av grad högst $n - 1$ med avseende på $w(x)$ och $q(x)$ har grad högst $n - 1$ gäller:

$$\int_a^b q(x)p_n(x) w(x) dx = 0.$$

Kvar blir:

$$\int_a^b f(x) w(x) dx = \int_a^b r(x) w(x) dx.$$

Polynomet $r(x)$ har grad högst $n-1$ och därför är det bestämt av de n värdena $r(x_i)$ t.ex. via Lagranges interpolationspolynom:

$$L_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j},$$

med formeln

$$r(x) = \sum_{i=1}^n r(x_i) L_i(x).$$

Om vi sätter detta uttryck för $r(x)$ i integralen får vi:

$$\int_a^b r(x) w(x) dx = \sum_{i=1}^n r(x_i) \int_a^b L_i(x) w(x) dx.$$

Vi definierar nu vikterna:

$$w_i = \int_a^b L_i(x) w(x) dx.$$

Så integralen blir:

$$\int_a^b r(x) w(x) dx = \sum_{i=1}^n r(x_i) w_i.$$

Eftersom $r(x_i) = f(x_i)$ (då $p_n(x_i) = 0$ vid rötterna) får vi:

$$\sum_{i=1}^n r(x_i) w_i = \sum_{i=1}^n f(x_i) w_i.$$

Alltså är:

$$\int_a^b f(x) w(x) dx = \sum_{i=1}^n w_i f(x_i),$$

vilket visar att Gauss-kvadratur är exakt för alla polynom av grad upp till $2n-1$.

□

Exempel 2.3

Säg att man vill approximera integralen:

$$\int_{-1}^1 x^2 dx$$

med Gauss-kvadratur och jämföra med det exakta värdet.

Analytiskt lösning:

$$\int_{-1}^1 x^2 dx = \left[\frac{x^3}{3} \right]_{-1}^1 = \frac{1^3}{3} - \frac{(-1)^3}{3} = \frac{1}{3} - \left(-\frac{1}{3} \right) = \frac{2}{3}$$

Gauss-kvadratur på intervallet $[-1, 1]$ med två noder ($n = 2$):

- Noder: $x_1 = -\frac{1}{\sqrt{3}}, \quad x_2 = \frac{1}{\sqrt{3}}$
- Vikter: $w_1 = w_2 = 1$

Approximationen blir:

$$\int_{-1}^1 x^2 dx \approx w_1 f(x_1) + w_2 f(x_2) = 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}$$

Observera att eftersom funktionen $f(x) = x^2$ är ett polynom av grad 2 och Gauss-kvadratur med $n = 2$ är exakt för alla polynom upp till grad $2n - 1 = 3$ erhålls ett exakt resultat.

En naturlig fråga som dyker upp är då vilka ortogonala polynom man ska använda? Det finns ju flera familjer av ortogonala polynom. Det är något som beror på viktfunktionen. Följande tabell visar de vanligaste familjer av ortogonala polynom och tillhörande viktfunktion:

Namn	Intervall	Viktfunktion
Legendre	$(-1, 1)$	1
Laguerre	$(0, \infty)$	$x^\alpha e^{-x}$
Hermite	$(-\infty, \infty)$	e^{-x^2}
Chebyshev	$(-1, 1)$	$(1 - x^2)^{-1/2}$
Jacobi	$(-1, 1)$	$(1 - x)^\alpha (1 + x)^\beta$

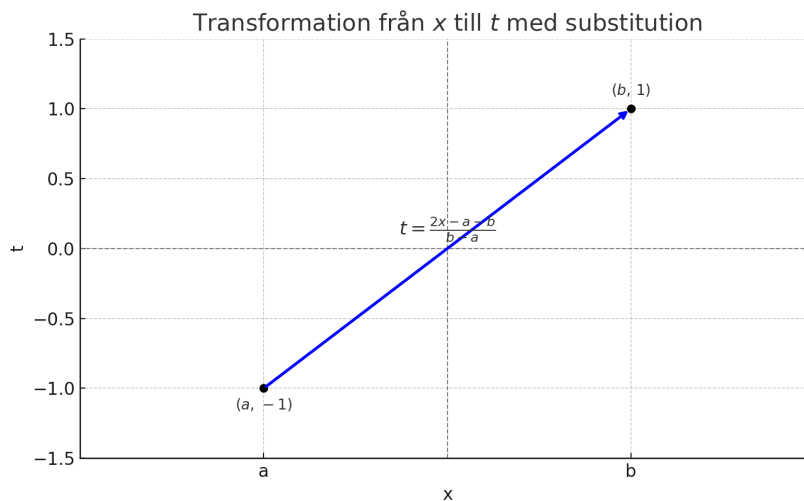
Tabell 2: Översikt över ortogonala polynom och deras viktfunktioner

3.2 Gauss-Legendre-kvadratur

Den mest klassiska varianten av *Gauss-kvadratur* är *Gauss-Legendre-kvadratur*. Anledningen till detta är att viktfunktionen $w(x) = 1$ är vanligast för beräkning av godtyckliga integraler. Genast uppstår det ett problem - *Gauss-Legendre-kvadratur* kräver att integreringsintervallen är $[-1, 1]$. Man kan komma runt detta genom att transformera ett godtyckligt intervall $[a, b]$ till $[-1, 1]$. Detta sker genom att man introducerar en ny variabel t där sambandet mellan x och t är en linjär transformation.

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

Denna transformation gör så att när $x = a$ så är $t = -1$ och när $x = b$ så är $t = 1$. På det sättet skapar man en linjär avbildning $[a, b] \rightarrow [-1, 1]$.



Figur 3: Transformationen från x till t

Definition 3.2: Transformation av intervallet för Gauss-Legendre-kvadratur

$$\int_a^b f(x) dx = \frac{b - a}{2} \int_{-1}^1 f\left(\frac{b - a}{2}v + \frac{b + a}{2}\right) dv \quad (6)$$

Exempel 2.4

Vi vill approximera integralen

$$\int_2^6 3x^4 - 2x^2 + x \, dx$$

med hjälp av Gauss-Legendre-kvadratur med 3 noder. Gauss-Legendre-kvadratur gäller på intervallet $[-1, 1]$ så vi gör först en ändring av variabel: Definera en ny variabel t sådan att:

$$x = \frac{b-a}{2}t + \frac{a+b}{2} = \frac{6-2}{2}t + \frac{6+2}{2} = 2t + 4$$

Därmed kan vi skriva om integralen som:

$$\int_2^6 3x^4 - 2x^2 + x \, dx = 2 \cdot \int_{-1}^1 3(2t+4)^4 - 2(2t+4)^2 + (2t+4) \, dt$$

Noder och vikter för $n = 3$ på intervallet $[-1, 1]$ är:

$$\begin{array}{ll} t_1 = -\sqrt{\frac{3}{5}}, & w_1 = \frac{5}{9} \\ t_2 = 0, & w_2 = \frac{8}{9} \\ t_3 = \sqrt{\frac{3}{5}}, & w_3 = \frac{5}{9} \end{array}$$

Vi räknar ut funktionens värden i de transformerade punkterna:

$$x_i = 2t_i + 4$$

Sedan approximeras integralen som:

$$\int_2^6 3x^4 - 2x^2 + x \, dx \approx 2 \cdot \sum_{i=1}^3 w_i \cdot 3x_i^4 - 2x_i^2 + x_i$$

Gauss-Legendre-approximation: 4523.7333

Analytisklösning lösning: 4523.7333

3.3 Hur hittar man vikter och noder?

Det finns två metoder för att explicit kunna hitta noder och vikter för *Gauss-kvadratur*. Antingen kan man ställa upp ett ekvationsystem eller använda nollställen till ortogonala polynom. För varje viktfunction $w(x)$ så finns det tillhörande familj av ortogonala polynom. För att göra det enklare kommer vi att undersöka *Legendre-polynom*. Låt oss kika på hur båda metoder fungerar och för enkelhets skull låt oss betrakta *Gauss-Legendre-kvadratur*.

Metoden med ekvationsystem:

Målet är att hitta sådana w_i och x_i att:

$$\int_{-1}^1 f(x)w(x) dx \approx \sum_{i=1}^n w_i f(x_i).$$

är exakt för alla polynom upp till grad $2n - 1$. Låt oss titta på fallet där man använder 2 noder. Det betyder att vi söker w_i och x_i sådana att:

$$\int_{-1}^1 f(x)w(x) dx = \sum_{i=1}^2 w_i f(x_i) = w_1 f(x_1) + w_2 f(x_2).$$

Enligt *sats 3.2* om Gauss-kvadraturens exakthet gäller att en kvadraturformel med n noder ger exakt resultat för alla polynom av grad upp till $2n - 1$. För att uppnå detta krav konstrueras ett ekvationssystem där man kräver att kvadraturen ger exakt resultat för varje monom x^k för $k = 0, 1, \dots, 2n - 1$.

Detta är precis vad som görs i metoden med ekvationssystem: vi ställer upp ett system med $2n$ ekvationer och $2n$ obekanta (noder x_i och vikter w_i) genom att kräva:

$$\int_{-1}^1 x^k dx = \sum_{i=1}^n w_i x_i^k, \quad \text{för } k = 0, 1, \dots, 2n - 1.$$

Denna metod är ett rent algebraiskt sätt att bestämma vikter och noder, i stället för att använda integraler och viktfunctioner direkt. Det kan dessutom innebära en viss fördel ur beräkningssynpunkt, eftersom man bara behöver lösa ett linjärt ekvationssystem. Observera att varje polynom kan skrivas som en linjär kombinationer av monomer, det vill säga:

$$f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_k x^k.$$

Observera att varje polynom kan skrivas som en linjär kombinationer av monomer och kvadraturet skall ge ett exakt resultat för varje polynom av grad $2n - 1$. Detta innebär att för att säkerställa detta räcker det att man kräver exaktheten för en uppsättning av monomer $1, x, x^2, x^3, \dots$ Upp till grad $2n - 1$. Detta fungerar eftersom om kvadraturen ger ett exakt resultat för varje sådan

monom kommer den också ge ett exakt resultat för summan av dessa monomer. Ekvationsystemet blir således följande:

$$\begin{cases} \int_{-1}^1 x^0 dx = w_1 x_1^0 + w_2 x_2^0 \\ \int_{-1}^1 x^1 dx = w_1 x_1^1 + w_2 x_2^1 \\ \int_{-1}^1 x^2 dx = w_1 x_1^2 + w_2 x_2^2 \\ \int_{-1}^1 x^3 dx = w_1 x_1^3 + w_2 x_2^3 \end{cases}$$

$$\begin{cases} 2 = w_1 + w_2 \\ 0 = w_1 x_1 + w_2 x_2 \\ \frac{2}{3} = w_1 x_1^2 + w_2 x_2^2 \\ 0 = w_1 x_1^3 + w_2 x_2^3 \end{cases}$$

Lösningen till detta systemet ges av:

$$\begin{cases} x_1 = -\frac{1}{\sqrt{3}} \\ x_2 = \frac{1}{\sqrt{3}} \\ w_1 = 1 \\ w_2 = 1 \end{cases}$$

Metoden med ekvationsystem är intuitiv, men inte så praktisk då ekvationer är inte linjära och antalet ekvationer växer väldigt fort ju fler noder man använder. Vi har tidigare visat att noderna ges av nollställena till ortogonala polynom. För att beräkna vikterna har man stor nytta att följande sats:

Sats 3.3: Vikter för Legendrepolytom

Antag att $P_n(x)$ är Legendrepolytommet av grad n och att $x_1, x_2, \dots, x_n$ är dess nollställena. Då ges vikterna A_k i Gauss-kvadratur enligt formeln:

$$A_k = \frac{1}{P'_n(x_k)} \int_{-1}^1 \frac{P_n(x)}{x - x_k} dx. \quad (7)$$

Bevis. Vi utgår från att $x_1, \dots, x_n$ är nollställen till Legendrepolynomet $P_n(x)$, och att $f(x)$ approximeras med ett Lagrange-interpolationspolynom:

$$P(x) = \sum_{k=1}^n L_k(x) f(x_k)$$

där $L_k(x)$ är de klassiska Lagrangepolynom.

Eftersom $f(x_k)$ är konstant med avseende på x kan vi skriva:

$$\int_{-1}^1 f(x) dx \approx \sum_{k=1}^n f(x_k) \int_{-1}^1 L_k(x) dx.$$

Detta leder till definitionen av vikterna:

$$A_k = \int_{-1}^1 L_k(x) dx.$$

Vi vet att $P_n(x)$ är ett polynom med nollställen i $x_1, \dots, x_n$ så:

$$P_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n).$$

Om vi kallar detta $F(x)$ så kan vi skriva:

$$L_k(x) = \frac{F(x)}{(x - x_k)F'(x_k)}$$

Uttrycket $F(x)/(x - x_k)$ är ett polynom som innehåller alla faktorer utom $(x - x_k)$ dvs:

$$\frac{F(x)}{x - x_k} = (x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n).$$

Dess värde i $x = x_k$ är just $F'(x_k)$ eftersom derivatan i en enkel nollpunkt ger produkten av de övriga faktorerna.

Därför får vi:

$$L_k(x) = \frac{P_n(x)}{(x - x_k)P'_n(x_k)}.$$

Vi sätter in detta i uttrycket för vikten:

$$A_k = \int_{-1}^1 L_k(x) dx = \frac{1}{P'_n(x_k)} \int_{-1}^1 \frac{P_n(x)}{x - x_k} dx.$$

□

Sats 3.3 är viktig, men inte så praktisk då den innehåller en integral. Det är möjligt att, genom att utnyttja egenskaper hos Legendre-polynomen att härleda en mer explicit form för (7). Detta presenteras som en sats och en fullständig härledning av formeln ges i [4].

Sats 3.4: Explicit formel för vikterna

Vikterna i Gauss-kvadratur kan också uttryckas som:

$$A_k = \frac{2}{(1 - x_k^2) [P'_n(x_k)]^2}, \quad (8)$$

där x_k är nollställena till Legendrepolynomet $P_n(x)$.

Metoden med ortogonala polynom:

Låt oss säga att vi arbetar med *Gauss-Legendre-kvadratur*. Söker således vikter w_i och noder x_i sådana att:

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^n w_i f(x_i)$$

Det tillhörande familjen av ortogonala polynom är således *Legendre-polynom*. Vi säger att vi vill använda 2 noder. Då kommer noderna ges av nollställena till det Legendre polynomet som har grad 2. I vårtfall blir det nollställena till polynomet:

$$P_2(x) = \frac{1}{2}(3x^2 - 1) = 0 \implies x_1 = \frac{1}{\sqrt{3}}, x_2 = -\frac{1}{\sqrt{3}}$$

Därefter hittas vikterna w_i genom:

$$w_i = \frac{2}{(1 - x_i^2) [P'_n(x_i)]^2}$$

$$w_1 = \frac{2}{\left(1 - \left(-\frac{1}{\sqrt{3}}\right)^2\right) \left[P'_2\left(-\frac{1}{\sqrt{3}}\right)\right]^2} = \frac{2}{\left(1 - \frac{1}{3}\right) (-\sqrt{3})^2} = \frac{2}{\frac{2}{3} \cdot 3} = \frac{2}{2} = 1$$

$$w_2 = \frac{2}{\left(1 - \left(\frac{1}{\sqrt{3}}\right)^2\right) \left[P'_2\left(\frac{1}{\sqrt{3}}\right)\right]^2} = \frac{2}{\left(1 - \frac{1}{3}\right) (\sqrt{3})^2} = \frac{2}{\frac{2}{3} \cdot 3} = \frac{2}{2} = 1$$

Observera att när man i teorin kan använda bägge metoder så är metoden med ekvationsystem mer intuitiv men samtidigt mer krävande då man får för stora värden på n lösa stora icke-linjära ekvationsystem som ofta kräver numeriska beräkningar. Metoden med ortogonala polynom är mindre intuitiv men är betydligt smidigare att använda. Notera att beräknar av noder och vikter för olika värden på n är universiell i den meningen att dessa kvantiteter är oberoende av funktionen som man vill integrera eller integreringsintervallet. Av den anledning är det praktiskt att tabulera dessa värden.

Tabell 3: Vikter och noder för Gauss–Legendre-kvadratur på intervallet $[-1, 1]$

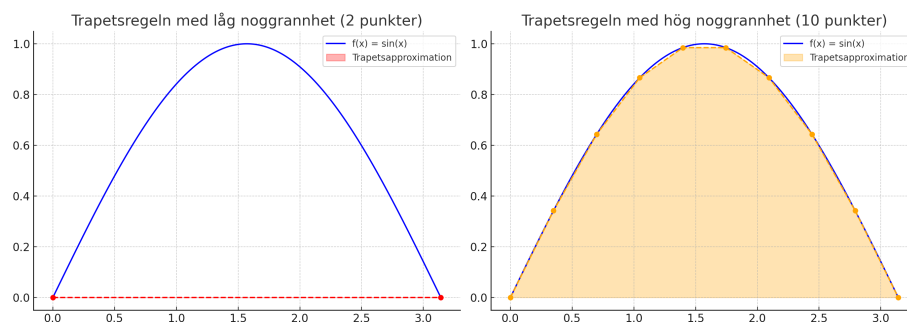
n	i	w_i	x_i
2	1	1.000000000000000	-0.5773502691896257
	2	1.000000000000000	0.5773502691896257
3	1	0.5555555555555556	-0.7745966692414834
	2	0.8888888888888888	0.0000000000000000
	3	0.5555555555555556	0.7745966692414834
4	1	0.3478548451374538	-0.8611363115940526
	2	0.6521451548625461	-0.3399810435848563
	3	0.6521451548625461	0.3399810435848563
	4	0.3478548451374538	0.8611363115940526
5	1	0.2369268850561891	-0.9061798459386640
	2	0.4786286704993665	-0.5384693101056831
	3	0.5688888888888889	0.0000000000000000
	4	0.4786286704993665	0.5384693101056831
	5	0.2369268850561891	0.9061798459386640
6	1	0.1713244923791704	-0.9324695142031521
	2	0.3607615730481386	-0.6612093864662645
	3	0.4679139345726910	-0.2386191860831969
	4	0.4679139345726910	0.2386191860831969
	5	0.3607615730481386	0.6612093864662645
	6	0.1713244923791704	0.9324695142031521

3.4 Jämförelse med traditionella numeriska metoder

En av de simplaste metoder i numerisk integration är *Trapetsregel* [2, s. 229]. Man approximerar hela arean under kurvan på $[a, b]$ genom att dra en linje från $a, f(a)$ till $b, f(b)$ och räknar ut arean som bildas. Högre noggrannhet uppnås genom att man använder fler delintervall.

Definition 3.3: Trapetsregeln

$$\int_a^b f(x) dx \approx \frac{b-a}{2} (f(a) + f(b))$$

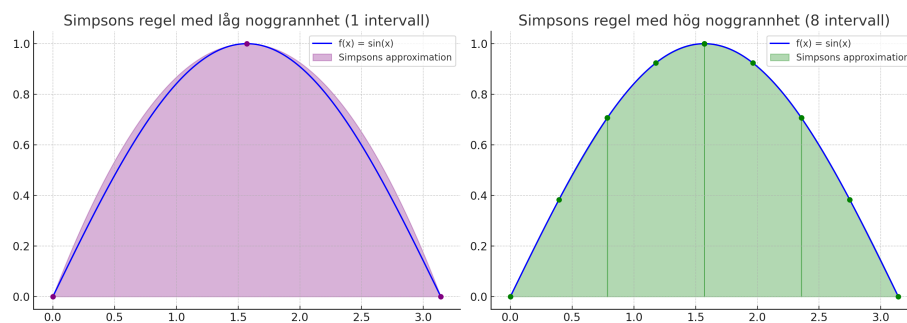


Figur 4: Trapetsregeln med låg noggrannhet(vänster) och hög noggrannhet(höger).

En annan metod för numerisk integrering är *Simpsonsregeln* [2, s. 230]. Till skillnad från trapetsregeln så försöker man att anpassa en andragradspolynom mellan tre punkter. Av den anledningen kräver simpsonsregel att antalet punkter är udda och antalet delintervall blir således jämnt.

Definition 3.4: Simpsons regel

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right)$$



Figur 5: Simpsonsregeln med låg noggrannhet(vänster) och hög noggrannhet(höger).

För att illustrera skillnaderna mellan trapetsregeln, Simpsonsregeln och Gausskvadratur på en funktion som saknar en primitiv funktion. Syftet är att visa hur noggrant varje metod kan approximera integralen med ett begränsat antal noder.

Trapetsregeln och Simpsons regel är klassiska metoder som bygger på att approximera funktionen med linjära respektive andragradspolynom. Dessa meto-

der kräver ofta ett stort antal delintervall för att uppnå hög noggrannhet.

Gauss-kvadratur utnyttjar däremot både optimalt valda noder och vikter och kan därför nå mycket hög noggrannhet även med få utvärderingspunkter.

För att konkret jämföra metoderna beräknas nedan en integral av funktionen $x^6 - x^2 \sin(2x)$ på intervallet $[1, 3]$ [?, s. 233] Resultaten presenteras i en tabell där antalet punkter hålls lika för alla metoder.

Antal punkter	Trapetsregel	Simpsons regel	Gauss-kvadratur
2	731.605442	333.238094	306.819934
4	369.725593	318.414648	317.345390
6	336.398103	317.557676	317.344247

Tabell 4: Approximativ beräkning av $\int_1^3 (x^6 - x^2 \sin(2x)) dx$. Det exakta värdet är 317.3442466.

4 Slutsats

I detta arbete har vi studerat teorin som ligger bakom ortogonala polynom och hur denna teori används i utvecklingen av Gauss-kvadratur för numerisk integration. Vi har sett att ortogonala polynom har en mängd kraftfulla egenskaper som möjliggör effektiv och i många fall exakt approximation av polynom.

Sammanfattningsvis erbjuder ortogonala polynom och Gauss-kvadratur en elegant och kraftfull verktygslåda för att lösa integraler inom många områden av matematiken och naturvetenskapen. Som ett slutsats kan läsaren se en enkel implementation av gauss kvadratur i Python [6]. Det hela sköts av biblioteket numpy, men förhoppningsvis förstår läsaren vad som händer bakom scenen.

Listing 1: Enkel implementation av Gauss-Legendre-kvadratur i Python

```
import numpy as np
from numpy.polynomial.legendre import leggauss

# Funktion som man vill integrera
def f(x):
    return x**2

# Antal noder
n = 2

# F r noder och vikter
nodes, weights = leggauss(n)

# Ber kna Gauss-Legendre-approximation p intervallet [-1, 1]
approx_integral = sum(w * f(x) for x, w in zip(nodes, weights))

print(f"Approximativt v rde p integralen: {approx_integral}")
```

Referenser

- [1] Sven Erick Alm and Tom Britton. *Stokastik : sannolikhets teori och statistik teori med tillämpningar*. Liber, Stockholm, 1. uppl. edition, 2008.
- [2] Richard L. Burden, J. Douglas Faires, and Annette M. Burden. *Numerical analysis*. Cengage Learning, Australia, tenth edition edition, 2016.
- [3] Walter. Gautschi. *Numerical Analysis [Elektronisk resurs]*. Birkhäuser Boston, Boston, second edition. edition, 2012.
- [4] Francis B. Hildebrand. *Introduction to Numerical Analysis*. Dover Publications, New York, 2nd edition, 1974.
- [5] LibreTexts. Legendre polynomials, 2025. Hämtad: 26 april 2025.
- [6] NumPy Developers. `numpy.polynomial.legendre.leggauss`. NumPy, 2024. Tillgänglig på <https://numpy.org/doc/2.1/reference/generated/numpy.polynomial.legendre.leggauss.html>.
- [7] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition, 1976.