

Kapittel 11

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.

Ortogonal projeksjon i \mathbb{R}^2

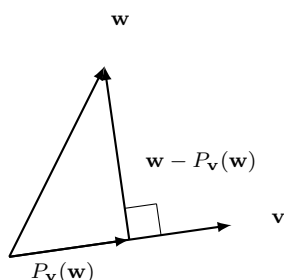
Vi husker skalarproduktet, eller prikkproduktet, fra gymnaset. Du har lært to måter å beregne indreproduktet, også kjent som prikkproduktet, nemlig

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

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} , eller

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

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



Hva er projeksjon?

Vi kan utlede en formel for lengden:

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

slik at

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

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 på \mathbf{v} er

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

Eksempel 11.1. 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 lengden $\mathbf{w} - P_{\mathbf{v}}(\mathbf{w})$:

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

Skalarproduktet i \mathbb{R}^n

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

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n,$$

som også kan uttrykkes ved matriseproduktet

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

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

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2},$$

eller ekvivalent

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

Formelen

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

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

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

og *parallele* hvis

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

– fordi dette betyr at vinkelen mellom dem er 0 eller 180 grader; de ligger på samme linje. Og 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 nullvektoren. 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 11.2. 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:

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

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} :

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

△

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 11.3. Den ortogonale projeksjonen på en vektor \mathbf{v} ,

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

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 $\mathbf{v}^T \mathbf{w}$. Det er med andre ord lineærtransformasjonen som er gitt av $1 \times n$ -matrisen \mathbf{v}^T . Dette gir automatisk at skalarproduktet er lineært (fordi matriseproduktet er distributivt). Du kan alternativt bevise påstanden direkte ved å bruke formelen $\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n$. Nå er det rett frem å vise at projeksjonen på \mathbf{v} er en lineærtransformasjon:

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

□

Eksempel 11.4. 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 lineærtransformasjon 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

$$[P_{\mathbf{v}}] = [P_{\mathbf{v}}(\mathbf{e}_1) \quad P_{\mathbf{v}}(\mathbf{e}_2) \quad P_{\mathbf{v}}(\mathbf{e}_3)]$$

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

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

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, sjekker vi 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 $P_{\mathbf{v}}(\mathbf{w}) = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$. Er $P_{\mathbf{v}}$ injektiv?

Hva med 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} , $\text{Sp}\{\mathbf{v}\}$, 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 $\text{Sp}\{\mathbf{v}\}$. 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 $[P_{\mathbf{v}}]$, er planet som står normalt på $\text{Sp}\{\mathbf{v}\}$ med origo som skjæringspunkt. Bildet til $P_{\mathbf{v}}$, eller kolonnerommet til $[P_{\mathbf{v}}]$, er $\text{Sp}\mathbf{v}$ – fordi vektorer projiseres ned på denne linjen. \triangle

Hvis du tenker deg om vil du etterhvert innse – eller kanskje bli lurt til å 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 skalarmultipliserer

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

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

Koordinater er altså bare 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 skalarmultiplikasjon. 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

Før matematikk 3 tenkte de fleste av oss på vektorer som piler i \mathbb{R}^2 , eller \mathbb{R}^3 . Disse pilene kan legges sammen og skaleres på en naturlig måte. I Kapittel 8 så vi hvordan disse ideene kan abstraheres til mer generelle vektorrom; mengder hvor man kan legge sammen og skalere elementer – som kalles vektorer – på en slik måte at regnereglene vi er vant med fra \mathbb{R}^2 , og \mathbb{R}^3 , fortsatt er gyldige.

Matematikere liker å generalisere.

Hvorfor? Et svar er at det ofte er nyttig å skjønne hvordan ting, som kanskje ser helt forskjellig ut, har like egenskaper. Til nå har vi blant annet sett at polynom, og mer generelt funksjoner, har visse egenskaper til felles med piler i \mathbb{R}^2 – alle er vektorrom.

Vi skal abstrahere ideen om skalarproduktet. For å bedre skjønne hvorfor dette er av interesse, kan det være lurt å undre litt over:

Hva tilfører skalarproduktet?

Et upresist svar er at det ser ut til å måle hvordan vektorer i \mathbb{R}^n orienterer seg i forhold til hverandre. Kanskje vi til og med kan komme med følgende vågale utsagn. Skalarproduktet er geometrien til \mathbb{R}^n . Hvis du tenker tilbake på forrige seksjon synes du kanskje ikke dette virker urimelig: Det vi ser for oss kan jo beskrives med skalarproduktet.

Alt dette er vel og bra. Men hvordan i alle dager kan man generalisere noe slikt? Det er kanskje litt enklere å tenke på:

Hvilke egenskaper til skalarproduktet er av geometrisk betydning?

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. For 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 $\mathbf{w} \cdot \mathbf{v} = 0$. I så fall måtte vi skilt mellom:

Vektoren \mathbf{v} er ortogonal med \mathbf{w} .

Vektoren \mathbf{w} er ortogonal med \mathbf{v} .

Noe som virker helt absurd. Hvordan kan \mathbf{v} være ortogonal med \mathbf{w} , mens \mathbf{w} ikke er ortogonal med \mathbf{v} ?

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

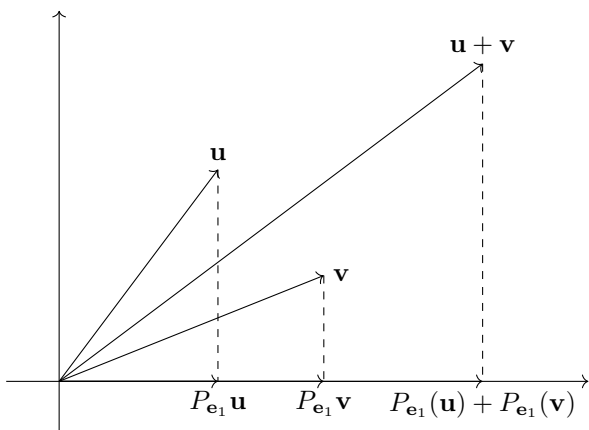
$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$$

siden kvadratrotten bare er definert for positive tall. Og 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.

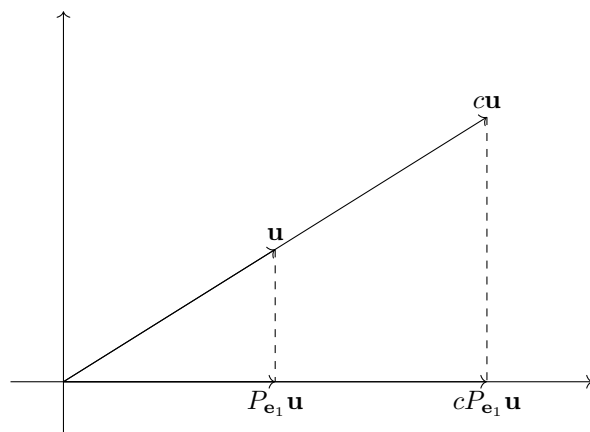
Den siste egenskapen vi skal tenke på er *linearitet*:

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

Skalarproduktet også er lineært i første faktor fordi det er symmetrisk. I beviset til Teorem 11.3 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

En liten oppsummering på diskusjonen ovenfor: 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 \mathbb{R}^n . Symmetri, linearitet og positivitet er tre meget viktige egenskaper til skalarproduktet i \mathbb{R}^n . Disse sørger for at ortogonalitet, projeksjon og lengde oppfører seg i henhold til intuisjonen vår. Derfor burde vi i alle fall kreve at generaliseringen skal tilfredstille disse tre. Dette er faktisk matematikerenes beste forslag på hva den riktige generaliseringen av skalarproduktet er. Konvensjonen er å kalle generaliseringen for et *indreprodukt* med notasjon $\langle \mathbf{v}, \mathbf{w} \rangle$ – indreproduktet mellom \mathbf{v} og \mathbf{w} .

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 ett valgt indreprodukt, er et *indreproduktrom*. \triangle

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 11.5. 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. Dette eksempelet var tross alt hele motivasjonen for definisjonen ovenfor. \triangle

En viktig lærdom fra seksjonen om skalarproduktet i \mathbb{R}^n , er at intuisjonen fra \mathbb{R}^2 fungerer helt ypperlig. Tror du dette også er sant for andre indreproduktrom?

Intuisjonen din fra \mathbb{R}^2 fungerer fortsatt helt ypperlig.

Det være lurt å tenke på $\mathbf{v} \cdot \mathbf{w}$ i stedet for $\langle \mathbf{v}, \mathbf{w} \rangle$ fremover.

Eksempel 11.6. I \mathbb{R}^2 er skalarproduktet med nullvektoren alltid null. Dette er også tilfellet i et generelt indreproduktrom. Et triks er å observere at

$$\begin{aligned} 2\langle \mathbf{v}, \mathbf{0} \rangle &= \langle \mathbf{v}, 2\mathbf{0} \rangle, & \text{linearitet} \\ &= \langle \mathbf{v}, \mathbf{0} \rangle, & 2\mathbf{0} = \mathbf{0} \end{aligned}$$

Men da tilfredstiller tallet $\langle \mathbf{v}, \mathbf{0} \rangle$ likheten

$$2\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{v}, \mathbf{0} \rangle.$$

Flytte-bytte gir

$$\langle \mathbf{v}, \mathbf{0} \rangle = 0.$$

△

Lengden er definert til å være

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

Eksempel 11.7. Hvis du deler en ikke-null vektor med lengden sin, så får du en vektor av lengde 1; $\frac{\mathbf{v}}{\|\mathbf{v}\|}$ har lengde lik 1. Algebraisk verifikasjon:

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

△

To vektorer er *ortogonale* hvis

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

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

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

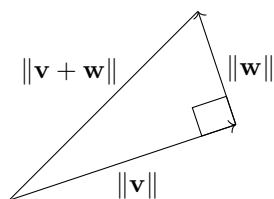
Teorem 11.8 (Pytagoras). *La V være et indreproduktrom. Dersom vektorene \mathbf{v} og \mathbf{w} er ortogonale, er*

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

Bevis. Vi regner litt på venstre side av ligningen.

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

□



Pytagoras' teorem

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

Teorem 11.9 (Cauchy-Schwarz). *La V være et 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

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

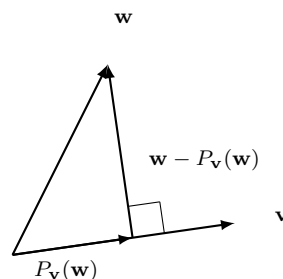
Spesielt er to vektorer *parallelle* hvis

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

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 11.3 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 for generelle indreproduktrom (bare bytt $\mathbf{v} \cdot \mathbf{w}$ med $\langle \mathbf{v}, \mathbf{w} \rangle$).

Teorem 11.10. *La V være et indreproduktrom. Den ortogonale projeksjonen på en vektor \mathbf{v} ,*

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

er en lineærtransformasjon.

Teorem 11.11. *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):

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

□

Indreprodukt mellom funksjoner

Denne seksjonen inneholder et eksempel på et nytt indreprodukt. Vi skal definere et indreprodukt mellom stykkevis kontinuerte 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 kontinuerte 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 utellbart 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 utellbart 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 utellbart 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 11.12. *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:

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

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 større eller lik null. Derfor blir arealet under grafen større eller lik null. Vi kan konkludere med at $\langle f, f \rangle \geq 0$. Når er $\langle f, f \rangle = 0$? Siden $f(x)^2$ alltid er positiv, må vi ha at $f(x) = 0$ for alle x , ellers får vi et ikke-null areal. Det er, med andre ord, kun null-funksjonen som tilfredstiller $\langle f, f \rangle = 0$. □

Merk. I praksis vil vi også tillate *stykkevis kontinuerte* funksjoner – kontinuerlig overalt, bortsett fra et endelig antall punkter – for vårt indreprodukt. Et teknisk problem gjør det litt vanskelig å formulere dette presist – men du trenger ikke å bekymre deg for denne typen detaljer. Fra nå av lar vi $\mathcal{C}_s([a, b])$ betegne rommet av alle stykkevis kontinuerte funksjoner. Den *spesielt interesserte* studenten kan merke seg at integralet av en funksjon som er null overalt, bortsett fra i ett punkt, er null. Det tekniske problemet er altså at ikke bare null-funksjonen gir $\langle f, f \rangle = 0$. Løsningen er å modifisere likhet i $\mathcal{C}_s([a, b])$: to funksjoner f og g er like hvis $f(x) = g(x)$ for alle x , bortsett fra et endelig antall punkter; de er så å si like. Faktisk er det mulig med et bedre integralbegrep enn det du lærte i matematikk 1 (Lebesgue integral) som tillater at to funksjoner f og g er like i $\mathcal{C}_s([a, b])$ hvis $f(x) = g(x)$ for alle x , bortsett fra et tellbart antall punkter. Men det er en historie for en annen dag.

Det er på tide med et eksempel.

Eksempel 11.13. Vi ser på indreproduktet når $a = 0$ og $b = 1$. Er x og x^2 ortogonale?

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

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

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

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

Lengdene er

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

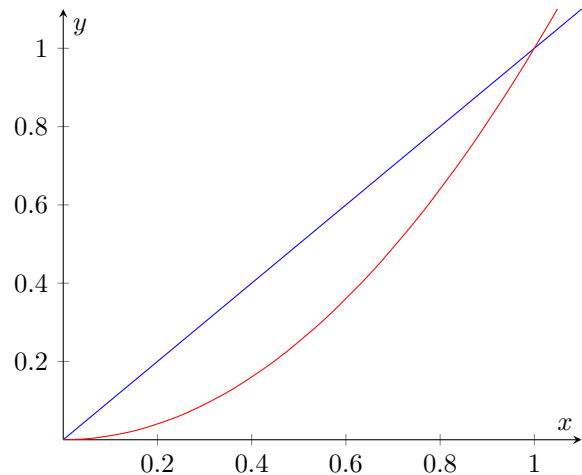
Hva er vinkelen mellom dem?

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

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 nærmere 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 nærmere hverandre, derfor er vinkelen mellom dem relativt liten. Hva er den ortogonale projeksjonen av x^2 på x ?

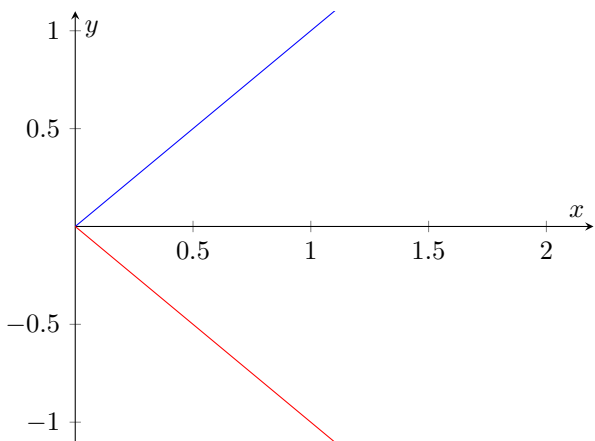
$$P_x(x^2) = \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} x = \frac{\frac{1}{4}}{\frac{1}{3}} x = \frac{3}{4} x.$$

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

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

△

Eksempel 11.14. 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

Men husk at vinkelen ikke er en vinkel mellom grafene. Den riktige inuisjonen her er at vinkelen burde være 180 grader. Hvorfor? Linjen utspent av $x - \text{Sp}\{x\}$ 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.

$$\begin{aligned} \cos \theta &= \frac{\langle x, -x \rangle}{\|x\| \| -x \|} \\ &= \frac{\int_0^1 x \cdot (-x) dx}{\sqrt{\int_0^1 x^2 dx} \sqrt{\int_0^1 (-x)^2 dx}} \\ &= \frac{-\int_0^1 x^2 dx}{\int_0^1 x^2 dx} \\ &= -1 \end{aligned}$$

Dette betyr at θ er 180 grader. \triangle

Eksempel 11.15. 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.

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

Husk den trigonometriske identiteten

$$\cos(2x) = 2 \cos^2 x - 1.$$

$$\begin{aligned} \|\cos x\|^2 &= \frac{1}{\pi - (-\pi)} \int_{-\pi}^{\pi} \cos x \cdot \cos x dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos^2 x dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 + \cos(2x)}{2} dx \\ &= \frac{1}{2\pi} \left[\frac{1}{2} x + \frac{1}{4} \sin(2x) \right]_{-\pi}^{\pi}, \quad \sin(\pm 2\pi) = 0 \\ &= \frac{1}{2\pi} \left(\frac{1}{2} \pi - \frac{1}{2} (-\pi) \right) \\ &= \frac{1}{2} \end{aligned}$$

På lignende vis regner vi ut

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

Vektorene 1 og $\cos x$ er ortogonale:

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

Tilsvarende regning gir

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

Bruk substitusjon for å regne ut det siste integralet:

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

\triangle

Du kan bruke lignende triks fra matematikk 1 for å bevise at $\cos(nx)$ og $\sin(mx)$, hvor $n, m = 2, 3, \dots$, også kan tas med i Eksempel 11.15.

Teorem 11.16. Vektorene 1 , $\cos(nx)$ og $\sin(mx)$, hvor $n, m = 1, 2, 3, \dots$, er parvis ortogonale i $\mathcal{C}_s([-\pi, \pi])$.

Mer om ortogonalitet

Definisjon. En *ortogonal mengde* er en mengde av ikke-null vektorer $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$, 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 standardbasen i \mathbb{R}^3 .

Teorem 11.17. *En ortogonal mengde er lineært uavhengig.*

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

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

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

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

Venstre side:

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

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 . \square

Eksempel 11.18. Standardbasen $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ for \mathbb{R}^n er en ortonormal mengde. \triangle

Eksempel 11.19. I forrige seksjon så vi at polynomene x og x^2 er ikke en ortogonal mengde i $\mathcal{C}_s([0, 1])$. \triangle

Eksempel 11.20. Teorem 11.16 kan omformuleres: Vektorene $1, \cos(nx)$ og $\sin(mx)$, hvor $n, m = 1, 2, 3, \dots$, er en ortogonal mengde i $\mathcal{C}_s([-\pi, \pi])$. \triangle

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

Teorem 11.17 sier at vi liksom godt kunne byttet ut 'basis' med 'spenner ut' i definisjonen av en ortogonal basis – en ortogonal mengde er lineært uavhengig, vi trenger kun å sjekke om den spenner ut.

Merk. Hvis indreproduktet er skalarproduktet i \mathbb{R}^n , så er det vanlig å sette opp $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ 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 basen $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$. Koordinatene, x_1, x_2, \dots, x_n , til \mathbf{v} i denne basen er gitt ved ligningen

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

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

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

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, så ser vi at

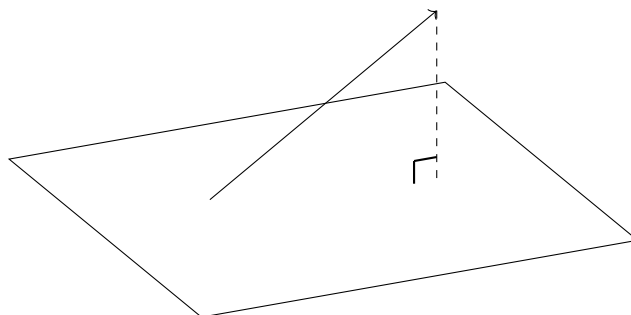
$$x_i \mathbf{u}_i = \frac{\langle \mathbf{u}_i, \mathbf{v} \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i = P_{\mathbf{u}_i}(\mathbf{v}).$$

Vi beviste altså:

Teorem 11.21. *Koordinatene til \mathbf{v} i en ortogonal basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ er*

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

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

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

er det punktet i V som har kortest avstand til \mathbf{v} :

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

Punktet $P_U(\mathbf{v})$ er det unike punktet som minimerer avstanden til U . Derfor er $P_U(\mathbf{v})$ uavhengig av hvilken ortogonal basis du velger for U . Dette definerer en lineærtransformasjon $P_U : V \rightarrow U$, 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 11.22. Hvordan velger jeg en ortogonal basis? Svaret er Gram-Schmidts metode, temaet i neste seksjon. Men vi klarer å regne på et enkelt eksempel:

Eksempel 11.23. 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

$$U^\perp = \{\text{vektorene i } V \text{ som er ortogonale på alle vektorer i } U\}.$$

En vektor \mathbf{v} er altså i U^\perp dersom $\langle \mathbf{v}, \mathbf{u} \rangle = 0$ for alle \mathbf{u} i U .

Merk. Det ortogonale komplementet er et underrom av V .

Det neste teoremet sier at det holder å sjekke om en vektor står normalt på en basis .

Teorem 11.24. La $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ 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 11.25. 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 11.26. La U være et underrom av et indreproduktrom V . Da gjelder

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

Gram-Schmidts metode

La $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ være en lineært uavhengig vektormengde. Vi skal lage oss en ortogonal basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ 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 . Nå kan vi fortsette slik, og definere rekursivt

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

Teorem 11.27. Mengden $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ er en ortogonal basis for rommet utspent av $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

Bevis. Vi bruker induksjon. Det er lett å se at \mathbf{u}_1 og \mathbf{u}_2 er ortogonale:

$$\begin{aligned} \langle \mathbf{u}_1, \mathbf{u}_2 \rangle &= \langle \mathbf{u}_1, (\mathbf{v}_2 - \frac{\langle \mathbf{u}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1) \rangle \\ &= \langle \mathbf{u}_1, \mathbf{v}_2 \rangle - \frac{\langle \mathbf{u}_1, \mathbf{v}_2 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \langle \mathbf{u}_1, \mathbf{u}_1 \rangle \\ &= \langle \mathbf{u}_1, \mathbf{v}_2 \rangle - \langle \mathbf{u}_1, \mathbf{v}_2 \rangle = 0 \end{aligned}$$

Siden \mathbf{u}_1 og \mathbf{u}_2 er ikke-trivielle lineærkombinasjoner av de lineært uavhengige vektorene \mathbf{v}_1 og \mathbf{v}_2 , er det

åpenbart at \mathbf{u}_1 og \mathbf{u}_2 spenner ut det samme rommet som \mathbf{v}_1 og \mathbf{v}_2 .

La nå

$$V_k = \text{Sp}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}.$$

Vi antar at $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ er en ortogonal basis for V_{k-1} . Vi må vise at \mathbf{u}_k står ortogonalt på V_{k-1} , og at $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ spenner ut V_k . Vi sjekker indreproduktet av \mathbf{u}_j med \mathbf{u}_k . Siden $\mathbf{u}_j^* \mathbf{u}_m = 0$ når $j \neq m$, får vi

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

Vi ser altså at \mathbf{u}_k står ortogonalt på alle \mathbf{u}_j , og siden $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$ er en ortogonal basis for V_{k-1} , står \mathbf{u}_k ortogonalt på V_{k-1} . Siden \mathbf{v}_k er lineært uavhengig av $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$, og \mathbf{u}_k er en lineærkombinasjon av V_k og $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{k-1}$, spenner $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ ut V_k . \square

La oss illustrere metoden med et par eksempler.

Eksempel 11.28. 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

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

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 . \triangle

Eksempel 11.29. 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 11.13. Ta $\mathbf{v}_2 = x^2$:

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

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

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

En ortogonal basis for U er

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

△

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

1. Bruk Gram-Schmidt til å finne en ortogonal basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ for U .

2. Den ortogonale projeksjonen er

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

3. Hvis $V = \mathbb{R}^n$, så er standardmatrisen

$$[P_U(\mathbf{e}_1) \quad P_U(\mathbf{e}_2) \quad \dots \quad P_U(\mathbf{e}_n)].$$

Eksempel 11.30. 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 11.28:

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

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

$$[P_U] = \frac{1}{6} \begin{bmatrix} 6 & 0 & 0 & 0 \\ 6 & 5 & 1 & 2 \\ 6 & 1 & 5 & -2 \\ 6 & 2 & 2 & 2 \end{bmatrix}.$$

△

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

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n,$$

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,

$$\mathbf{v}^* = [\overline{v_1} \quad \overline{v_2} \quad \dots \quad \overline{v_n}].$$

Vektoren \mathbf{v}^* kalles den *adjungerte* til \mathbf{v} . Nå kan vi bruke matrisemultiplikasjon til å definere et produkt

$$\mathbf{v}^* \mathbf{w} = \overline{v_1} w_1 + \overline{v_2} w_2 + \dots + \overline{v_n} w_n.$$

Merk deg at

$$\begin{aligned} \mathbf{v}^* \mathbf{v} &= \overline{v_1} v_1 + \overline{v_2} v_2 + \dots + \overline{v_n} v_n \\ &= |v_1|^2 + |v_2|^2 + \dots + |v_n|^2 \end{aligned}$$

alltid et ikke-negativt reelt tall; summen av lengden til komponentene. Et eksempel 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

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

Vi sier at V , sammen med ett valgt indreprodukt, er et *indreproduktrom*. \triangle

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

Du trenger kun å vite om ett komplekst eksempel, nemlig \mathbb{C}^n .

Eksempel 11.31. 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. \triangle

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

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

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 Cauchy-Schwarz, Teorem 11.9, impliserer at $\langle \mathbf{v}, \mathbf{w} \rangle$ er et komplekst tall med lengde mindre eller lik 1. I det reelle tilfellet befinner oss dermed i intervallet $[0, 1]$, men i det komplekse tilfellet er det flere tall som er inkludert; disken som er sentrert i 0 med radius 1.

Teorem 11.32. *Alt som gjelder for reelle indreproduktrom, med unntak av vinkelen, gjelder også for komplekse indreproduktrom.*

Eksempel 11.33. Vektorene

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

er ortogonale:

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

\triangle

Eksempel 11.34. 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

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

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

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

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

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

Mer generelt, la V og W være indreproduktrom, og T er 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 11.35. 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:

$$(A^*)^* = A \quad \triangle$$

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

Det siste resultatet i denne seksjonen er grunnlaget for minste kvadraters metode – som er en anvendelse i neste 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 $\text{Col } A^*$, er ortogonal med $\text{Null } A$. Faktisk gjelder følgende likheter:

Teorem 11.36. La A være en $m \times n$ -matrise. Da vet vi

$$\begin{aligned} (\text{Col } A)^\perp &= \text{Null } A^* \\ (\text{Null } A)^\perp &= \text{Col } A^*. \end{aligned}$$

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

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

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

Men dette betyr ved definisjon av multiplikasjonen matrise gang vektor at $A^* \mathbf{x} = \mathbf{0}$. Altså ligger \mathbf{x} i $\text{Null } A^*$.

En annen måte å vise dette er å observere at $\mathbf{x} \in (\text{Col } A)^\perp$ 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 \mathbf{x} ligger i $\text{Null } A^*$.

Omvendt la oss anta at \mathbf{x} er en vektor i $\text{Null } A^*$. 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 $(\text{Col } A)^\perp$. □

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

$$\begin{aligned} (\text{Col } A)^\perp &= \text{Null } A^T \\ (\text{Null } A)^\perp &= \text{Col } A^T. \end{aligned}$$

△