Kapittel 9

Projeksjon

En projeksjon er en lineærtransformasjon P som tilfredsstiller

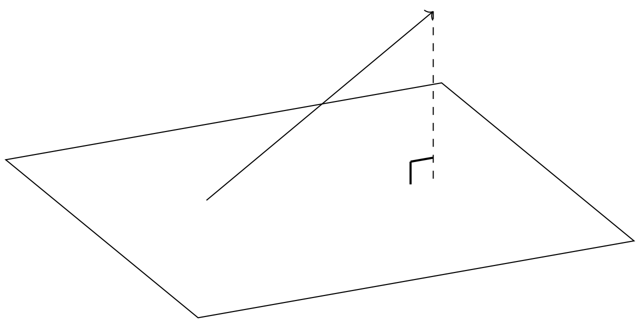
$$P\mathbf{x} = P^2\mathbf{x}$$

for alle $\mathbf{x}$. Denne ligningen sier at intet nytt skjer om du benytter lineærtransformasjonen for andre gang, og man kan tenke at $P\mathbf{x}$ er skyggen $\mathbf{x}$ kaster dersom man lyser på $\mathbf{x}$ med en lommelykt. Vi skal begrense oss til å studere ortogonale projeksjoner. Dette betyr at lommelykten står slik at $\mathbf{x}$ og $P\mathbf{x}$ danner en rettvinklet trekant, se figur lengre nede.

La oss først prøve å forstå hva en projeksjon er, ved å tenke litt over følgende oppgave:

> La $\mathbf{v} = [x, y, z]^T$ være en vektor i $\mathbb{R}^3$. Hva er
>
> 1. vektoren $P\mathbf{v}$ av $\mathbf{v}$, hvor $P : \mathbb{R}^3 \to \mathbb{R}^3$ sender $\mathbf{v}$ vinkelrett ned på xy-planet?
> 2. koordinatene til $P\mathbf{v}$ i basisen $\{\mathbf{e}_1, \mathbf{e}_2\}$?
> 3. koordinatene til $P\mathbf{v}$ i basisen for xy-planet gitt ved $[1, 1, 0]^T$ og $[1, -1, 0]^T$?

I løpet av dette kapittelet skal vi blant annet studere følgende figur. Her har vi et plan og en linje i $\mathbb{R}^3$. Vi har lyst til å projisere linjen ned i planet. Dette kan også forstås mer generelt: La V være et vektorrom, og la U være et underrom av V som er utspent av vektorene $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n$. La $\mathbf{v}$ være en vektor i V. Hva er projeksjonen $P_U(\mathbf{v})$ av $\mathbf{v}$ ned på U? Altså, hva er koordinatene til $P_U(\mathbf{v})$ med hensyn til $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_n$?



Projisere en vektor ned på et plan.

Det første vi skal lære oss er å forstå projeksjonen $P_\mathbf{u}(\mathbf{v})$ av en vektor $\mathbf{v}$ ned på en vilkårlig vektor $\mathbf{u}$. Så skal vi se at dersom vi har en basis $U = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, hvor hver av basisvektorene står vinkelrett på hverandre, så kan vi projisere $\mathbf{v}$ ned på U på følgende måte, fordi P er en lineærtransformasjon:

$$P_U(\mathbf{v}) = P_{\mathbf{u}_1}(\mathbf{v}) + P_{\mathbf{u}_2}(\mathbf{v}) + ... + P_{\mathbf{u}_n}(\mathbf{v}).$$

Dette er et av de viktigste resultatene i dette kapittelet. Aller først skal vi diskutere skalarprodukt og projeksjon i $\mathbb{R}^2$.

Ortogonal projeksjon i $\mathbb{R}^2$

Vi husker skalarproduktet, eller prikkproduktet, fra videregående skole. Du har lært to måter å beregne dette på, nemlig

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta,$$

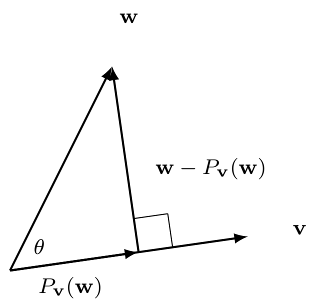
der $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$ er lengden til $\mathbf{v}$ og $\|\mathbf{w}\|$ er lengden til $\mathbf{w}$ og θ er vinkelen mellom $\mathbf{v}$ og $\mathbf{w}$, og

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2.$$

Ut ifra den andre formelen så observerer vi at dersom vi tar skalarproduktet av $\mathbf{v}$ med seg selv, så får vi $v_1^2 + v_2^2$. Det følger dermed at lengden $\|\mathbf{v}\|$ av $\mathbf{v}$ er:

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

Videre kan vi bruke skalarproduktet til å projisere vektorer ortogonalt på hverandre. Det sentrale spørsmålet er: hvordan kan vi skrive vektoren $P_\mathbf{v}(\mathbf{w})$, projeksjonen av vektoren $\mathbf{w}$ ned på $\mathbf{v}$, i figuren under?



Hva er projeksjon?

Vi starter med å utlede en formel for lengden $\|P_\mathbf{v}(\mathbf{w})\|$, som vi etterpå skal bruke i utledningen av $P_\mathbf{v}(\mathbf{w})$:

$$\|P_\mathbf{v}(\mathbf{w})\| = \|\mathbf{w}\| \cos\theta = \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|}\|\mathbf{w}\| \cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|}.$$

Da følger det videre, siden $P_{\mathbf{v}}(\mathbf{w})$ ligger langs $\mathbf{v}$, at

$$P_{\mathbf{v}}(\mathbf{w}) = \frac{\|P_{\mathbf{v}}(\mathbf{w})\|}{\|\mathbf{v}\|}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{v}\|^2}\mathbf{v} = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}.$$

Denne vektoren kalles gjerne $\mathbf{w}$ sin komponent i retningen gitt av $\mathbf{v}$, eller den *ortogonale projeksjonen* av $\mathbf{w}$ på $\mathbf{v}$. Fra figuren ovenfor ser vi at komponenten til $\mathbf{w}$ ortogonalt, eller vinkelrett, på $\mathbf{v}$ er vektoren

$$\mathbf{w} - P_{\mathbf{v}}(\mathbf{w}).$$

Eksempel 9.1. Vektoren

$$\mathbf{w} = \begin{bmatrix}2\\1\end{bmatrix}$$

sin komponent i retningen gitt av

$$\mathbf{v} = \begin{bmatrix}1\\2\end{bmatrix}$$

er:

$$P_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} = \frac{4}{5}\begin{bmatrix}1\\2\end{bmatrix}.$$

Vi kan også beregne vektoren $\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})$:

$$\mathbf{w} - P_{\mathbf{v}}(\mathbf{w}) = \begin{bmatrix}2\\1\end{bmatrix} - \frac{4}{5}\begin{bmatrix}1\\2\end{bmatrix} = \frac{3}{5}\begin{bmatrix}2\\-1\end{bmatrix}. \qquad \triangle$$

Skalarproduktet i $\mathbb{R}^n$

Skalarproduktet i $\mathbb{R}^2$ kan med ord beskrives ved at vi legger sammen produktet av komponentene til to vektorer. Dette er en sterk indikasjon på at den naturlige generaliseringen i $\mathbb{R}^n$ er

$$\mathbf{v}\cdot\mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n,$$

som også kan uttrykkes ved matriseproduktet

$$\mathbf{v}\cdot\mathbf{w} = \mathbf{v}^\top\mathbf{w}.$$

Hvordan skal man tenke på dette produktet? Svaret er at intuisjonen din fra $\mathbb{R}^2$ fungerer helt fint. Husk at lengden til en vektor er gitt som

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Formelen

$$\mathbf{v}\cdot\mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\|\cos\theta$$

gjelder fortsatt. Formelen for vinkelen mellom to vektorer gir spesielt at to vektorer er *ortogonale* hvis

$$\mathbf{v}\cdot\mathbf{w} = 0,$$

og *parallelle* hvis

$$\mathbf{v}\cdot\mathbf{w} = \pm\|\mathbf{v}\|\|\mathbf{w}\|.$$

Andre ligning stemmer fordi dette betyr at vinkelen mellom dem er 0 eller 180 grader; de ligger på samme linje. Akkurat som i $\mathbb{R}^2$ blir den *ortogonale projeksjonen* av $\mathbf{w}$ på $\mathbf{v}$

$$P_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}.$$

Intuisjonen vår – som illustrert i figuren fra forrige seksjon – forteller oss at en vektor $\mathbf{w}$ som er parallell med $\mathbf{v}$ burde være lik sin egen projeksjon på $\mathbf{v}$. Vi kan sjekke at algebraen stemmer overens med intuisjonen: parallellitet betyr at $\mathbf{w}$ er et skalarmultiplum av $\mathbf{v}$; $\mathbf{w} = t\mathbf{v}$. Sett inn i formelen for å se at

$$\begin{aligned}P_{\mathbf{v}}(\mathbf{w}) &= \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\\ &= \frac{\mathbf{v}\cdot t\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\\ &= t\frac{\mathbf{v}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\\ &= t\mathbf{v}\\ &= \mathbf{w}.\end{aligned}$$

På samme måte virker det rimelig at projeksjonen av en vektor $\mathbf{w}$ som er ortogonal med $\mathbf{v}$ er null-vektoren. Igjen, algebraen er enig: vi antar at $\mathbf{v}\cdot\mathbf{w} = 0$ slik at

$$\begin{aligned}P_{\mathbf{v}}(\mathbf{w}) &= \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\\ &= \frac{0}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\\ &= 0\cdot\mathbf{v}\\ &= \mathbf{0}.\end{aligned}$$

Eksempel 9.2. Vektorene $\mathbf{v} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ og $\mathbf{w} = \begin{bmatrix}1\\-1\\6\end{bmatrix}$ er ikke ortogonale fordi

$$\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\-1\\6\end{bmatrix} = 1\cdot 1 + 1\cdot(-1) + 1\cdot 6 = 6.$$

Den ortogonale projeksjonen av $\mathbf{w}$ på $\mathbf{v}$ er

$$\begin{aligned}P_{\mathbf{v}}(\mathbf{w}) &= \frac{\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\-1\\6\end{bmatrix}}{\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix}\\ &= \frac{6}{1^2+1^2+1^2}\begin{bmatrix}1\\1\\1\end{bmatrix}\\ &= \frac{6}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}\\ &= \begin{bmatrix}2\\2\\2\end{bmatrix}.\end{aligned}$$

Vi ser at $P_{\mathbf{v}}(\mathbf{w})$ og $\mathbf{w}$ er parallelle siden $P_{\mathbf{v}}(\mathbf{w}) = 2\mathbf{w}$, og at $P_{\mathbf{v}}(\mathbf{w})$ og $\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})$ er ortogonale, fordi:

$$\begin{aligned}\begin{bmatrix}2\\2\\2\end{bmatrix}\cdot\left(\begin{bmatrix}1\\-1\\6\end{bmatrix} - \begin{bmatrix}2\\2\\2\end{bmatrix}\right) &= \begin{bmatrix}2\\2\\2\end{bmatrix}\cdot\begin{bmatrix}-1\\-3\\4\end{bmatrix}\\ &= 2\cdot(-1) + 2\cdot(-3) + 2\cdot 4\\ &= 0.\end{aligned}$$

Det er også verdt å merke seg at vi kan skrive $\mathbf{w}$ som en sum av komponenten $P_{\mathbf{v}}(\mathbf{w})$ parallell med $\mathbf{v}$ og komponenten $\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})$ ortogonal på $\mathbf{w}$:

$$\mathbf{w} = P_{\mathbf{v}}(\mathbf{w}) + (\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})) \qquad \triangle$$

Fikser en vektor $\mathbf{v}$ i $\mathbb{R}^n$. Du kan tenkte på $P_{\mathbf{v}}$ som en funksjon fra $\mathbb{R}^n$ til $\mathbb{R}^n$ som tar inn en vektor, $\mathbf{w}$, for å produsere den ortogonale projeksjonen

$$P_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}.$$

Men den er mer enn en vanlig funksjon. Den er av typen vi liker aller best; den er en lineærtransformasjon.

Teorem 9.3. *Den ortogonale projeksjonen på en vektor* $\mathbf{v}$,

$$\begin{aligned} P_{\mathbf{v}} &: \mathbb{R}^n \to \mathbb{R}^n \\ P_{\mathbf{v}}(\mathbf{w}) &= \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}, \end{aligned}$$

er en lineærtransformasjon.

Bevis. Resultatet følger essensielt av at skalarproduktet er lineært i andre faktor. For å se at skalarproduktet er lineært i andre faktor, må vi vise at

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

Husk at $\mathbf{v}\cdot\mathbf{w}$ er matriseproduktet $\mathbf{v}^\top\mathbf{w}$. Det er med andre ord lineærtransformasjonen som er gitt av $1\times n$-matrisen $\mathbf{v}^\top$. Dette gir automatisk at skalarproduktet er lineært som følge av at matriseproduktet er distributivt. Du kan alternativt bevise påstanden direkte ved å bruke formelen $\mathbf{v}\cdot\mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n$. Nå er det rett frem å vise at projeksjonen på $\mathbf{v}$ er en lineærtransformasjon:

$$\begin{aligned} P_{\mathbf{v}}(a\mathbf{w}+b\mathbf{u}) &= \frac{\mathbf{v}\cdot(a\mathbf{w}+b\mathbf{u})}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \\ &= \frac{a\mathbf{v}\cdot\mathbf{w}+b\mathbf{v}\cdot\mathbf{u}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \qquad \text{(linearitet)} \\ &= a\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} + b\frac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \quad (\mathbf{v}\cdot\mathbf{v} \text{ er et tall}) \\ &= aP_{\mathbf{v}}(\mathbf{w}) + bP_{\mathbf{v}}(\mathbf{u}). \ \square \end{aligned}$$

Eksempel 9.4. La $\mathbf{v} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ være vektoren fra forrige eksempel. Vi så nettopp at projeksjonen $P_{\mathbf{v}}$ er en lineærtransormasjon fra $\mathbb{R}^3$ til $\mathbb{R}^3$. Fra teorien om lineærtransformasjoner vet vi at det er mulig å finne en 3×3-matrise $[P_{\mathbf{v}}]$ – som kalles standardmatrisen til $P_{\mathbf{v}}$ – slik at $P_{\mathbf{v}}(\mathbf{x}) = [P_{\mathbf{v}}]\mathbf{x}$. Denne matrisen er alltid gitt som

$$[P_{\mathbf{v}}] = \begin{bmatrix} P_{\mathbf{v}}(\mathbf{e}_1) & P_{\mathbf{v}}(\mathbf{e}_2) & P_{\mathbf{v}}(\mathbf{e}_3)\end{bmatrix}$$

hvor $\mathbf{e}_i$-ene er standardbasisen til $\mathbb{R}^3$. Regn ut:

$$\begin{aligned} P_{\mathbf{v}}(\mathbf{e}_1) &= \frac{\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\0\\0\end{bmatrix}}{\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\\1\end{bmatrix}}\begin{bmatrix}1\\1\\1\end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} \end{aligned}$$

Tilsvarende blir også

$$P_{\mathbf{v}}(\mathbf{e}_2) = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} \text{ og } P_{\mathbf{v}}(\mathbf{e}_3) = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}.$$

Standardmatrisen er derfor

$$[P_{\mathbf{v}}] = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}.$$

Hvis $\mathbf{w} = \begin{bmatrix}1\\-1\\6\end{bmatrix}$ også er som i forrige eksempel, så kan vi verifisere at

$$[P_{\mathbf{v}}]\mathbf{w} = \frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}1\\-1\\6\end{bmatrix} = \frac{1}{3}\begin{bmatrix}6\\6\\6\end{bmatrix} = \begin{bmatrix}2\\2\\2\end{bmatrix}$$

stemmer overens med $P_{\mathbf{v}}(\mathbf{w}) = \begin{bmatrix}2\\2\\2\end{bmatrix}$. Det gjør den.

Videre kan vi spørre: Er $P_{\mathbf{v}}$ injektiv? Er $P_{\mathbf{v}}$ surjektiv? Vi kan selvfølgelig svare på disse spørsmålene ved å radredusere $[P_{\mathbf{v}}]$, men dette kan også forstås geometrisk. Tenk litt på hvorfor et plan som står normalt på linjen utspent av $\mathbf{v}$, $\mathrm{Sp}\{\mathbf{v}\}$, er alle vektorene som projiseres ned på skjæringspunktet. Nå skjønner du kanskje at det er mange vektorer – et helt plan – som treffer en gitt vektor på $\mathrm{Sp}\{\mathbf{v}\}$. Men dette betyr jo at ulike vektorer sendes til like vektorer; transformasjonen er ikke injektiv. Spesielt er kjernen til $P_{\mathbf{v}}$, eller nullrommet til $[P_{\mathbf{v}}]$, planet som står normalt på $\mathrm{Sp}\{\mathbf{v}\}$ med origo som skjæringspunkt. Bildet til $P_{\mathbf{v}}$, eller kolonnerommet til $[P_{\mathbf{v}}]$, er $\mathrm{Sp}\,\mathbf{v}$ – fordi vektorer projiseres ned på denne linjen. Da $\mathbf{v}$ ikke spenner ut hele $\mathbb{R}^3$ så er $P_{\mathbf{v}}$ ikke surjektiv. △

Nå kan vi gå tilbake til hva vi skrev i introduksjonen, og vise at ligningen $P\mathbf{x} = P^2\mathbf{x}$ faktisk holder for alle projeksjoner P.

Teorem 9.5. *La* $\mathbf{v}$ *og* $\mathbf{w}$ *være to vektorer og la* $P_{\mathbf{v}}(\mathbf{w})$ *være projeksjonen av* $\mathbf{w}$ *ned på* $\mathbf{v}$. *Da gjelder det at* $P_{\mathbf{v}}(\mathbf{w}) = P_{\mathbf{v}}^2(\mathbf{w})$, *og generelt, at* $P\mathbf{x} = P^2\mathbf{x}$ *for alle projeksjoner* P *og alle vektorer* $\mathbf{x}$.

Bevis. La $P_{\mathbf{v}}(\mathbf{w}) = \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$. Da trenger vi bare sjekke hva som skjer dersom vi beregner

$$P_{\mathbf{v}}^2(\mathbf{w}) = P_{\mathbf{v}}\left(\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\right).$$

Det følger at

$$\begin{aligned} P_{\mathbf{v}}\left(\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\right) &= \frac{\mathbf{v}\cdot\left(\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}\right)}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \\ &= \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\frac{\mathbf{v}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \\ &= \frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v} \\ &= P_{\mathbf{v}}(\mathbf{w}). \end{aligned}$$

Teoremet er altså sant for vilkårlige vektorpar $\mathbf{v}$ og $\mathbf{w}$. Ettersom en projeksjon P er en lineærtransform, så kan vi projisere $\mathbf{v}$ ned på en hvilkårlig mengde vektorer ved å ta lineærkombinasjonen av $\mathbf{v}$ projisert ned på hver enkelt vektor i mengden. Derfor gjelder det generelt at $P\mathbf{x} = P^2\mathbf{x}$ for alle projeksjoner P og alle vektorer $\mathbf{x}$. $\square$

Eksempel 9.6. Vi sjekker at teoremet gjelder for projeksjonen $P_{\mathbf{v}}$ fra forrige eksempel:

$$\begin{aligned} P_{\mathbf{v}}^2(\mathbf{w}) = [P_{\mathbf{v}}]^2\mathbf{w} &= \frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}^2 \\ &= \frac{1}{9}\begin{bmatrix}3 & 3 & 3\\ 3 & 3 & 3\\ 3 & 3 & 3\end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} = [P_{\mathbf{v}}]\mathbf{w} = P_{\mathbf{v}}(\mathbf{w}). \triangle \end{aligned}$$

Hvis du tenker deg om vil du etterhvert innse at geometrien til vektorer i $\mathbb{R}^n$ kommer fra skalarproduktet. Vi har sett at vinkel, projeksjon og lengde kan uttrykkes ved skalarproduktet. Til og med selve koordinatsystemet vi ser for oss er beskrevet av skalarproduktet: Standardbasisen i $\mathbb{R}^n$ er vektorene $\mathbf{e}_i$ som har komponent i lik 1, og null ellers. Merk at hvis du tar skalarproduktet av en vilkårlig vektor $\mathbf{v} = \begin{bmatrix}v_1\\ v_2\\ \vdots\\ v_n\end{bmatrix}$ med $\mathbf{e}_i$, så får du enkelt og greit v_i, den i-te komponenten til $\mathbf{v}$. Et eksempel fra $\mathbb{R}^3$ er

$$\begin{bmatrix}1\\ 2\\ 3\end{bmatrix} \cdot \begin{bmatrix}0\\ 1\\ 0\end{bmatrix} = 1 \cdot 0 + 2 \cdot 1 + 3 \cdot 0 = 2.$$

Derfor kan vi skrive $\mathbf{v}$ som en lineærkombinasjon

$$\mathbf{v} = v_1\mathbf{v}_{\mathbf{e}_1} + v_2\mathbf{v}_{\mathbf{e}_2} + \cdots + v_n\mathbf{v}_{\mathbf{e}_n}.$$

Fra dette kan vi konkludere at de to ulike definisjonene vi gav for skalarproduktet i introduksjonen faktisk er identiske. Først ser vi at for alle i har vi

$$\mathbf{v} \cdot \mathbf{e}_i = \|\mathbf{v}\|\|\mathbf{e}_i\| \cos\theta = \|\mathbf{v}\| \cos\theta = v_i,$$

og da følger det at

$$\begin{aligned} \mathbf{v} \cdot \mathbf{w} &= \mathbf{v} \cdot \sum_i w_i\mathbf{e}_i \\ &= \sum_i (\mathbf{v} \cdot \mathbf{e}_i)w_i \\ &= \sum_i (\|\mathbf{v}\|\|\mathbf{e}_i\| \cos\theta)w_i \\ &= \sum_i (\|\mathbf{v}\| \cos\theta)w_i \\ &= \sum_i v_iw_i \\ &= v_1w_1 + ...v_nw_n. \end{aligned}$$

I tillegg ser vi at koordinater altså bare er summen av projeksjoner ned på vektorer i standardbasisen. Grunnen til dette er at standardbasisen er et eksempel på en *ortogonal basis*, noe vi skal komme tilbake til senere i kapittelet. Alt av geometri ser ut til å komme fra at $\mathbb{R}^n$ har skalarprodukt. Målet med neste seksjon er å generalisere skalarproduktet – som vi kommer til å kalle et indreprodukt – for å finne lignende geometri i mer abstrakte vektorrom.

Indreproduktrom

Vi ønsker å definere et produkt som tar inn to vektorer for å produsere en skalar. Dette produktet skal gi en geometri som minner om prikkproduktet i $\mathbb{R}^n$. Konvensjonen er å kalle generaliseringen for et *indreprodukt*, med notasjon $\langle \mathbf{v}, \mathbf{w} \rangle$ – indreproduktet mellom $\mathbf{v}$ og $\mathbf{w}$.

Skalarproduktet i $\mathbb{R}^n$ er en operasjon som tar inn to vektorer, $\mathbf{v}$ og $\mathbf{w}$, for å produsere en skalar $\mathbf{v} \cdot \mathbf{w}$. Denne operasjonen er *symmetrisk*:

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$$

Husk at to vektorer i $\mathbb{R}^n$ er ortogonale hvis $\mathbf{v} \cdot \mathbf{w} = 0$. Det er nettopp symmetrien til skalarproduktet som gjør definisjonen veldefinert. Uten symmetrien kunne det tenkes at $\mathbf{v} \cdot \mathbf{w} = 0$ og $\mathbf{w} \cdot \mathbf{v} \neq 0$, eller $\mathbf{v} \cdot \mathbf{w} \neq 0$ og $\mathbf{w} \cdot \mathbf{v} = 0$. Dette virker helt absurd.

Videre tilfredstiller skalarproduktet en egenskap vi kaller *positivitet*:

$$\mathbf{v} \cdot \mathbf{v} \geq 0, \text{ og } \mathbf{v} \cdot \mathbf{v} = 0 \text{ kun hvis } \mathbf{v} = \mathbf{0}$$

Kravet $\mathbf{v} \cdot \mathbf{v} \geq 0$ er rett og slett det som lar oss definere lengden

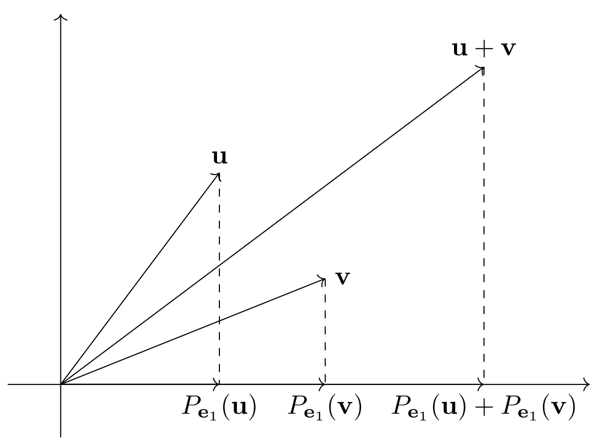
$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}},$$

siden kvadratroten bare er definert for positive tall. Det andre kravet betyr at det kun er nullvektoren som har lengde lik null. Positivitet er med andre ord essensielt for at lengde-begrepet skal gi mening.

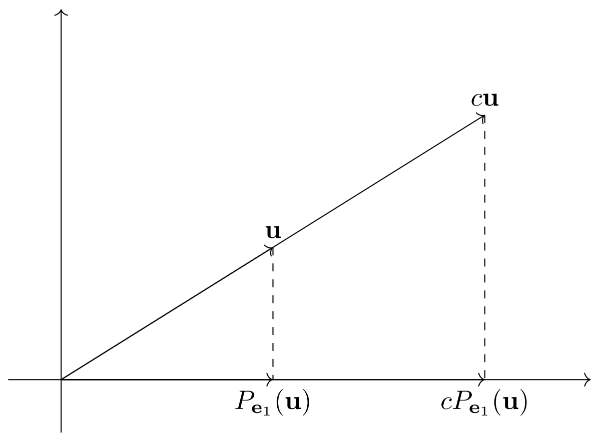
Den siste egenskapen vi skal tenke på er *linearitet*:

$$\mathbf{v} \cdot (a\mathbf{w} + b\mathbf{u}) = a\mathbf{v} \cdot \mathbf{w} + b\mathbf{v} \cdot \mathbf{u}$$

Skalarproduktet også er lineært i første faktor fordi det er symmetrisk. I beviset til Teorem 9.3 kommer det frem at nettopp denne egenskapen gjør projeksjon lineær. Rent geometrisk får vi naturlige bilder å la:



Å projisere for så å addere, eller å addere for så å projisere, er det samme



Å projisere for så å skalere, eller å skalere for så å projisere, er det samme

Definisjon. La V være et reelt vektorrom. Et *indreprodukt* i V er en operasjon som tar inn to vektorer, $\mathbf{v}$ og $\mathbf{w}$, for å gi ut et reelt tall $\langle \mathbf{v}, \mathbf{w} \rangle$. Operasjonen tilfredstiller

$$
\begin{aligned}
&\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle && \text{(symmetri)} \\
&\langle \mathbf{v}, (a\mathbf{w} + b\mathbf{u}) \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{u} \rangle && \text{(linearitet)} \\
&\langle \mathbf{v}, \mathbf{v} \rangle \geq 0, \text{ og } \langle \mathbf{v}, \mathbf{v} \rangle = 0 \text{ kun hvis } \mathbf{v} = \mathbf{0} && \text{(positivitet)}
\end{aligned}
$$

Vi sier at V, sammen med et valgt indreprodukt, er et *indreproduktrom*. △

Merk. Hvis du kombinerer linearitet og symmetri, så får du også at indreproduktet er lineært i første faktor:

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

Eksempel 9.7. Skalarproduktet i $\mathbb{R}^n$,

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot \mathbf{w},$$

er et indreprodukt i $\mathbb{R}^n$. Til sammen utgjør de et indreproduktrom. △

Merk. I $\mathbb{R}^2$ er skalarproduktet med nullvektoren alltid null. Dette er også tilfellet i et generelt indreproduktrom. Positivitet viser at den eneste vektoren som er ortogonal med *alle* vektorer i V er nullvektoren. △

Eksempel 9.8. Hvis du deler en ikke-null vektor på lengden sin, får du en vektor av lengde 1; $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ har lengde lik 1. Algebraisk verifikasjon:

$$
\begin{aligned}
\|\frac{\mathbf{v}}{\|\mathbf{v}\|}\| &= \sqrt{\langle \frac{\mathbf{v}}{\|\mathbf{v}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|} \rangle} \\
&= \sqrt{\frac{1}{\|\mathbf{v}\|^2} \langle \mathbf{v}, \mathbf{v} \rangle} \\
&= \sqrt{\frac{1}{\|\mathbf{v}\|^2} \|\mathbf{v}\|^2} \\
&= \sqrt{1} = 1.
\end{aligned}
$$

△

Basert på disse definisjonene kan vi allerede bevise Pytagoras' teorem.

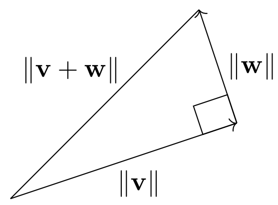
Teorem 9.9 (Pytagoras)**.** *La V være et reelt indreproduktrom. Dersom vektorene $\mathbf{v}$ og $\mathbf{w}$ er ortogonale, er*

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

Bevis. Vi regner litt på venstre side av ligningen.

$$
\begin{aligned}
\|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \\
&+ \langle \mathbf{w}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle, && \text{linearitet} \\
&= \langle \mathbf{v}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle, && \text{ortogonalitet} \\
&= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.
\end{aligned}
$$

□



Pytagoras' teorem

Det neste teoremet er en ulikhet som lar oss definere vinkelen mellom vektorer.

Teorem 9.10 (Cauchy–Schwarz)**.** *La V være et reelt indreproduktrom. Alle vektorer $\mathbf{v}$ og $\mathbf{w}$ tilfredstiller*

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

Fra ulikheten følger det at

$$-1 \leq \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1.$$

Cosinus ligger også mellom -1 og 1. Derfor kan vi definere *vinkelen* mellom $\mathbf{v}$ og $\mathbf{w}$ til å være løsningen av

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

Vi har med andre ord fortsatt formelen

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

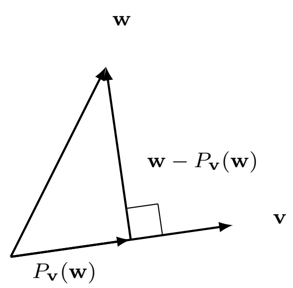
Spesielt er to vektorer *parallelle* hvis

$$\langle \mathbf{v}, \mathbf{w} \rangle = \pm\|\mathbf{v}\|\|\mathbf{w}\|,$$

eller ekvivalent, $\mathbf{w} = t\mathbf{v}$ for et reelt tall t; de ligger på samme linje. På dette tidspunktet blir du kanskje ikke overrasket over at den *ortogonale projeksjonen* av $\mathbf{w}$ på $\mathbf{v}$ er

$$P_{\mathbf{v}}(\mathbf{w}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}\mathbf{v}.$$

Bildet er fortsatt det samme som i $\mathbb{R}^2$:



Den ortogonale projeksjonen i et indreproduktrom

Teorem 9.3 sier at den ortogonale projeksjonen i $\mathbb{R}^n$ er en lineærtransformasjon. Hvis du leste beviset, så er du enig i at alt koker ned til at skalarproduktet er lineært. Samme bevis fungerer dermed for generelle indreproduktrom ved å bytte $\mathbf{v} \cdot \mathbf{w}$ med $\langle \mathbf{v}, \mathbf{w} \rangle$.

Teorem 9.11. *La V være et indreproduktrom. Den ortogonale projeksjonen på en vektor $\mathbf{v}$,*

$$\begin{aligned} P_{\mathbf{v}} &: V \to V \\ P_{\mathbf{v}}(\mathbf{w}) &= \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \end{aligned}$$

er en lineærtransformasjon.

Teorem 9.12. *La $\mathbf{v}$ og $\mathbf{w}$ være to vektorer i et indreproduktrom V. Da er $P_{\mathbf{v}}(\mathbf{w})$ og $\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})$ ortogonale.*

Bevis. Bruk linearitet (og symmetri, for å få linearitet i første faktor):

$$\begin{aligned} &\langle P_{\mathbf{v}}(\mathbf{w}), \mathbf{w} - P_{\mathbf{v}}(\mathbf{w}) \rangle \\ =&\langle \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{w} - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \rangle \\ =&\langle \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{w} \rangle - \langle \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \rangle \\ =&\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{w} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{v} \rangle \\ =&\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{w} \rangle - \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \langle \mathbf{v}, \mathbf{w} \rangle \\ &= 0 \end{aligned}$$

□

Ortogonal projeksjon

Definisjon. En *ortogonal mengde* er en mengde av ikke-null vekorer $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$, slik at

$$\langle \mathbf{u}_i, \mathbf{u}_k \rangle = 0$$

for alle vektorer $\mathbf{u}_i$ og $\mathbf{u}_k$ i mengden med $i \neq k$. Dersom i tillegg $\|\mathbf{u}_j\| = 1$ for alle vektorene, sier vi at mengden er *ortonormal*.

Intuisjonen bak det neste teoremet er klar: Vektorer som parvis står 90 grader på hverandre er lineært uavhengig. Du kan f. eks. tenke på standardbasisen i $\mathbb{R}^3$.

Teorem 9.13. *En ortogonal mengde er lineært uavhengig.*

Bevis. La $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ være en vilkårlig ortogonal mengde i et indreproduktrom. Vi ønsker å vise at ligningen

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$$

kun har triviell løsning – definisjonen på lineær uavhengighet. Trikset er å anvende indreproduktet med alle $\mathbf{u}_i$-ene på ligningen. Høyre side:

$$\langle \mathbf{0}, \mathbf{u}_i \rangle = 0$$

Venstre side:

$$\begin{aligned} &\langle x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n, \mathbf{u}_i \rangle \\ =&x_1\langle \mathbf{u}_1, \mathbf{u}_i \rangle + x_2\langle \mathbf{u}_2, \mathbf{u}_i \rangle + \cdots + x_n\langle \mathbf{u}_n, \mathbf{u}_i \rangle \\ =&x_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle. \end{aligned}$$

Til sammen får vi at

$$x_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 0.$$

Her kan ikke $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 0$ fordi alle $\mathbf{u}_i$-ene er ikke nullvektoren (positivitet). Dermed må $x_i = 0$. Vi har kun triviell løsning $x_i = 0$ for alle i.

□

Eksempel 9.14. Standardbasisen $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for $\mathbb{R}^n$ er en ortonormal mengde. △

Definisjon. Dersom en ortogonal mengde $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ i V også er en basis, sier vi at mengden er en *ortogonal basis* for V.

Merk. Hvis indreproduktet er skalarproduktet i $\mathbb{R}^n$, så er det vanlig å sette opp $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ som kolonner i en matrise U. Vi sier da at U er en *ortogonal matrise.*

Hvis vi har en ortogonal basis for et rom, er det veldig lett å finne en vektors komponenter i rommet. La oss si at vi ønsker å finne vektoren $\mathbf{v}$ sine koordinater i basisen $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Koordinatene, $x_1, x_2, \ldots, x_n$, til $\mathbf{v}$ i denne basisen er gitt ved ligningen

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

I motivasjonen for indreproduktet så vi at når $\mathbf{u}_i$-ene er standardbasisen til $\mathbb{R}^n$, så er x_i-ene bare komponentene til vektoren $\mathbf{v}$; projeksjonen ned på hvert element i standardbasisen. Det samme gjelder for en ortonogonal basis. Anvend indreproduktet med $\mathbf{u}_i$ for å få ut den i-te komponenten:

$$\begin{aligned} \langle \mathbf{u}_i, \mathbf{v} \rangle &= \langle \mathbf{u}_i, x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + ...x_n\mathbf{u}_n \rangle \\ &= x_1\langle \mathbf{u}_1, \mathbf{u}_i \rangle + x_2\langle \mathbf{u}_2, \mathbf{u}_i \rangle \\ &+ \cdots + x_n\langle \mathbf{u}_n, \mathbf{u}_i \rangle, && \text{linearitet} \\ &= x_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle, && \text{ortogonalitet} \end{aligned}$$

Altså er

$$\langle \mathbf{u}_i, \mathbf{v} \rangle = x_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle$$

for hver i. Fordi $\mathbf{u}_i$ ikke er nullvektoren, er $\langle \mathbf{u}_i, \mathbf{u}_i \rangle > 0$. Altså er løsningen

$$x_i = \frac{\langle \mathbf{u}_i, \mathbf{v} \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}.$$

Når vi husker formelen for projeksjonen på en vektor, ser vi at

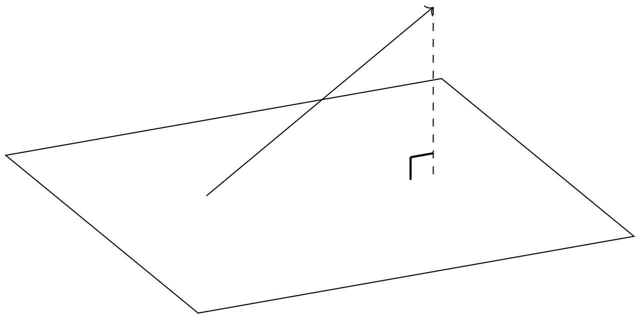
$$x_i\mathbf{u}_i = \frac{\langle \mathbf{u}_i, \mathbf{v} \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}\mathbf{u}_i = P_{\mathbf{u}_i}(\mathbf{v}).$$

Vi beviste altså følgende teorem:

Teorem 9.15. *Koordinatene til $\mathbf{v}$ i en ortogonal basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ er*

$$\begin{aligned} \mathbf{v} &= P_{\mathbf{u}_1}(\mathbf{v}) + P_{\mathbf{u}_2}(\mathbf{v}) + ... + P_{\mathbf{u}_n}(\mathbf{v}) \\ &= \frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{v} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{u}_n, \mathbf{v} \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\mathbf{u}_n. \end{aligned}$$

Vi kan også projisere en vektor ned i et underrom der den ikke hører hjemme. Projeksjonen minimerer avstanden fra underrommet til vektoren. Det neste teoremet sier at vi må finne en ortogonal basis for å få til situasjonen som er illustrert i bildet nedenfor.



Ortogonal projeksjon minimerer avstanden til et underrom

Teorem 9.16. *La* $\mathbf{u}_1$*,* $\mathbf{u}_2$*, ...,*$\mathbf{u}_n$ *være en ortogonal basis for* U*, et underrom av* V*. Punktet*

$$\begin{aligned}P_U(\mathbf{v}) &= P_{\mathbf{u}_1}(\mathbf{v}) + P_{\mathbf{u}_2}(\mathbf{v}) + ... + P_{\mathbf{u}_n}(\mathbf{v}) \\ &= \frac{\langle \mathbf{u}_1, \mathbf{v}\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{v}\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 + ... + \frac{\langle \mathbf{u}_n, \mathbf{v}\rangle}{\langle \mathbf{u}_n, \mathbf{u}_n\rangle}\mathbf{u}_n\end{aligned}$$

er det punktet i U *som har kortest avstand til* $\mathbf{v}$*:*

$$\|\mathbf{v} - P_U(\mathbf{v})\| = \min_{\mathbf{u}\in U} \|\mathbf{v} - \mathbf{u}\|.$$

Punktet $P_U(\mathbf{v})$ er det unike punktet som minimerer avstanden fra $\mathbf{v}$ til U. Derfor er $P_U(\mathbf{v})$ uavhengig av hvilken ortogonal basis du velger for U. Dette definerer en lineærtransformasjon $P_U : V \to U$, *den ortogonale projeksjonen* ned på underrommet U. Du kan regne den ut ved å velge en ortogonal basis for så å bruke formelen i Teorem 9.16. Hvordan velger man en ortogonal basis? Svaret er Gram–Schmidts metode, temaet i neste seksjon. Men vi klarer å regne på et enkelt eksempel:

Eksempel 9.17. Betrakt (x, y)–planet som et underrom av $\mathbb{R}^3$ – ved å sette $z = 0$. Vektorene

$$\mathbf{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \text{ og } \mathbf{e}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

er en ortonormal basis. Formelen gir

$$P_{(x,y)\text{–planet}}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) = P_{\mathbf{e}_1}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) + P_{\mathbf{e}_2}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right).$$

Projeksjonen $P_{\mathbf{e}_i}$ gir enkelt og greit gir ut den i-te komponenten:

$$P_{(x,y)\text{–planet}}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) = \begin{bmatrix}v_1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\v_2\\0\end{bmatrix} = \begin{bmatrix}v_1\\v_2\\0\end{bmatrix}.$$

Resultatet er – naturlig nok – lineærtransformasjonen som dropper siste komponent. Standardmatrisen er

$$[P_{(x,y)\text{–planet}}] = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}.$$

△

La U være et underrom av et indreproduktrom V. Vi definerer det *ortogonale komplementet* til U i V som

$$U^\perp = \{\text{ Alle } \mathbf{v} \text{ i } V \text{ som er ortogonal på alle } \mathbf{u} \text{ i } U\}.$$

En vektor $\mathbf{v}$ er altså i $U^\perp$ dersom $\langle \mathbf{v}, \mathbf{u}\rangle = 0$ for alle $\mathbf{u}$ i U. Det kan sjekkes at det ortogonale komplementet faktisk er et underrom av V. Det neste teoremet sier at det holder å sjekke om en vektor står normalt på en basis.

Teorem 9.18. *La* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *være en basis for et underrom* U*. En vektor* $\mathbf{v}$ *er i det ortogonale komplementet hvis og bare hvis* $\langle \mathbf{u}_i, \mathbf{v}\rangle = 0$ *for alle* i*.*

Eksempel 9.19. La oss finne det ortogonale komplementet til (x, y)–planet i $\mathbb{R}^3$. Det består av vektorer $\begin{bmatrix}a\\b\\c\end{bmatrix}$ slik at

$$\begin{bmatrix}a\\b\\c\end{bmatrix} \cdot \begin{bmatrix}1\\0\\0\end{bmatrix} = 0$$

og

$$\begin{bmatrix}a\\b\\c\end{bmatrix} \cdot \begin{bmatrix}0\\1\\0\end{bmatrix} = 0.$$

Venstre side er henholdsvis a og b. Kravet for å være i det ortogonale komplementet er $a = b = 0$; det ortogonale komplementet er z–aksen. △

Før neste teorem kan det være nyttig å tenke over hvorfor det ortogonale komplementet til en linje i $\mathbb{R}^3$ er et plan; hvorfor det ortogonale komplementet til et plan i $\mathbb{R}^3$ er en linje.

Teorem 9.20. *La* U *være et underrom av et indreproduktrom* V*. Da gjelder*

$$\dim U + \dim U^\perp = \dim V.$$

Gram-Schmidts metode

La $\mathbf{v}_1$, $\mathbf{v}_2$, ...,$\mathbf{v}_n$ være en lineært uavhengig vektormengde. Vi skal lage oss en ortogonal basis $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ for rommet utspent av vektorene i mengden. Vi begynner med å definere

$$\mathbf{u}_1 = \mathbf{v}_1$$

Vektoren $\mathbf{v}_2$ er ikke nødvendigvis ortogonal på $\mathbf{u}_1$, men

$$\mathbf{u}_2 = \mathbf{v}_2 - P_{\mathbf{u}_1}(\mathbf{v}_2) = \mathbf{v}_2 - \frac{\langle \mathbf{u}_1, \mathbf{v}_2\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1$$

er – fordi vi trakk fra komponenten langs $\mathbf{u}_1$. Vektoren

$$\begin{aligned}\mathbf{u}_3 &= \mathbf{v}_3 - P_{\mathbf{u}_1}(\mathbf{v}_3) - P_{\mathbf{u}_2}(\mathbf{v}_3) \\ &= \mathbf{v}_3 - \frac{\langle \mathbf{u}_1, \mathbf{v}_3\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 - \frac{\langle \mathbf{u}_2, \mathbf{v}_3\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2\end{aligned}$$

er ortogonal på både $\mathbf{u}_1$ og $\mathbf{u}_2$ – fordi vi trakk fra komponentene langs $\mathbf{u}_1$ og $\mathbf{u}_2$. De tre vektorene $\mathbf{u}_1$,

$\mathbf{u}_2$ og $\mathbf{u}_3$ spenner ut det samme rommet som $\mathbf{v}_1$, $\mathbf{v}_2$ og $\mathbf{v}_3$, for opplagt er $\mathbf{v}_3$ en lineærkombinasjon av $\mathbf{u}_1, \mathbf{u}_2$ og $\mathbf{u}_3$. Nå kan vi fortsette slik, og definere rekursivt

$$\begin{aligned}\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} P_{\mathbf{u}_j}(\mathbf{v}_k)\\ &= \mathbf{v}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{u}_j, \mathbf{v}_k\rangle}{\langle \mathbf{u}_j, \mathbf{u}_j\rangle}\mathbf{u}_j.\end{aligned}$$

Det er fortsatt klart at $\mathbf{u}_1, \cdots, \mathbf{u}_n$ spenner ut det samme rommet som $\mathbf{v}_1, \cdots \mathbf{v}_n$.

Teorem 9.21. *Mengden* $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ *er en ortogonal basis for rommet utspent av* $\mathbf{v}_1$, $\mathbf{v}_2$, ...,$\mathbf{v}_n$.

Bevis. Vi mangler bare å se at basisen er ortogonal. Det er klart at $\mathbf{u}_1$ er en ortogonal basis for $\mathrm{Span}(\mathbf{v}_1)$. Vi antar derfor at det er sant at mengden $\{\mathbf{u}_1, \cdots, \mathbf{u}_{k-1}\}$ er ortogonal og viser at for $j < k$ er $\langle \mathbf{u}_j, \mathbf{v}_k\rangle = 0$.

$$\begin{aligned}\langle \mathbf{u}_j, \mathbf{u}_k\rangle &= \langle \mathbf{u}_j, (\mathbf{v}_k - \sum_{m=1}^{k-1} \frac{\langle \mathbf{u}_m, \mathbf{v}_k\rangle}{\langle \mathbf{u}_m, \mathbf{u}_m\rangle}\mathbf{u}_m)\rangle\\ &= \langle \mathbf{u}_j, \mathbf{v}_k\rangle - \sum_{m=1}^{k-1} \frac{\langle \mathbf{u}_m, \mathbf{v}_k\rangle}{\langle \mathbf{u}_m, \mathbf{u}_m\rangle}\langle \mathbf{u}_j, \mathbf{u}_m\rangle\\ &= \langle \mathbf{u}_j, \mathbf{v}_k\rangle - \langle \mathbf{u}_j, \mathbf{v}_k\rangle = 0.\end{aligned}$$

□

Eksempel 9.22. Vi finner en ortogonal basis for underrommet U av $\mathbb{R}^4$ utspent av

$$\mathbf{v}_1 = \begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}1\\0\\-2\\1\end{bmatrix} \text{ og } \mathbf{v}_3 = \begin{bmatrix}1\\1\\1\\0\end{bmatrix}.$$

Ta $\mathbf{u}_1 = \mathbf{v}_1$. Observer at

$$\mathbf{u}_1 \cdot \mathbf{v}_2 = 2\cdot 1 + 1\cdot 0 + 1\cdot(-2) + 0\cdot 1 = 0.$$

De er allerede ortogonale, så neste basisvektor blir bare

$$\mathbf{u}_2 = \mathbf{v}_2 - \frac{\mathbf{u}_1\cdot\mathbf{v}_2}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 = \mathbf{v}_2.$$

Siste basisvektor er

$$\begin{aligned}\mathbf{u}_3 &= \mathbf{v}_3 - \frac{\mathbf{u}_1\cdot\mathbf{v}_3}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 - \frac{\mathbf{u}_2\cdot\mathbf{v}_3}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2\\
&= \begin{bmatrix}1\\1\\1\\0\end{bmatrix} - \frac{\begin{bmatrix}2\\1\\1\\0\end{bmatrix}\cdot\begin{bmatrix}1\\1\\1\\0\end{bmatrix}}{\begin{bmatrix}2\\1\\1\\0\end{bmatrix}\cdot\begin{bmatrix}2\\1\\1\\0\end{bmatrix}}\begin{bmatrix}2\\1\\1\\0\end{bmatrix}\\
&\quad - \frac{\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}\cdot\begin{bmatrix}1\\1\\1\\0\end{bmatrix}}{\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}\cdot\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}}\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}\\
&= \begin{bmatrix}1\\1\\1\\0\end{bmatrix} - \frac{4}{6}\begin{bmatrix}2\\1\\1\\0\end{bmatrix} - \frac{-1}{6}\begin{bmatrix}1\\0\\-2\\1\end{bmatrix}\\
&= \frac{1}{6}(\begin{bmatrix}6\\6\\6\\0\end{bmatrix} + \begin{bmatrix}-8\\-4\\-4\\0\end{bmatrix} + \begin{bmatrix}1\\0\\-2\\1\end{bmatrix})\\
&= \frac{1}{6}\begin{bmatrix}-1\\2\\0\\1\end{bmatrix}.\end{aligned}$$

Vektorene

$$\begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-2\\1\end{bmatrix} \text{ og } \frac{1}{6}\begin{bmatrix}-1\\2\\0\\1\end{bmatrix}$$

er en basis for U. Den blir finere dersom vi skalerer $\mathbf{u}_3$ med 6: samlingen

$$\begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-2\\1\end{bmatrix} \text{ og } \begin{bmatrix}-1\\2\\0\\1\end{bmatrix}$$

er også en ortogonal basis for U. △

Beregne den ortogonale projeksjonen

La U være et underrom til et indreproduktrom V. Her er en metode for å regne ut den ortogonale projeksjonen P_U:

1. Ta utgangspunkt i en basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ som spenner ut U.
2. Bruk så Gram-Schmidt til å finne en ortogonal basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ for U.
3. Da kan man, for alle $\mathbf{x}$ i V, beregne den ortogonale projeksjonen

$$P_U(\mathbf{x}) = P_{\mathbf{u}_1}(\mathbf{x}) + P_{\mathbf{u}_2}(\mathbf{x}) + \cdots + P_{\mathbf{u}_n}(\mathbf{x}).$$

4. Hvis $V = \mathbb{R}^n$, så får vi standardmatrisen
$$[P_U] = \begin{bmatrix} P_U(\mathbf{e}_1) & P_U(\mathbf{e}_2) & \dots & P_U(\mathbf{e}_n) \end{bmatrix},$$
og da kan man, for alle $\mathbf{x}$ i V, beregne den ortogonale projeksjonen $P_U(\mathbf{x}) = [P_U]\mathbf{x}$.

Eksempel 9.23. Finn standardmatrisen til P_U hvor U er underrommet av $\mathbb{R}^4$ utspent av
$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \text{ og } \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$
Vi fant en ortogonal basis i Eksempel 9.22:
$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \text{ og } \mathbf{u}_3 = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$
Regn ut $P_U(\mathbf{e}_1)$:
$$\begin{aligned}
P_U(\mathbf{e}_1) &= P_{\mathbf{u}_1}(\mathbf{e}_1) + P_{\mathbf{u}_2}(\mathbf{e}_1) + P_{\mathbf{u}_3}(\mathbf{e}_1) \\
&= \frac{\begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}{\begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \frac{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \\
&\quad + \frac{\begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}{\begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \\
&= \frac{2}{6} \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} + \frac{-1}{6} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \\
&= \frac{1}{6} \begin{bmatrix} 6 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned}$$
Siden $P_U(\mathbf{e}_1) = \mathbf{e}_1$ skjønner vi at $\mathbf{e}_1$ ligger i U. En del regning gir at $P_U(\mathbf{e}_2)$, $P_U(\mathbf{e}_3)$ og $P_U(\mathbf{e}_4)$ er henholdsvis
$$\frac{1}{6} \begin{bmatrix} 0 \\ 5 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{6} \begin{bmatrix} 0 \\ 1 \\ 5 \\ 2 \end{bmatrix} \text{ og } \frac{1}{6} \begin{bmatrix} 0 \\ 2 \\ -2 \\ 2 \end{bmatrix}.$$
Standardmatrisen er
$$[P_U] = \frac{1}{6} \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 5 & 1 & 2 \\ 0 & 1 & 5 & -2 \\ 0 & 2 & 2 & 2 \end{bmatrix}.$$
△

Komplekse indreprodukt

Hva skjer om vi definerer prikkproduktet i $\mathbb{C}^n$ på samme måte som i $\mathbb{R}^n$? La oss prøve: Et produkt
$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n,$$
hvor v_i-ene og w_j-ene er komplekse tall, tilfredstiller ikke nødvendigvis at $\mathbf{v} \cdot \mathbf{v}$ er et reelt tall. Eksempelvis er
$$\begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix} = e^{i\frac{\pi}{4}} e^{i\frac{\pi}{4}} = i.$$
Dette er et problem hvis vi ønsker å definere lengde som i det reelle tilfellet. Vi kan med andre ord ikke definere
$$\| \begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix} \|^2 = \begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix}.$$
Det finnes en enkel måte å fikse dette problemet. Husk at prikkproduktet i $\mathbb{R}^n$ er et matriseprodukt
$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^\top \mathbf{w}.$$
Nå konjugerer vi $\mathbf{v}$, i tillegg til å transponere den,
$$\mathbf{v}^* = \begin{bmatrix} \overline{v_1} & \overline{v_2} & \dots & \overline{v_n}. \end{bmatrix}$$
Vektoren $\mathbf{v}^*$ kalles den *adjungerte* til $\mathbf{v}$. Nå kan vi bruke matrisemultiplikasjon til å definere et produkt
$$\mathbf{v}^* \mathbf{w} = \overline{v_1} w_1 + \overline{v_2} w_2 + \cdots + \overline{v_n} w_n.$$
Merk deg at
$$\begin{aligned} \mathbf{v}^* \mathbf{v} &= \overline{v_1} v_1 + \overline{v_2} v_2 + \cdots + \overline{v_n} v_n \\ &= |v_1|^2 + |v_2|^2 + \cdots + |v_n|^2 \end{aligned}$$
alltid er et ikke-negativt reelt tall; summen av lengden til komponentene. Eksempelvis er
$$\begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix}^* \begin{bmatrix} e^{i\frac{\pi}{4}} \\ 0 \end{bmatrix} = \overline{e^{i\frac{\pi}{4}}} e^{i\frac{\pi}{4}} = e^{-i\frac{\pi}{4}} e^{i\frac{\pi}{4}} = 1.$$
Vi har ikke lengre symmetri, $\mathbf{v}^* \mathbf{w} \neq \mathbf{w}^* \mathbf{v}$, men vi har *konjugert symmetri*
$$\mathbf{v}^* \mathbf{w} = \overline{\mathbf{w}^* \mathbf{v}}.$$
Ellers kan du sjekke at vi fortsatt har linearitet i andre variabel
$$\mathbf{v}^*(a\mathbf{w} + b\mathbf{u}) = a\mathbf{v}^* \mathbf{w} + b\mathbf{v}^* \mathbf{u},$$
og positivitet
$$\mathbf{v}^* \mathbf{v} \geq 0, \text{ og } \mathbf{v}^* \mathbf{v} = 0 \text{ kun hvis } \mathbf{v} = \mathbf{0}.$$
Basert på motivasjonen fra det reelle indreproduktet virker det rimelig at dette fortsatt gir en rik geometri som blant annet inneholder lengde, ortogonalitet og projeksjoner.

Definisjon. La V være et komplekst vektorrom. Et *indreprodukt* i V er en operasjon som tar inn to vektorer, $\mathbf{v}$ og $\mathbf{w}$, for å gi ut et komplekst tall $\langle \mathbf{v}, \mathbf{w} \rangle$. Operasjonen tilfredstiller

$$\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle} \qquad \text{(konj. sym.)}$$
$$\langle \mathbf{v}, (a\mathbf{w} + b\mathbf{u}) \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{u} \rangle \qquad \text{(linearitet)}$$
$$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0, \text{ og } \langle \mathbf{v}, \mathbf{v} \rangle = 0 \text{ kun hvis } \mathbf{v} = \mathbf{0} \qquad \text{(positivitet)}$$

Vi sier at V, sammen med et valgt indreprodukt, er et *indreproduktrom.* △

Merk. Positivitet viser igjen at den eneste vektoren som er ortogonal med *alle* vektorer i V er nullvektoren. △

Vi skal kun fokusere på ett komplekst eksempel, nemlig $\mathbb{C}^n$.

Eksempel 9.24. Operasjonen

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v} \cdot \mathbf{w} = \mathbf{v}^* \mathbf{w},$$

er et indreprodukt i $\mathbb{C}^n$. Til sammen utgjør de et komplekst indreproduktrom. △

Merk. Dersom $\mathbf{v}$ og $\mathbf{w}$ er reelle, blir

$$\mathbf{v}^* \mathbf{w} = \mathbf{v}^T \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots + v_n w_n = \mathbf{v} \cdot \mathbf{w},$$

det gode gamle skalarproduktet. Fra nå av er $\mathbf{v}^*\mathbf{w}$ en fellesbetegnelse for skalarproduktet i $\mathbb{R}^n$ og indreproduktet i $\mathbb{C}^n$.

Merk. Nesten alt vi gjorde for reelle indreproduktrom gjelder også for komplekse indreproduktrom. Det er ett unntak, vinkelen er ikke definert. Grunnen er at Caucy–Schwarz, Teorem 9.10, impliserer at $\langle \mathbf{v}, \mathbf{w} \rangle$ er et komplekst tall med lengde mindre eller lik 1. I det reelle tilfellet befinner vi oss dermed i intervallet $[0, 1]$, men i det komlekse tilfellet er det flere tall som er inkludert; disken som er sentrert i 0 med radius 1.

Teorem 9.25. *Alt som gjelder for reelle indreproduktrom, med unntak av vinkelen, gjelder også for komplekse indreproduktrom.*

Eksempel 9.26. Vektorene

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{og} \quad \begin{bmatrix} i \\ 1 \end{bmatrix}$$

er ortogonale:

$$\begin{bmatrix} 1 \\ i \end{bmatrix}^* \begin{bmatrix} i \\ 1 \end{bmatrix} = \overline{1} \cdot i + \overline{i} \cdot 1 = i - i = 0.$$

△

Eksempel 9.27. La oss projisere vektoren

$$\mathbf{w} = \begin{bmatrix} -5i \\ 0 \\ 2i \end{bmatrix}$$

både på og normalt på

$$\mathbf{v} = \begin{bmatrix} 3 \\ -i \\ 4 \end{bmatrix}.$$

Vi beregner:

$$\mathbf{v}^* \mathbf{v} = 3 \cdot 3 + i \cdot (-i) + 4 \cdot 4 = 26$$

og

$$\mathbf{v}^* \mathbf{w} = 3 \cdot (-5i) + i \cdot 0 + 4 \cdot 2i = -7i$$

slik at

$$P_{\mathbf{v}} \mathbf{w} = \frac{\mathbf{v}^* \mathbf{w}}{\mathbf{v}^* \mathbf{v}} \mathbf{v} = \frac{-7i}{26} \begin{bmatrix} 3 \\ -i \\ 4 \end{bmatrix}$$

og

$$\begin{aligned} \mathbf{w} - P_{\mathbf{v}}(\mathbf{w}) &= \mathbf{w} - \frac{\mathbf{v}^* \mathbf{w}}{\mathbf{v}^* \mathbf{v}} \mathbf{v} \\ &= \begin{bmatrix} -5i \\ 0 \\ 2i \end{bmatrix} - \frac{-7i}{26} \begin{bmatrix} 3 \\ -i \\ 4 \end{bmatrix} = \frac{1}{26} \begin{bmatrix} -109i \\ 7 \\ 80i \end{bmatrix} \end{aligned}$$ △

Vi avslutter seksjonen med en liten diskusjon om adjungering, noe du kanskje kommer til å møte på i fremtiden. Man kan definere den *adjungerte* til

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

som matrisen

$$A^* = \begin{bmatrix} \overline{a}_{11} & \overline{a}_{21} & \cdots & \overline{a}_{m1} \\ \overline{a}_{12} & \overline{a}_{22} & \cdots & \overline{a}_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{1n} & \overline{a}_{2n} & \cdots & \overline{a}_{mn} \end{bmatrix}$$

der radene og kolonnene i A er byttet om, og alt er komplekskonjugert. Du kan sjekke at siden $(AB)^\top = B^\top A^\top$, så gjelder

$$(A\mathbf{x})^* \mathbf{y} = \mathbf{x}^* (A^* \mathbf{y}).$$

Mer generelt, la V og W være indreproduktrom, og la T være en lineærtransformasjon mellom dem. Den *adjungerte* til T, T^*, er lineærtransformasjonen som tilfredstiller

$$\langle T(\mathbf{x}), \mathbf{y} \rangle = \langle \mathbf{x}, T^*(\mathbf{y}) \rangle.$$

Eksempel 9.28. Hvis vi lar A være matrisen

$$A = \begin{bmatrix} 5i & 0 & -2i \\ 3 & i & 4 \end{bmatrix},$$

så er den adjungerte av A^* gitt ved

$$A^* = \begin{bmatrix} -5i & 3 \\ 0 & -i \\ 2i & 4 \end{bmatrix}.$$

Hvis vi adjungerer denne matrisen igjen, så kommer vi tilbake til utgangspunktet:

$$(A^*)^* = A$$ △

Merk. $A = A^{**}$

Det siste resultatet i denne seksjonen er grunnlaget for minste kvadraters metode – som er en anvendelse i et senere kapittel.

La A være en $m \times n$-matrise. Da observerer vi en sammenheng mellom kolonnerommet til A^* og nullrommet til A. For hvis $\mathbf{w} = A^* \mathbf{u}$ og $A\mathbf{v} = \mathbf{0}$, så er

$$\begin{aligned} \mathbf{w}^* \mathbf{v} &= (A^* \mathbf{u})^* (\mathbf{v}) = \mathbf{u}^* (A^{**} \mathbf{v}) \\ &= \mathbf{u}^* (A\mathbf{v}) = \mathbf{u}^* \mathbf{0} = 0. \end{aligned}$$

Dette viser at $\mathbf{w}$, som ligger i $\operatorname{Col} A^*$, er ortogonal med $\operatorname{Null} A$. Faktisk gjelder følgende likheter:

Teorem 9.29. *La A være en $m \times n$-matrise. Da vet vi*

$$(\operatorname{Col} A)^{\perp} = \operatorname{Null} A^*$$
$$(\operatorname{Null} A)^{\perp} = \operatorname{Col} A^*.$$

Bevis. Det holder å bevise en av påstandene. For den andre følger ved å se på A^* istedet for A. Så la oss bevise $(\operatorname{Col} A)^{\perp} = \operatorname{Null} A^*$. Vi skal vise denne påstanden ved å sjekke at $(\operatorname{Col} A)^{\perp}$ er en delmengde av $\operatorname{Null} A^*$ og at samtidig $\operatorname{Null} A^*$ er en delmengde av $(\operatorname{Col} A)^{\perp}$.

Vi begynner med at $\mathbf{x}$ er en vektor i $(\operatorname{Col} A)^{\perp}$. La $\mathbf{a}_1, \ldots, \mathbf{a}_n$ være kolonnevektorene i A. Kolonnerommet $\operatorname{Col} A$ er utspent av kolonnene i A. Altså er $\mathbf{x}$ ortogonal med alle $\mathbf{a}_i$, dvs

$$0 = \langle \mathbf{a}_i, \mathbf{x} \rangle = \mathbf{a}_i^* \mathbf{x} \text{ for alle } i = 1, \ldots, n.$$

Men dette betyr per definisjon at $A^*\mathbf{x} = \mathbf{0}$. Altså ligger $\mathbf{x}$ i $\operatorname{Null} A^*$.

En annen måte å vise dette er å observere at $\mathbf{x} \in (\operatorname{Col} A)^{\perp}$ betyr at

$$\langle A\mathbf{u}, \mathbf{x} \rangle = 0 \text{ for alle } \mathbf{u} \in \mathbb{C}^n.$$

Men vi vet

$$\langle A\mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{u}, A^*\mathbf{x} \rangle.$$

Altså har vi

$$\langle \mathbf{u}, A^*\mathbf{x} \rangle = 0 \text{ for alle } \mathbf{u} \in \mathbb{C}^n.$$

Vektoren $A^*\mathbf{x}$ er altså ortogonal med *alle* vektorer i $\mathbb{C}^n$. På grunn av positivitet må $A^*\mathbf{x}$ være nullvektoren, dvs. at $\mathbf{x}$ ligger i $\operatorname{Null} A^*$.

Omvendt la oss anta at $\mathbf{x}$ er en vektor i $\operatorname{Null} A^*$. Dette betyr at $A^*\mathbf{x} = \mathbf{0}$ og viser ved definisjon av multiplikasjonen matrise gang vektor at $\mathbf{x}$ ortogonal med alle kolonnene i A. Med andre ord $\mathbf{x}$ ligger i $(\operatorname{Col} A)^{\perp}$. $\square$

Merk. For en *reell* $m \times n$-matrise A sier teoremet:

$$(\operatorname{Col} A)^{\perp} = \operatorname{Null} A^T$$
$$(\operatorname{Null} A)^{\perp} = \operatorname{Col} A^T.$$

$\triangle$

Indreprodukt mellom funksjoner

Denne seksjonen inneholder et eksempel på et nytt indreprodukt. Vi skal definere et indreprodukt av kontinuerlige funksjoner. Med konseptuelt små, men litt tekniske justeringer kan man definere et indreprodukt av stykkevis kontinuerlige funksjoner. Dette indreproduktet har mange anvendelser, spesielt innen signalbehandling. Du skal lære mer om dette i matematikk 4, stikkordet er *fourieranalyse*: Ideen er å projisere signaler – les stykkevis kontinuerlige 2π-periodiske funksjoner – ned på de elementære signalene $\cos(nx)$ og $\sin(mx)$ for å plukke ut hvor mye av hver frekvens signalet består av. Det viser seg at fine nok signaler kan rekonstrueres på denne måten.

Skalarproduktet i $\mathbb{R}^n$ er en sum. Det er nærmere bestemt en sum av produktet mellom komponentene til to vektorer. En funksjon fra et intervall $[a, b]$ til $\mathbb{R}$ – en vektor i $\mathcal{C}([a, b])$ – kan tenkes på som en 'klassisk vektor', eller pil, med overtellbart mange komponenter: En funksjon f, fra $[a, b]$ til $\mathbb{R}$, er en regel som gir ut ett tall $f(x)$ for hver x i $[a, b]$, og det er overtellbart mange tall i dette intervallet. Vi kunne godt ha skrevet dette som en samling av komponenter

$$(f(x) \text{ hvor } x \text{ ligger i } [a, b]).$$

Det er umulig å summere overtellbart mange tall; umulig å summere over alle komponentene – funksjonsverdiene – til en funksjon. Men i matematikk 1 så vi at det finnes noe som minner om en slik sum, nemlig integralet over $[a, b]$: Integralet er arealet under grafen, summen av uendelig tynne stolper med tilhørende funksjonsverdi som høyde.

Basert på motivasjonen ovenfor forsøker vi å definere indreproduktet mellom f og g som

$$\langle f, g \rangle = \frac{1}{b-a} \int_a^b f(x)g(x)dx.$$

Faktoren $\frac{1}{b-a}$ er ikke nødvendig, men den gir finere formler i praksis. Vi må sjekke at aksiomene for et indreprodukt holder.

Teorem 9.30. *Operasjonen*

$$\langle f, g \rangle = \frac{1}{b-a} \int_a^b f(x)g(x)dx.$$

er et indreprodukt på $\mathcal{C}([a, b])$.

Bevis. Symmetri:

$$\begin{aligned}
\langle f, g \rangle &= \frac{1}{b-a} \int_a^b f(x)g(x)dx \\
&= \frac{1}{b-a} \int_a^b g(x)f(x)dx \\
&= \langle g, f \rangle.
\end{aligned}$$

Linearitet: I matematikk 1 lærte du at integralet er lineært, det vil si

$$\int (cg(x) + dh(x))dx = c \int g(x)dx + d \int h(x)dx,$$

som gir

$$\begin{aligned}
\langle f, cg + dh \rangle &= \frac{1}{b-a} \int_a^b f(x)(cg(x) + dh(x))dx \\
&= \frac{1}{b-a} \int_a^b (cf(x)g(x) + df(x)h(x))dx \\
&= c\frac{1}{b-a} \int_a^b f(x)g(x)dx \\
&\quad + d\frac{1}{b-a} \int_a^b f(x)h(x)dx \\
&= c\langle f, g \rangle + d\langle f, h \rangle.
\end{aligned}$$

Positivitet:

$$\langle f, f\rangle = \frac{1}{b-a}\int_a^b f(x)^2 dx$$

Funksjonen $f(x)^2$ er alltid minst null. Derfor blir arealet under grafen minst null. Det vil si at $\langle f, f\rangle \geq 0$. Når er $\langle f, f\rangle = 0$? Intuitivt er det klart at om en ikke-negativ kontinuerlig funksjon ikke er konstant null vil arealet under grafen være større enn null. Mer presist: dersom $f(x_0) \neq 0$ for en eller annen x_0 setter vi $\epsilon = \frac{1}{2}f(x_0)^2$. Kontinuitet av $f(x)^2$ gir oss en δ, $0 < \delta < 1$, slik at $f(x)^2 > \epsilon$ for $|x - x_0| < \delta$. Dermed er arealet under grafen til $f(x)^2$ minst $\delta\epsilon > 0$ og vi konkluderer med at dersom $\langle f, f\rangle = 0$ må vi ha $f = 0$. □

Det er på tide med et eksempel.

Eksempel 9.31. Vi ser på indreproduktet når $a = 0$ og $b = 1$. Er x og x^2 ortogonale?

$$\begin{aligned}\langle x, x^2\rangle &= \frac{1}{1-0}\int_0^1 x \cdot x^2 dx = \int_0^1 x^3 dx \\ &= [\frac{1}{4}x^4]_0^1 = \frac{1}{4}\end{aligned}$$

Nei, de er ikke ortogonale. Hva er lengden til x og x^2?

$$\begin{aligned}\langle x, x\rangle &= \frac{1}{1-0}\int_0^1 x \cdot x dx = \int_0^1 x^2 dx \\ &= [\frac{1}{3}x^3]_0^1 = \frac{1}{3}\end{aligned}$$

$$\begin{aligned}\langle x^2, x^2\rangle &= \frac{1}{1-0}\int_0^1 x^2 \cdot x^2 dx = \int_0^1 x^4 dx \\ &= [\frac{1}{5}x^5]_0^1 = \frac{1}{5}\end{aligned}$$

Lengdene er

$$\|x\| = \frac{1}{\sqrt{3}} \text{ og } \|x^2\| = \frac{1}{\sqrt{5}}.$$

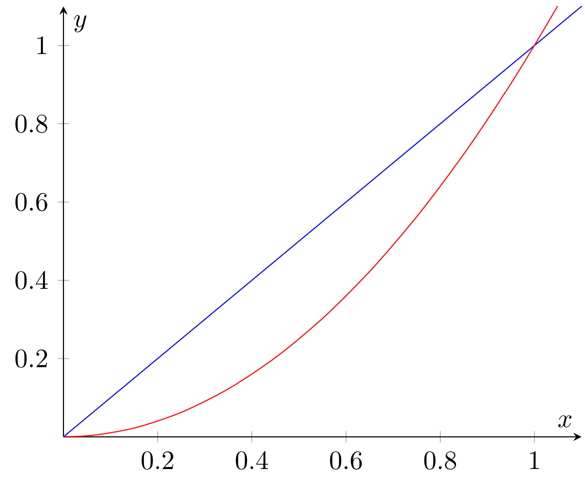
Hva er vinkelen mellom dem?

$$\begin{aligned}\cos\theta &= \frac{\langle x, x^2\rangle}{\|x\|\|x^2\|} = \frac{\frac{1}{4}}{\frac{1}{\sqrt{3}}\frac{1}{\sqrt{5}}} \\ &= \frac{\sqrt{3}\sqrt{5}}{4} = \frac{\sqrt{15}}{4}.\end{aligned}$$

Merk at $\sqrt{15} \approx 3.87 \leq 4$ som stemmer overens med Cauchy–Schwarz. En kalkulator gir $\theta \approx 14.48$ grader. Vi kan ikke se denne vinkelen i figuren nedenfor, men vi kan forstå at de er nærmere å være parallelle, enn de er å være ortogonale. Hvis de hadde vært parallelle, ville θ vært 0 eller 180 grader, og hvis de hadde vært ortogonale ville θ vært 90 grader. Dette gir også et bevis på at x og x^2 er lineært uavhengige; de er ikke på samme linje/parallelle. La oss tenke litt mer på denne vinkelen. Siden lengden av x og x^2 er henholdsvis $\frac{1}{\sqrt{2}}$ og $\frac{1}{\sqrt{3}}$, og vinkelen mellom dem er 14.48 grader, ser de ut til å være nærme hverandre. Avstanden mellom dem er – som i $\mathbb{R}^2$ – lengden til differansen;

$$\|x - x^2\|^2 = \int_0^1 (x - x^2)^2 dx$$

Fra figuren ser dette målet på avstand ut til å være lite:



Arealet mellom x (blå) og x^2 (rød) er lite

Mer presist:

$$\begin{aligned}\|x - x^2\|^2 &= \int_0^1 (x - x^2)^2 dx \\ &= \int_0^1 (x^2 - 2x^3 + x^4) dx \\ &= \frac{1}{3} - \frac{2}{4} + \frac{1}{5} \\ &= \frac{1}{30}.\end{aligned}$$

Avstanden mellom x og x^2 er $\frac{1}{\sqrt{30}} \approx 0.183$, et lite tall. For å oppsummere, grafene i figuren over ser ut til å være nærme hverandre, derfor er vinkelen mellom dem relativt liten. Hva er den ortogonale projeksjonen av x^2 på x?

$$P_x(x^2) = \frac{\langle x, x^2\rangle}{\langle x, x\rangle}x = \frac{\frac{1}{4}}{\frac{1}{3}}x = \frac{3}{4}x.$$

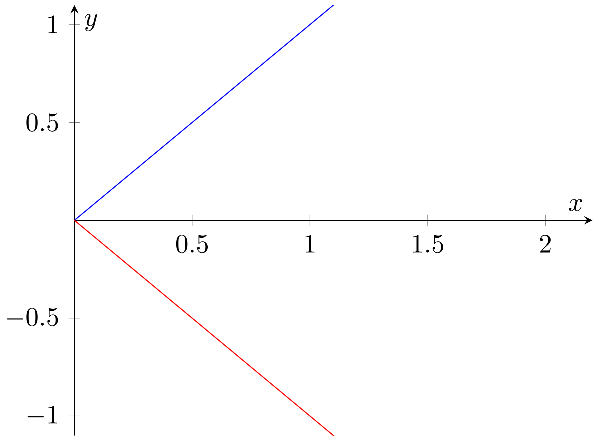
Vi sjekker at $P_x(x^2)$ og $x^2 - P_x(x^2)$ faktisk er ortogonale:

$$\begin{aligned}\langle P_x(x^2), x^2 - P_x(x^2)\rangle &= \frac{1}{1-0}\int_0^1 \frac{3}{4}x(x^2 - \frac{3}{4}x) dx \\ &= \frac{3}{4}\int_0^1 x^3 - \frac{9}{16}\int_0^1 x^2 dx \\ &= \frac{3}{4}\frac{1}{4} - \frac{9}{16}\frac{1}{3} \\ &= \frac{3}{16} - \frac{3}{16} \\ &= 0\end{aligned}$$

△

Merk. Husk at vi jobber i $\mathcal{C}([a,b])$, ikke i $\mathbb{R}^2$. Derfor stemmer ikke lengder og vinkler med hvordan vi forestiller oss funksjonene i $\mathbb{R}^2$. Det er dermed enkelt å bli lurt til å la vår geometriske intuisjon fra det reelle planet til å spille oss et puss når vi jobber med funksjonsrom. Husk dette når vi går videre.

Eksempel 9.32. Hva er vinkelen mellom x og $-x$ i $\mathcal{C}_s([0,1])$? Man skulle kanskje tro at vinkelen er 90 grader:



Vinkelen mellom grafene til x (blå) og $-x$ (rød) er 90 grader i $\mathbb{R}^2$

Men husk at vinkelen ikke er en vinkel mellom grafene. Den riktige intuisjonen her er at vinkelen burde være 180 grader. Hvorfor? Linjen utspent av x – $\mathrm{Sp}\{x\}$ – består av alle ax hvor a er et reelt tall. Spesielt ligger $-x = (-1)\cdot x$ på linjen. Tenk på en vektor $\mathbf{v}$ i $\mathbb{R}^2$, vinkelen mellom $\mathbf{v}$ og $-\mathbf{v}$ er 180 grader.

$$\begin{aligned}\cos\theta &= \frac{\langle x,-x\rangle}{\|x\|\|-x\|}\\ &= \frac{\int\limits_0^1 x\cdot(-x)dx}{\sqrt{\int\limits_0^1 x^2dx}\sqrt{\int\limits_0^1(-x)^2dx}}\\ &= \frac{-\int\limits_0^1 x^2dx}{\int\limits_0^1 x^2dx}\\ &= -1\end{aligned}$$

Dette betyr at θ er 180 grader. △

Eksempel 9.33. Vi finner en ortogonal basis for underrommet U av $\mathcal{C}_s([0,1])$ utspent av 1, x og x^2. Start med $\mathbf{u}_1 = x$ for å kunne gjenbruke utregninger fra Eksempel 9.31. Ta $\mathbf{v}_2 = x^2$:

$$\begin{aligned}\mathbf{u}_2 = \mathbf{v}_2 - \frac{\langle\mathbf{u}_1,\mathbf{v}_2\rangle}{\langle\mathbf{u}_1,\mathbf{u}_1\rangle}\mathbf{u}_1 &= x^2 - \frac{\langle x,x^2\rangle}{\langle x,x\rangle}x\\ &= x^2 - \frac{\frac{1}{4}}{\frac{1}{3}}x = x^2 - \frac{3}{4}x\end{aligned}$$

Du kan sjekke at

$$\langle x^2-\frac{3}{4}x, x^2-\frac{3}{4}x\rangle = \frac{1}{80}$$

og

$$\langle x^2-\frac{3}{4}x, 1\rangle = -\frac{1}{24}.$$

Siste basisvektor er

$$\begin{aligned}\mathbf{u}_3 &= 1 - \frac{\langle x,1\rangle}{\langle x,x\rangle}x - \frac{\langle x^2-\frac{3}{4},1\rangle}{\langle x^2-\frac{3}{4}x, x^2-\frac{3}{4}x\rangle}(x^2-\frac{3}{4}x)\\ &= 1 - \frac{\frac{1}{2}}{\frac{1}{3}}x - \frac{-\frac{1}{24}}{\frac{1}{80}}(x^2-\frac{3}{4}x)\\ &= \frac{10}{3}x^2 - \frac{12}{3}x + 1\end{aligned}$$

En ortogonal basis for U er

$$x,\ x^2-\frac{3}{4}x \text{ og } \frac{10}{3}x^2 - \frac{12}{3}x + 1,$$

eller, dersom vi skalerer de leddene som har brøker:

$$x,\ 4x^2 - 3x \text{ og } \frac{1}{0}x^2 - \frac{1}{2}3x + 3.$$

△

Eksempel 9.34. I dette eksempelet er $a = -\pi$ og $b = \pi$. Vi skal regne ut lengden til vektorene 1, $\cos x$ og $\sin x$, og se at de er parvis ortogonale.

$$\begin{aligned}\|1\|^2 &= \frac{1}{\pi-(-\pi)}\int\limits_{-\pi}^{\pi} 1\cdot 1dx\\ &= \frac{1}{2\pi}[x]_{-\pi}^{\pi}\\ &= \frac{\pi-(-\pi)}{2\pi}\\ &= 1.\end{aligned}$$

Husk den trigonometriske identiteten $\cos(2x) = 2\cos^2 x - 1$. Da blir

$$\begin{aligned}\|\cos x\|^2 &= \frac{1}{\pi-(-\pi)}\int\limits_{-\pi}^{\pi}\cos x\cdot\cos x dx\\ &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\cos^2 x dx\\ &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\frac{1+\cos(2x)}{2}dx\\ &= \frac{1}{2\pi}[\frac{1}{2}x+\frac{1}{4}\sin(2x)]_{-\pi}^{\pi}, \qquad \sin(\pm 2\pi) = 0\\ &= \frac{1}{2\pi}(\frac{1}{2}\pi - \frac{1}{2}(-\pi))\\ &= \frac{1}{2}.\end{aligned}$$

På lignende vis regner vi ut

$$\|\sin x\|^2 = \frac{1}{2}.$$

Vektorene 1 og $\cos x$ er ortogonale:

$$\begin{aligned}
\langle 1, \cos x\rangle &= \frac{1}{\pi-(-\pi)}\int\limits_{-\pi}^{\pi} 1\cdot\cos x dx \\
&= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\cos x dx \\
&= \frac{1}{2\pi}[\sin x]_{-\pi}^{\pi}, \qquad\qquad \sin(\pm\pi) = 0 \\
&= 0
\end{aligned}$$

Tilsvarende regning gir

$$\langle 1, \sin x\rangle = 0.$$

Bruk substitusjon for å regne ut det siste integralet:

$$\begin{aligned}
\langle \cos x, \sin x\rangle &= \frac{1}{\pi-(-\pi)}\int\limits_{-\pi}^{\pi}\cos x\cdot\sin x dx \\
&= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi}\cos x\sin x dx, \qquad\qquad u = \cos x \\
&= -\frac{1}{2\pi}\int\limits_{-1}^{1} u du \\
&= -\frac{1}{2\pi}[\frac{1}{2}u^2]_{-1}^{1} \\
&= -\frac{1}{2\pi}(\frac{1}{2}1^2 - \frac{1}{2}(-1)^2) \\
&= 0
\end{aligned}$$

$\triangle$

Du kan bruke lignende triks fra matematikk 1 for å bevise at $\cos(nx)$ og $\sin(mx)$, hvor $n, m = 2, 3, \ldots$, også kan tas med i Eksempel 9.34.

Teorem 9.35. *Vektorene* 1, $\cos(nx)$ *og* $\sin(mx)$, *hvor* $n, m = 1, 2, 3, \ldots$, *er parvis ortogonale i* $\mathcal{C}_s([-\pi, \pi])$.