UiO : Matematisk institutt
Det matematisk-naturvitenskapelige fakultet

Kapittel 7: Indreproduktrom

MAT1060: Matematikk for anvendelser 2

Ole Fredrik Brevig

Versjon 1



Innhold

Innhold i

7 Indreproduktrom 1
7.1 Definisjon og eksempler 3
7.2 Norm og geometri 7
7.3 Ortogonal dekomposisjon og projeksjon 8
7.4 Fourier-rekker 11

KAPITTEL 7

Indreproduktrom

Det er natt i stoveglaset,
men dei gamle har sin gud.
Og vi to, vi to har kvarandre,
la meg ha deg heilt til slutt.

Saft - Fjøsvise

Vi starter med litt repetisjon fra Kapittel 1 og Kapittel 3 hvor vi betraktet det Euklidske rommet $\mathbb{R}^n$ utstyrt med tre binære operasjoner:

- *Addisjon* av to vektorer $\mathbf{u}$ og $\mathbf{v}$ som gir *vektoren* $\mathbf{u}+\mathbf{v}$.
- *Skalarmultiplikasjon* av et tall c og en vektor $\mathbf{u}$ som gir *vektoren* $c\mathbf{u}$.
- *Prikkprodukt* av to vektorer $\mathbf{u}$ og $\mathbf{v}$ som gir *tallet* $\mathbf{u}\cdot\mathbf{v}$.

De to første operasjonene oppfyller følgende egenskaper:

(i) Addisjon er *assosiativ*, slik at $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.

(ii) Addisjon er *kommutativ*, slik at $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.

(iii) Det finnes en additiv *identitet* $\mathbf{0}$ som oppfyller $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$.

(iv) For hver vektor $\mathbf{u}$ finnes en additiv invers $-\mathbf{u}$ slik at $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.

(v) Skalarmultiplikasjon er *assosiativ*: $a(b\mathbf{u})=(ab)\mathbf{u}$.

(vi) Det finnes *multiplikativ* identitet 1 som oppfyller $1\mathbf{u}=\mathbf{u}$.

(viii) Den distributive lovene

$$(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u} \qquad \text{og} \qquad a(\mathbf{u}+\mathbf{v})=a\mathbf{u}+a\mathbf{v}$$

holder.

Den additive identiten er nullvektorn (vektoren der alle komponentene er 0) og den multiplikative identiten er tallet 1.

Prikkproduktet oppfyller følgende egenskaper:

(ix) *Lineæritet* i hver variabel

$$(a\mathbf{u}+b\mathbf{v})\cdot\mathbf{w} = a\mathbf{u}\cdot\mathbf{w}+b\mathbf{v}\cdot\mathbf{w} \qquad \text{og} \qquad \mathbf{u}\cdot(a\mathbf{v}+b\mathbf{w}) = a\mathbf{u}\cdot\mathbf{v}+b\mathbf{u}\cdot\mathbf{w}.$$

(x) *Symmetrisk*, slik at $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$.

(xi) *Positivt definitt*, det vil si at at

$$\mathbf{u}\cdot\mathbf{u} \geq 0$$

med likhet hvis og bare hvis $\mathbf{u} = 0$.

Her er er antaglesen (xi) spesielt viktig, fordi det lar oss regne ut *normen* (eller lengden) til en vektor ved

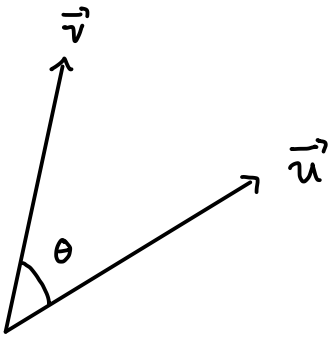
$$\|\mathbf{u}\| = \sqrt{\mathbf{u}\cdot\mathbf{u}}.$$

Merk at det er bare nullvektoren som har norm 0.

Prikkproduktet er også veldig nyttig fra et geometrisk perspektiv: To vektorer $\mathbf{u} \neq \mathbf{0}$ og $\mathbf{v} \neq \mathbf{0}$ står ortogonalt på hverandre hvis og bare hvis $\mathbf{u}\cdot\mathbf{v} = 0$. Mer generelt har vi at *vinkelen* θ mellom $\mathbf{u}$ og $\mathbf{v}$ oppfyller

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\mathbf{v}\|}.$$

Se Figur 7.1 for en illustrasjon som jeg har tegnet selv.



Figur 7.1: Vinkelen mellom to vektorer i $\mathbb{R}^2$.

7.1 Definisjon og eksempler

Formålet med dette kapittelet er å ta se hvordan egenskapene til $\mathbb{R}^n$ gjengitt ovenfor leder til en abstrakt definisjon som gir opphav til flere interessante eksempler med oppførsel lignende $\mathbb{R}^n$.

Definisjon. Et (reelt) *indreproduktrom* V er en mengde V med tre binære operasjoner:

- *Addisjon* $+\colon V \times V \to V$ og vi skriver $\mathbf{u} + \mathbf{v}$.
- *Skalarmultiplikasjon* $\colon \mathbb{R} \times V \to V$ og vi skriver $c\mathbf{u}$.
- *Indreprodukt* $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ og vi skriver $\langle \mathbf{u}, \mathbf{v} \rangle$.

De binære operasjonene oppfyller:

(i) Addisjon er *assosiativ*, slik at $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

(ii) Addisjon er *kommutativ*, slik at $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(iii) Det finnes en additiv *identitet* $\mathbf{0}$ som oppfyller $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$.

(iv) For hver vektor $\mathbf{u}$ finnes en additiv invers $-\mathbf{u}$ slik at $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

(v) Skalarmultiplikasjon er *assosiativ*: $a(b\mathbf{u}) = (ab)\mathbf{u}$.

(vi) Det finnes *multiplikativ* identitet 1 som oppfyller $1\mathbf{u} = \mathbf{u}$.

(viii) Den distributive lovene

$$(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u} \qquad \text{og} \qquad a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$$

holder.

(ix) Indreproduktet er *lineært* i hver variabel

$$\langle a\mathbf{u}+b\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{w} \rangle \qquad \text{og} \qquad \langle \mathbf{u}, a\mathbf{v}+b\mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle + b\langle \mathbf{u}, \mathbf{w} \rangle.$$

(x) Indreproduktet er *symmetrisk*, slik at $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

(xi) Indreproduktet er *positivt definitt*, det vil si at at

$$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$$

med likhet hvis og bare hvis $\mathbf{u} = 0$.

Vi definerer også *normen* (eller lengden) til en vektor $\mathbf{u}$ i V ved

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

Vi har valgt definisjonen på indreproduktrom slik at vi kan ta $V = \mathbb{R}^n$

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix}, \qquad c\mathbf{u} = c \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{pmatrix},$$

og indreprodukt definert på samme måte som prikkproduktet

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{j=1}^{n} u_j v_j.$$

I tillegg til dette skal vi se på to andre eksempler. For å definere den første trenger vi en definisjon som omhandler kvadratiske matriser.

Definisjon. La $A = (a_{jk})_{j,k=1}^n$. Da er *trasen* til A definert ved

$$\operatorname{tr} A = \sum_{j=1}^{n} a_{jj}.$$

Altså er trasen summen av tallene på diagonalen av matrisen.

Eksempel 7.1.1. Vi har

$$\operatorname{tr} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9 = 15.$$

Før vi går videre til neste eksempel, så minner vi at om dersom A og B er $m \times n$-matriser, så er $A^T B$ en $n \times n$-matrise. Det er viktig for at følgende definsjon gir mening.

Eksempel 7.1.2. La m og n være positive heltall og la V være mengden av alle $m \times n$-matriser. Da blir V et indreproduktrom med matriseaddisjon $A + B$, skalarmultiplikasjon cA og indreproduktet

$$\langle A, B \rangle = \operatorname{tr} A^T B.$$

Hvis $A = \begin{pmatrix} \mathbf{a}_1 \,|\, \mathbf{a}_2 \,|\, \cdots \,|\, \mathbf{a}_n \end{pmatrix}$ og $B = \begin{pmatrix} \mathbf{b}_1 \,|\, \mathbf{b}_2 \,|\, \cdots \,|\, \mathbf{b}_n \end{pmatrix}$, så er

$$\operatorname{tr} A^T B = \mathbf{a}_1 \cdot \mathbf{b}_1 + \mathbf{a}_2 \cdot \mathbf{b}_2 + \cdots + \mathbf{a}_n \cdot \mathbf{b}_n,$$

så i dette indreproduktet regner vi regner ut summen av prikkproduktet av radene. Dette indreproktet kalles ofte *Frobenius*-indreproduktet og normen

$$\|A\| = \sqrt{\langle A, A \rangle}$$

kalles *Hilbert–Schmidt* normen.

Det er hverken vanskelig eller givende å sjekke at Frobenius-indreproduktet faktisk oppfyller kravene (ix), (x) og (xi), så vi overlater dette prosjektet til den spesielt interesserte leser.

Eksempel 7.1.3. La $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ og $B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Da blir

$$\langle A, B \rangle = \operatorname{tr} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \operatorname{tr} \begin{pmatrix} 1 & 4 & 1 \\ 2 & 5 & 2 \\ 3 & 6 & 3 \end{pmatrix} = 1 + 5 + 3.$$

Det er lettere å regne $\mathbf{a}_1 \cdot \mathbf{b}_1 = 1$, $\mathbf{a}_2 \cdot \mathbf{b}_2 = 5$ og $\mathbf{a}_3 \cdot \mathbf{b}_3 = 3$. På tilsvarende måte kan vi regne

$$\|A\| = \sqrt{\langle A, A \rangle} = \sqrt{1^2 + 4^2 + 2^2 + 5^2 + 6^2} = \sqrt{91}.$$

Vi definerer *lineær uavhengighet*, *basis* og *dimension* på nøyaktig samme måte som for $\mathbb{R}^n$.

Eksempel 7.1.4. La V være indreproduktrommet av 2×3-matriser. Da er

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}$$

en basis for V og, følgelig, er $\dim V = 6$.

På mange måter kan vi tenke på V fra Eksempel 7.1.4 som $\mathbb{R}^6$, bare at vi har skrevet om vektorene

$$(a_1, a_2, a_3, a_4, a_5, a_6) \qquad \text{til matrisene} \qquad \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \end{pmatrix}.$$

Det kan være gode grunner til å gjøre dette, men for oss virker det kanskje litt underlig. Vi går derfor raskt videre til det viktigste eksempelet i dette kapittelet.

Eksempel 7.1.5. La V være mengden av alle kontinuerlige funksjoner på intervallet $[a, b]$ og definerer addisjon og skalarmultiplikasjon ved

$$(f + g)(x) = f(x) + g(x) \qquad \text{og} \qquad (cf)(x) = cf(x).$$

Her er nullvektoren funksjonen som er lik 0 for alle x på $[a, b]$. Det er ikke vanskelig å sjekke at kravene (i)–(viii) holder. Indreproduktet på V er

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

Her følger (ix) av egenskapene til integralet, (x) siden $f(x)g(x) = g(x)f(x)$ for alle x på $[a, b]$. Til slutt har vi

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

Dette er åpenbart større enn eller lik 0. Siden f er kontinuerlig, kan integralet bare være lik 0 om f er lik 0 for alle x på $[a, b]$. Altså er f nullvektoren.

Vi skriver ofte $C([a, b])$ for indreproduktrommet i Eksempel 7.1.5. Hva er dimensjonen til dette rommet? Det er ikke så vanskelig å sjekke at mengden

$$\{1, x, x^2, \ldots, x^n\}$$

er lineært uavhengig. Det er mulig å bruke Taylors teorem (fra MAT1050) eller man kan argumentere som følger:

- Det er klart at 1 og x er lineært uavhengige.
- Hvis x^2 kunne skrives som en lineærkombinasjon av $\{1, x\}$ ville alle andregradspolynomer være lik et førstegradspolynom på $[a, b]$. Det er åpenbart ikke mulig.
- Hvis x^n kunne skrives som en lineærkombinasjon av $\{1, x, x^2, \ldots, x^{n-1}\}$ så ville alle polynomer av grad n være lik et polynom av grad $n - 1$ på $[a, b]$, som igjen ikke er mulig.

Siden vi for hvert heltall n har en lineært uavhengig undermengde av $C([a, b])$ som inneholder n elementer er det naturlig å si at $C([a, b])$ har *uendelig dimensjon*. Vi skal ikke gå inn på hva det vil si å være en *basis* for et uendelig-dimensjonalt rom, da dette omhandler konvergens og kan bli ganske vanskelig.

Eksempel 7.1.6. Betrakt $f(x) = \sin x$ og $g(x) = \cos x$ som vektorer i $C([-\pi, \pi])$. Da er

$$\langle f, g \rangle = \int_{-\pi}^{\pi} \sin x \cos x \, dx = \int_{-\pi}^{\pi} \frac{\sin 2x}{2} \, dx = 0.$$

Videre er

$$\|f\| = \sqrt{\int_{-\pi}^{\pi} \sin^2 x \, dx} = \sqrt{\int_{-\pi}^{\pi} \frac{1 - \cos 2x}{2} \, dx} = \sqrt{\pi},$$
$$\|g\| = \sqrt{\int_{-\pi}^{\pi} \cos^2 x \, dx} = \sqrt{\int_{-\pi}^{\pi} \frac{1 + \cos 2x}{2} \, dx} = \sqrt{\pi}.$$

7.2 Norm og geometri

Vi skal nå se at kravene (ix)–(xi) på indreproduktet og definisjonen av normen som

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

gjør at normen i et indreproduktrom oppfyller mange av de samme egenskapene som vi kjenner igjen fra lengdebegrepet i $\mathbb{R}^2$ og $\mathbb{R}^3$. Det viktigste resultatet i den sammenheng er som følger.

Teorem 7.2.1 (Cauchy–Schwarz ulikhet)**.** *Dersom* $\mathbf{u}$ *og* $\mathbf{v}$ *er vektorer i et indreproduktrom* V*, så er*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Bevis. Hvis enten $\mathbf{u} = \mathbf{0}$ eller $\mathbf{v} = \mathbf{0}$ er begge sider lik 0. La oss derfor anta $\mathbf{v} \neq \mathbf{0}$. La c være et tall som vi bestemmer senere. Vi regner

$$\begin{aligned} 0 \leq \|\mathbf{u} - c\mathbf{v}\|^2 &= \langle \mathbf{u} - c\mathbf{v}, \mathbf{u} - c\mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} - c\mathbf{v} \rangle - c\langle \mathbf{v}, \mathbf{u} - c\mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{u} \rangle - c\langle \mathbf{u}, \mathbf{v} \rangle - c\langle \mathbf{v}, \mathbf{u} \rangle + c^2 \langle \mathbf{v}, \mathbf{v} \rangle \\ &= \|\mathbf{u}\|^2 - 2c\langle \mathbf{u}, \mathbf{v} \rangle + c^2 \|\mathbf{v}\|^2. \end{aligned}$$

Hvis vi velger $c = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}$ så blir dette

$$0 \leq \|\mathbf{u}\|^2 - \frac{\big(\langle \mathbf{u}, \mathbf{v} \rangle\big)^2}{\|\mathbf{v}\|^2} \qquad \Longleftrightarrow \qquad \big(\langle \mathbf{u}, \mathbf{v} \rangle\big)^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2.$$

Det gjenstår bare å ta kvadratroten. □

Siden cosinus kun tar verdier mellom -1 og 1 er ulikheten i Teorem 7.2.1 en nødvendighet for at følgende definisjon skal gi mening.

Definisjon. La $\mathbf{u}$ og $\mathbf{v}$ være to vektorer fra et indreproduktrom V. Da er *vinkelen* θ mellom $\mathbf{u}$ og $\mathbf{v}$ definert ved

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

Vi sier at vektorene $\mathbf{u}$ og $\mathbf{v}$ er *ortogonale* (eller står vinkelrett på hverandre) dersom $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

I Eksempel 7.1.6 så vi at $f(x) = \sin x$ og $g(x) = \cos x$ er ortogonale i $C([a, b])$. Vi har også følgende konsekvens av Cauchy–Schwarz ulikhet.

Teorem 7.2.2 (Trekantulikheten)**.** *Hvis* $\mathbf{u}$ *og* $\mathbf{v}$ *er vektorer i et indreproduktrom* V*, så er*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Bevis. Vi regner først

$$\|\mathbf{u}+\mathbf{v}\|^2 = \langle \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{v}\rangle = \langle \mathbf{u},\mathbf{u}\rangle + 2\langle \mathbf{u},\mathbf{v}\rangle + \langle \mathbf{v},\mathbf{v}\rangle \leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2,$$

hvor vi brukte Teorem 7.2.1. Vi bruker så regelen $a^2 + 2ab + b^2 = (a+b)^2$ for å skrive

$$\|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = \big(\|\mathbf{u}\| + \|\mathbf{v}\|\big)^2.$$

Det gjenstår bare å ta kvadratroten. □

Beviset for de to neste resultatene følger de samme linjene, men er noe enklere. Vi lar derfor den interesserte leseren fylle inn detaljene (eller slå det opp i Kapittel 7 i *Kompendium, MAT1060*).

Teorem 7.2.3 (Pytagoras)**.** *Hvis* $\mathbf{u}$ *og* $\mathbf{v}$ *er to ortogonale vektorer i et indreproduktrom* V*, så er*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Teorem 7.2.4 (Parallellogram-loven)**.** *Hvis* $\mathbf{u}$ *og* $\mathbf{v}$ *er vektorer i et indreproduktrom* V*, så gjelder*

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

Selv om Teorem 7.2.2, Teorem 7.2.3 og Teorem 7.2.4 holder for et generelt indreproduktrom kan det være lurt å visualisere resultatene geometrisk i $\mathbb{R}^2$ og $\mathbb{R}^3$. Filosofien er at definisjonen av et indreproduktrom er såpass likt definisjonen av $\mathbb{R}^n$ at vår "klassiske" geometriske intuisjon fungerer mer eller mindre som utmerket.

7.3 Ortogonal dekomposisjon og projeksjon

Vi begynner med generaliseringen av ortonormale mengder til indreprodukt.

Definisjon. La V være et indreproduktrom. En mengde $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ av vektorer fra V kalles *ortonormal* dersom

- $\|\mathbf{u}_j\| = 1$ for $j = 1, 2, \ldots, n$,
- $\langle \mathbf{u}_j, \mathbf{u}_k\rangle = 0$ for alle $j \neq k$.

Hvis i tillegg $\operatorname{span}\mathcal{U} = V$ så kaller vi $\mathcal{U}$ en *ortonormal basis* for V.

Vi har allerede sett et eksempel på en ortonormal mengde.

Eksempel 7.3.1. Utregningene i Eksempel 7.1.6 viser at $\mathcal{U} = \{u_1, u_2\}$ for

$$u_1(x) = \frac{\sin x}{\sqrt{\pi}} \qquad \text{og} \qquad u_2(x) = \frac{\cos x}{\sqrt{\pi}}$$

er ortonormal i $C([a,b])$. Mengden er ikke en basis for $C([a,b])$ siden dette rommet er uendelig-dimensjonalt.

Hvis $\operatorname{span}\mathcal{U} \neq V$, så vil $U = \operatorname{span}\mathcal{U}$ være et underrom av V og $\mathcal{U}$ er en ortonormal basis for dette underrommet.

Eksempel 7.3.2. Mengden $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ for

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{og} \qquad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

er ortonormal i $\mathbb{R}^3$, men $U = \operatorname{span}\mathcal{U}$ er xy-planet (om vi identifiserer $\mathbb{R}^3$ som punktene (x, y, z)). Altså er $\mathcal{U}$ en ortonormal basis for dette planet.

La $\mathbf{v}$ være i $U = \operatorname{span}\mathcal{U}$ for en ortonormal mengde $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$. Det betyr at $\mathbf{v}$ kan skrives som en lineærkombinasjon av vektorene i $\mathcal{U}$, slik at

$$\mathbf{v} = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n$$

for tall $x_1, x_2, \ldots, x_n$. Det finnes en veldig enkel måte å finne disse tallene på. Vi har nemlig

$$\langle \mathbf{v}, \mathbf{u}_1 \rangle = x_1\langle \mathbf{u}_1, \mathbf{u}_1 \rangle + x_2\langle \mathbf{u}_2, \mathbf{u}_1 \rangle + \cdots + x_n\langle \mathbf{u}_n, \mathbf{u}_1 \rangle = x_1$$

og tilsvarende er $x_j = \langle \mathbf{v}, \mathbf{u}_j \rangle$ for $j = 2, 3, \ldots, n$. Totalt får vi den *ortogonale dekomposisjonen*

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n.$$

Det er klart at vi kan regne ut høyresiden selv om $\mathbf{v}$ ikke skulle være med i U, men da vil likheten ikke lenger holde.

Definisjon. La $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ være en ortonormal mengde i et indreproduktrom V og la $U = \operatorname{span}\mathcal{U}$. Den *ortogonale projeksjonen* på U er

$$\operatorname{proj}_U \mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n$$

for $\mathbf{v}$ i V.

Merk at $\operatorname{proj}_U \mathbf{v} = \mathbf{v}$ hvis og bare hvis $\mathbf{v}$ er i U. Følgende eksempel er kanskje enkelt, men det gir en god intuisjon om hvordan ortogonale projeksjoner fungerer.

Eksempel 7.3.3. La $\mathcal{U}$ og U være som i Eksempel 7.3.2. Hvis $\mathbf{v} = (x, y, z)$, så er

$$\operatorname{proj}_U \mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Her ser vi at den ortogonale projeksjonen på xy-planet tar et punkt (x, y, z) og sender det vinkelrett ned (hvis $z > 0$) eller vinkelrett opp (hvis $z < 0$) i xy-planet. Hvis $z = 0$, så er $\operatorname{proj}_U \mathbf{v} = \mathbf{v}$.

Et litt mer eksotisk eksempel er som følger.

Eksempel 7.3.4. La $\mathcal{U}$ være som i Eksempel 7.3.1 og la $U = \operatorname{span}\mathcal{U}$. Vi lar videre $h(x) = x$ og vil regne ut den ortogonale projeksjonen av h på U. Da trenger vi indreproduktene

$$\langle h, u_1\rangle = \int_{-\pi}^{\pi} x\frac{\sin x}{\sqrt{\pi}}\,dx = \frac{2}{\sqrt{\pi}} \qquad \text{og} \qquad \langle h, u_2\rangle = \int_{-\pi}^{\pi} x\frac{\cos x}{\sqrt{\pi}}\,dx = 0.$$

Altså blir

$$\operatorname{proj}_U h(x) = \langle h, u_1\rangle u_1(x) + \langle h, u_2\rangle u_2(x) = \frac{2}{\pi}\sin x.$$

Det viktigste egenskapene til den ortogonale projeksjonen er som følger.

Teorem 7.3.5. *La $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ være en ortonormal mengde i et indreproduktrom V og la $U = \operatorname{span}\mathcal{U}$. For hver $\mathbf{v}$ i V holder følgende:*

(a) *Vektoren $\mathbf{v} - \operatorname{proj}_U \mathbf{v}$ står ortogonalt på enhver vektor $\mathbf{u}$ i U.*

(b) *Den vektoren $\mathbf{u}$ i U som minimerer $\|\mathbf{v} - \mathbf{u}\|$ er $\mathbf{u} = \operatorname{proj}_U \mathbf{v}$.*

Bevis. For (a) er det nok å sjekke at $\langle \mathbf{v} - \operatorname{proj}_U \mathbf{v}, \mathbf{u}_j\rangle$ for $j = 1, 2, \ldots, n$. Vi har formelen

$$\operatorname{proj}_U \mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1\rangle\mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2\rangle\mathbf{u}_2 + \cdots + \langle \mathbf{v}, \mathbf{u}_n\rangle\mathbf{u}_n.$$

For $j = 1, 2, \ldots, n$ får vi derfor

$$\langle \mathbf{v} - \operatorname{proj}_U \mathbf{v}, \mathbf{u}_j\rangle = \langle \mathbf{v}, \mathbf{u}_j\rangle - \langle \operatorname{proj}_U \mathbf{v}, \mathbf{u}_j\rangle = \langle \mathbf{v}, \mathbf{u}_j\rangle - \langle \mathbf{v}, \mathbf{u}_j\rangle = 0$$

siden $\mathcal{U}$ er en ortonormal mengde. For (b) lar i $\mathbf{u}$ være en vilkårlig vektor i U og skriver

$$\mathbf{u} = \operatorname{proj}_U \mathbf{v} - \mathbf{u}_0.$$

Siden $\operatorname{proj}_U \mathbf{v}$ er i U må også $\mathbf{u}_0$ være i U. Det følger da av Pytagoras (Teorem 7.2.3) og (a) at

$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v} - \operatorname{proj}_U \mathbf{v} + \mathbf{u}_0\|^2 = \|\mathbf{v} - \operatorname{proj}_U \mathbf{v}\|^2 + \|\mathbf{u}_0\|^2.$$

Denne blir minst mulig om $\|\mathbf{u}_0\| = 0$ som gir $\mathbf{u}_0 = \mathbf{0}$ og $\mathbf{u} = \operatorname{proj}_U \mathbf{v}$. □

Tallet $\|\mathbf{v} - \mathbf{u}\|$ fra Teorem 7.3.5 er naturlig å tenke på som *avstanden* mellom vektoren $\mathbf{v}$ og underromet U.

Utregningen i Eksempel 7.3.4 og Teorem 7.3.5 gir oss følgende.

Eksempel 7.3.6. De reelle tallene a og b som minimerer integralet

$$\int_{-\pi}^{\pi} (x - a\sin x - b\cos x)^2\,dx$$

er $a = \frac{2}{\pi}$ og $b = 0$.

Vi kan også ta et regneksempel i trygge omgivelser.

Eksempel 7.3.7. Vi betrakter her $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ for

$$\mathbf{u}_1 = \frac{1}{3}\begin{pmatrix}1\\2\\2\end{pmatrix} \qquad \text{og} \qquad \mathbf{u}_2 = \frac{1}{3}\begin{pmatrix}2\\1\\-2\end{pmatrix}.$$

Siden $\|\mathbf{u}_1\| = \|\mathbf{u}\|_2 = 1$ og $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$ er $\mathcal{U}$ en ortonormal mengde i $\mathbb{R}^3$. Siden den består av to vektorer er $U = \operatorname{span}\mathcal{U}$ et plan. La oss betrakte vektoren

$$\mathbf{v} = \begin{pmatrix}1\\1\\3\end{pmatrix}.$$

Da blir $\mathbf{v} \cdot \mathbf{u}_1 = 3$ og $\mathbf{v} \cdot \mathbf{u}_2 = -1$, altså er

$$\operatorname{proj}_U \mathbf{v} = 3\mathbf{u}_1 - \mathbf{u}_2 = \begin{pmatrix}1\\2\\2\end{pmatrix} - \frac{1}{3}\begin{pmatrix}2\\1\\-2\end{pmatrix} = \frac{1}{3}\begin{pmatrix}1\\5\\8\end{pmatrix}.$$

Altså er avstanden mellom planet U og $\mathbf{v}$ lik

$$\|\mathbf{v} - \operatorname{proj}_U \mathbf{v}\| = \left\| \begin{pmatrix}1\\1\\3\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\5\\8\end{pmatrix} \right\| = \sqrt{\left(\frac{2}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2} = 1.$$

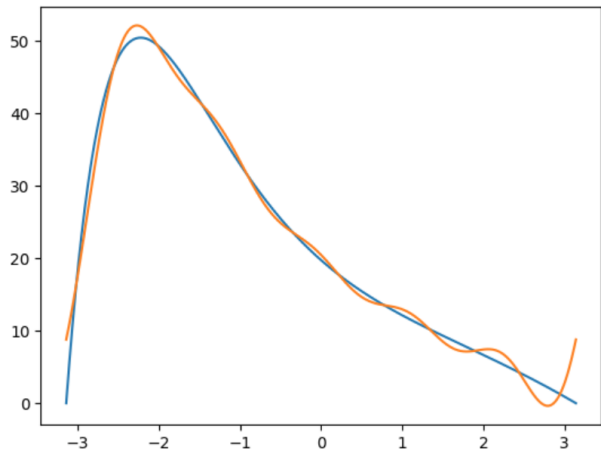
7.4 Fourier-rekker

Vi runder av med en liten smakebit på hva indreproduktrom og ortogonalitet kan anvendes til. Vi skal begrense oss til klassisk og—relativt sett—enkelt oppsett. Mer kompliserte oppsett basert på samme matematiske idégrunnlag brukes i anvendelser som:

- Lydfiltrering.
- Bildebehandling.
- Røntgenkrystallografi (bestem tredimensjonal struktur til en krystall).
- Computertomografi (CT-skan).

Hovedidéen bak alle anvendelsene er å dekomponere et komplisert *signal* i enklere bestanddeler, for eksempel for å bli kvitt støy eller for å overføre informasjon med minimalt tap.

Definisjon. En funksjon $f\colon \mathbb{R} \to \mathbb{R}$ sies å ha *periode* T dersom $f(x+T) = f(x)$ for alle x.



Figur 7.2: Plott av $y = f(x)$ og $y = \text{proj}_{U_5} f(x)$ på $[-\pi, \pi]$ for funksjonen $f(x) = (x^2 - \pi^2)(1 + \exp(-x))$.

Poenget med definisjonen er at hvis vi vet hvordan f ser ut på intervallet $[a, a+T]$ for et tall a, så vet vi hvordan f ser ut på hele $\mathbb{R}$. Noen ganger vil man være interessert i den minste perioden, men for oss er det mindre viktig.

Eksempel 7.4.1. Funksjonene $\cos x$ og $\sin x$ har periode 2π. Det har også $\cos nx$ og $\sin nx$ for $n = 1, 2, 3, \ldots$.

For enkelhets skyld begrenser vi oss til *kontinuerlige* funksjoner som har periode 2π. Det er naturlig å tenke på disse funksjonene som vektorer i indreproduktet $C([-\pi, \pi])$.

Teorem 7.4.2. *For hvert tall $n \geq 1$ så er mengden*

$$\mathcal{U}_n = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}, \frac{\cos nx}{\sqrt{\pi}} \right\}$$

ortonormal i indreproduktrommet $C([-\pi, \pi])$.

Hvis vi nå tenker på en funksjon f i $C([-\pi, \pi])$ som et signal, kan vi forsøke å dekomponere signalet ved funksjonene fra mengden $\mathcal{U}_n$. Vi vil bruke Teorem 7.3.5, som leder oss til følgende.

Definisjon. La f være en funksjon i $C([-\pi, \pi])$. Da er Fourier-koeffisientene til f gitt ved

$$\begin{aligned}
a_0 &= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x)\, dx, \\
a_n &= \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos nx\, dx, && \text{for } n = 1, 2, 3, \ldots, \\
b_n &= \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin nx\, dx, && \text{for } n = 1, 2, 3, \ldots.
\end{aligned}$$

Vi har da Fourier-rekken

$$f(x) = \frac{a_0}{\sqrt{2\pi}} + \sum_{n=1}^{\infty} \left(a_n \frac{\cos nx}{\sqrt{\pi}} + b_n \frac{\sin nx}{\sqrt{\pi}} \right).$$

Vi tillegger ikke dette likhetstegnet noen stor vekt. Hvis ikke f er for skrekkelig vil det konvergere fint, men det finnes også eksempler på funksjoner i $C([-\pi, \pi])$ hvor rekken ikke konvergerer i mange punkter. I anvendelser er vi ofte mye mer interessert i å ta en endelig sum det vil si at vi regner ut

$$\operatorname{proj}_{U_n} f(x) = \frac{a_0}{\sqrt{2\pi}} + \sum_{k=1}^{n} \left(a_k \frac{\cos kx}{\sqrt{\pi}} + b_k \frac{\sin kx}{\sqrt{\pi}} \right),$$

hvor $U_n = \operatorname{span} \mathcal{U}_n$. Generelt vil ikke $\operatorname{proj}_{U_n} f$ være lik f, men den vil ofte være veldig nærme. Se Figur 7.2.

Hvis vi er fornøyd med å erstatte f med $\operatorname{proj}_{U_5} f$, så kan vi forkaste f og ta ta vare på de elleve tallene $a_0, a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5$.