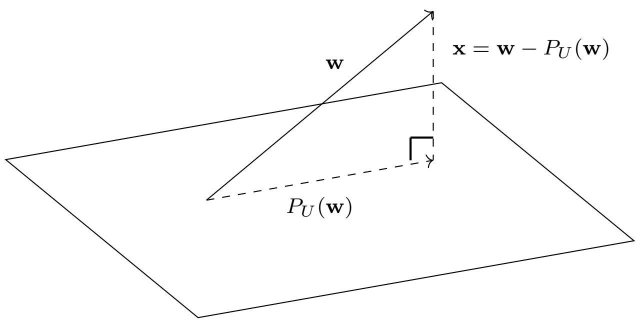
Kapittel 9

Projeksjon

La V være et vektorrom og la $\mathbf{w}$ være en vektor i V. Hvis vi har et underrom U av V, kan vi spørre: Hva er vektoren i U som *ligger nærmest* $\mathbf{w}$? Denne vektoren kaller vi *projeksjonen* av $\mathbf{w}$ ned i U, eller $P_U(\mathbf{w})$.

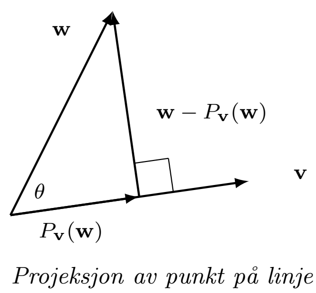
For eksempel kan U være et plan i $\mathbb{R}^3$, og $\mathbf{w}$ en vilkårlig vektor i $\mathbb{R}^3$. Følgende figur illustrerer dette.



Projisere en vektor ned på et plan.

Vi ser at vi kan finne det nærmeste punktet i planet, altså projeksjonen $P_U(\mathbf{w})$, ved å finne en vektor $\mathbf{x}$ som står normalt på U, og slik at $\mathbf{w} = P_U(\mathbf{w}) + \mathbf{x}$. Det er lengden av denne vektoren $\mathbf{x}$ som er den korteste avstanden fra $\mathbf{w}$ til et punkt i U.

Vi kan også tenke på projeksjon fra et punkt $\mathbf{w}$ i $\mathbb{R}^2$ til ei linje i $\mathbb{R}^2$. Vi kan tenke på linja som underrommet utspent av en vektor $\mathbf{v}$. Vi skal da altså finne en vektor $P_U(\mathbf{w}) = P_\mathbf{v}(\mathbf{w})$ som er parallell med $\mathbf{v}$ og slik at $\mathbf{v}$ og $\mathbf{w} - P_\mathbf{v}(\mathbf{w})$ er ortogonale (vinkelrette).



Projeksjon av punkt på linje

Dette kapitlet inneholder mye informasjon og mange formler. Her er hovedpunktene vi skal gjennom:

I. Vise hvordan vi projiserer en vektor $\mathbf{w}$ ned på en annen vektor $\mathbf{v}$, eller egentlig: projiserer $\mathbf{w}$ i underrommet (linja) utspent av $\mathbf{v}$.

II. Forklare hva det betyr at vektorer er ortogonale og innføre begrepet ortogonal basis for et vektorrom.

III. Vise hvordan vi projiserer en vektor $\mathbf{w}$ ned på et underrrom U, ved å bruke en ortogonal basis for U.

IV. Lære hvordan vi ved hjelp av Gram-Schmidt metoden finner en ortogonal basis.

V. Illustrere hvordan dette kan brukes i vektorrom som består av visse funksjoner, og der underrommene består av spesielt *fine funksjoner*, som polynomer eller trigonometriske funksjoner.

Intuisjonen og motivasjonen kommer igjen fra $\mathbb{R}^2$ og $\mathbb{R}^3$, men vi skal altså etterhvert jobbe med generelle vektorrom, blant annet for å kunne hanskes med punkt V.

La oss først minne oss på skalarprodukt i $\mathbb{R}^2$, og se hvordan vi kan bruke dette til å finne projeksjonen som vi nevnte i punkt I.

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

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

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

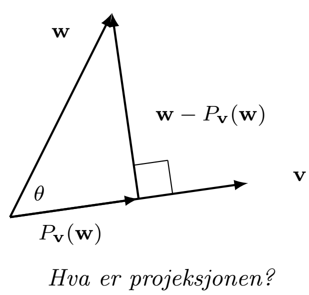
der $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2}$ er lengden til $\mathbf{v}$ og $\|\mathbf{w}\|$ er lengden til $\mathbf{w}$ og θ er vinkelen mellom $\mathbf{v}$ og $\mathbf{w}$, og

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

Ut ifra den andre formelen ser 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:

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

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 projeksjonen?

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

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

For å finne $P_{\mathbf{v}}(\mathbf{w})$ multipliserer vi $\|P_{\mathbf{v}}(\mathbf{w})\|$ med $\frac{\mathbf{v}}{\|\mathbf{v}\|}$, siden dette er en vektor med lengde 1 parallell med $\mathbf{v}$, og får

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

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. Den naturlige generaliseringen i $\mathbb{R}^n$ er

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

som også kan uttrykkes ved matriseproduktet

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

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 også i $\mathbb{R}^3$.

For $n > 3$ har ikke vinkelen θ lenger noen geometrisk tolkning, men vi kan jo definere et *tall* θ mellom 0 og π slik at

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

også i $\mathbb{R}^n$. For at dette skal gi mening må vi vite at tallet til høyre i ligningen har absoluttverdi mindre enn eller lik 1. Det kan bevises og kalles Cauchy-Schwarz-teoremet. Vi kommer tilbake til det senere i en mer generell sammenheng.

Da kan vi fortsatt si at to vektorer er *ortogonale* (som nå da betyr at dette tallet $\cos\theta = 0$) hvis

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

og *parallelle* (altså $\cos\theta = \pm 1$) hvis

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

Akkurat som i $\mathbb{R}^2$ kan vi definere den *ortogonale projeksjonen* av $\mathbf{w}$ på $\mathbf{v}$ som

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

Dette vil gi det *nærmeste* punktet til $\mathbf{w}$ på linja utstpent av $\mathbf{v}$. 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 $\mathbf{v}$. Heldigvis stemmer formelen overens med intuisjonen. La oss se hvorfor. Parallellitet betyr at $\mathbf{w}$ er et skalarmultiplum av $\mathbf{v}$; altså $\mathbf{w} = t\mathbf{v}$. 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 $\mathbf{v}$ er parallelle siden $P_{\mathbf{v}}(\mathbf{w}) = 2\mathbf{v}$, 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 $\mathbf{v}$:

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

Vi holder fast en vektor $\mathbf{v}$ i $\mathbb{R}^n$. Du kan tenke 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:

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

Vi skal etterhvert komme fram til hva projeksjoner er også i andre vektorrom enn $\mathbb{R}^n$. Vi har sett at projeksjon og lengde kan uttrykkes ved 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 v_i, den i-te komponenten til $\mathbf{v}$.

Et eksempel fra $\mathbb{R}^3$ er

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

Vi skriver $\mathbf{v}$ som en lineærkombinasjon

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

der vi altså har at $P_{\mathbf{e}_i}(\mathbf{v}) = v_i\mathbf{e}_i$, siden $\mathbf{e}_i\cdot\mathbf{e}_i = 1$. Det følger at

$$\mathbf{v} = P_{\mathbf{e}_1}(\mathbf{v}) + P_{\mathbf{e}_2}(\mathbf{v}) + \cdots + P_{\mathbf{e}_n}(\mathbf{v}).$$

Altså: $\mathbf{v}$ kan skrives som *summen av projeksjonene ned på hver vektor i standardbasisen.* Vi skal senere se at standardbasisen er et eksempel på en *ortogonal basis*, og at denne måten å skrive en vektor uttrykt som en sum av projeksjoner, vil fungere generelt, så lenge vi har en ortogonal basis.

Målet med neste seksjon er å generalisere skalarproduktet – som vi kommer til å kalle et indreprodukt – for å kunne snakke om avstander og projeksjoner i generelle abstrakte vektorrom.

Indreproduktrom

Vi ønsker å definere et produkt som tar inn to vektorer for å produsere en skalar. Dette produktet skal oppføre seg som prikkproduktet i $\mathbb{R}^n$. Generaliseringen kalles 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 $\mathbf{v}\cdot\mathbf{v} \geq 0$ lar oss definere lengden

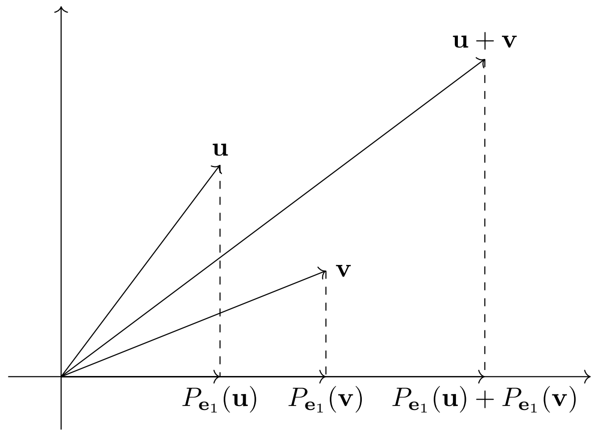
$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}},$$

siden kvadratroten av reelle tall 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.

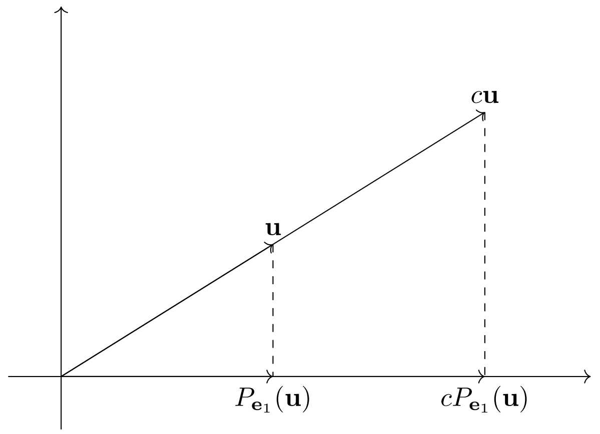
Den siste egenskapen vi skal tenke på er *linearitet*:

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

Skalarproduktet er også 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 à 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

Vi *definerer* nå abstrakte indreprodukt slik at disse egenskapene: symmetri, linearitet og positivitet holder.

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

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

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

Merk. Hvis du kombinerer linearitet og symmetri, 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 9.4. 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. △

Som for $\mathbb{R}^n$ kan vi nå definere lengden til en vektor $\mathbf{v}$ som kvadratrota av indreproduktet av $\mathbf{v}$ med seg selv, altså:

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

Merk at denne lengden ikke har en direkte geometrisk tolkning i V.

Eksempel 9.5. 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:

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

△

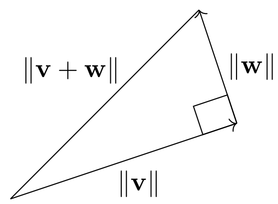
Basert på disse definisjonene kan vi allerede bevise Pytagoras' teorem.

Teorem 9.6 (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.

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

□



Pytagoras' teorem

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

Teorem 9.7 (Cauchy–Schwarz)**.** *La V være et reelt indreproduktrom. Alle vektorer $\mathbf{v}$ og $\mathbf{w}$ tilfredstiller*

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

Hvis ingen av vektorene er $\mathbf{0}$ så 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

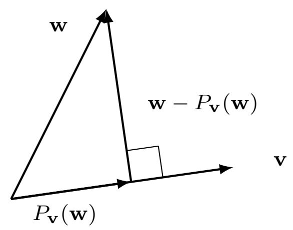
$$\langle \mathbf{v}, \mathbf{w} \rangle = \pm \|\mathbf{v}\| \|\mathbf{w}\|,$$

eller ekvivalent, $\mathbf{w} = t\mathbf{v}$ for et reelt tall t; de ligger på samme linje. Videre sier vi at to vektorer $\mathbf{u}, \mathbf{v}$ i V er *ortogonale* hvis $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Vi kan nå definere den *ortogonale projeksjonen* av $\mathbf{w}$ på $\mathbf{v}$ som

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

Den geometriske intuisjonen 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, og beviset følger av lineariteten til skalarproduktet. Samme bevis fungerer dermed for generelle indreproduktrom ved å bytte $\mathbf{v} \cdot \mathbf{w}$ med $\langle \mathbf{v}, \mathbf{w} \rangle$.

Teorem 9.8. *La V være et indreproduktrom. Den ortogonale projeksjonen på en vektor $\mathbf{v} \neq \mathbf{0}$,*

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

er en lineærtransformasjon.

Teorem 9.9. *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

Vi tenker altså på $P_{\mathbf{v}}$ som en funksjon (faktisk en lineær transformasjon) som projiserer alle vektorer i V ned på det én-dimensjonale underrommet U utspent av $\mathbf{v}$. Vi har stilltiende antatt $\mathbf{v} \neq \mathbf{0}$. Dermed er spesielt $\mathbf{v}$ en basis for underrommet $U = \mathrm{Sp}\{\mathbf{v}\}$. Dette har vi nå vist hvordan fungerer både i $\mathbb{R}^n$ og i abstrakte vektorrom.

Vi skal nå se at for å projisere ned på underrom U av dimensjon større enn én, vil vi trenge en spesiell type basis for U, som vi vil kalle en *ortogonal* basis.

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.*

Merk at vi alltid kan erstatte en vektor $\mathbf{v}$ som ikke nullvektoren med en vektor som peker i samme retning og har lengde 1 ved å gange Vv med tallet $\frac{1}{\|\mathbf{v}\|}$.

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 9.10. *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 9.11. 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.

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 ortogonal basis. Anvend indreproduktet med $\mathbf{u}_i$ for å få ut den i-te komponenten:

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

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

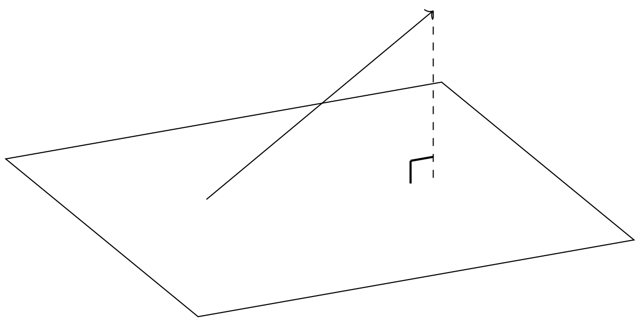
$$x_i\mathbf{u}_i = \frac{\langle \mathbf{u}_i, \mathbf{v}\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\mathbf{u}_i = P_{\mathbf{u}_i}(\mathbf{v}).$$

Vi beviste altså følgende teorem:

Teorem 9.12. *Koordinatene til* $\mathbf{v}$ *i en ortogonal basis* $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ *er*

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

Når vi projiserer vektorer $\mathbf{v}$ ned i et underrom U, der $\mathbf{v}$ ikke er i U, ønsker vi at resultatet $P_U(\mathbf{v})$ skal være vektoren i U med minst mulig avstand til $\mathbf{v}$. Det neste teoremet, som vi ikke skal bevise, forteller oss hvordan vi skal definere projeksjonen får å få til dette. Merk at vi må ha en *ortogonal basis* for underrommet U, for at teoremet skal holde.



Ortogonal projeksjon minimerer avstanden til et underrom

Teorem 9.13. *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* $\mathbf{v}$*, altså*

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

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 $P_U\colon V \to V$, *den ortogonale projeksjonen* ned på underrommet U. Bildet til P_U er altså U. Du kan regne den ut ved å velge en ortogonal basis for så å bruke formelen i Teorem 9.13. I neste seksjon skal vi se hvordan vi finner en ortogonal basis ved hjelp av Gram–Schmidts metode. La oss først se på et enkelt eksempel hvor vi allerde har en ortogonal basis:

Eksempel 9.14. Betrakt (x, y)–planet som et underrom U av $\mathbb{R}^3$. 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 for $U = \mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2\}$. Formelen gir

$$P_U\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) = P_{\mathbf{e}_1}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) + P_{\mathbf{e}_2}\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right).$$

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

$$P_U\left(\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right) = \begin{bmatrix}v_1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\v_2\\0\end{bmatrix} = \begin{bmatrix}v_1\\v_2\\0\end{bmatrix}.$$

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

$$[P_U] = \begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}.$$

△

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.

Teorem 9.15. *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 til* U *hvis og bare hvis* $\langle \mathbf{u}_i, \mathbf{v}\rangle = 0$ *for alle* i*.*

Eksempel 9.16. 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 9.17. *La U være et underrom av et indreproduktrom V. Da gjelder*

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

I et senere kapittel skal vi se på *minste kvadraters metode* som er en anvendelse av projeksjon.

For å beregne det ortogonale komplementet til et underrom U i $\mathbb{R}^n$ husker vi at hvert underrom er kolonnerommet til en matrise som har som kolonner vektorer som utspenner U. Og for kolonnerommet til en matrise observerer vi:

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

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

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

Teorem 9.18. *La A være en* reell *$m \times n$-matrise. Da vet vi*

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

Bevis. Det holder å bevise en av påstandene, siden den andre da følger ved å se på A^T istedet for A. Så la oss bevise $(\operatorname{Null} A)^{\perp} = \operatorname{Col} A^T$. Vi har allerede sett at $\operatorname{Col} A^T \subseteq (\operatorname{Null} A)^{\perp}$, og vi ser dermed lett at $\operatorname{Col} A^T$ er et underrom av $(\operatorname{Null} A)^{\perp}$.

Vi har at

$$\begin{aligned}\dim \operatorname{Col} A^T &= \dim \operatorname{Row} A = \operatorname{rank} A \\ &= n - \dim \operatorname{Null} A = \dim(\operatorname{Null} A)^{\perp}\end{aligned}$$

Dermed må vi ha at $\operatorname{Col} A^T = (\operatorname{Null} A)^{\perp}$.

Hvorfor: La $(\mathbf{x}_1, \ldots, \mathbf{x}_t)$ være en basis for $\operatorname{Col} A^T$. Siden disse er lineært uavhengige kan de utvides til en basis for $(\operatorname{Null} A)^{\perp}$. Men siden $\dim(\operatorname{Null} A)^{\perp} = t$, må dette allerede være en basis. Altså er

$$(\operatorname{Null} A)^{\perp} \subseteq \operatorname{Sp}\{\mathbf{x}_1, \ldots, \mathbf{x}_t\} = \operatorname{Col} A^T.$$

□

Finne en ortognal basis: Gram-Schmidts metode

I forrige avsnitt så vi at vi kan beskrive en projeksjon ned i et underrom U, hvis vi har en ortogonal basis for U. Vi skal nå se hvordan vi kan skaffe oss en ortogonal basis.

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 $\mathbf{u}_3$. Og vi har også at hver $\mathbf{u}_i$ kan skrives som en lineærkominasjon av $\mathbf{v}_1$, $\mathbf{v}_2$ og $\mathbf{v}_3$. Nå kan vi fortsette slik, og definere rekursivt

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

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 9.19. *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 trenger bare å vise at mengden $\mathbf{u}_1$, $\mathbf{u}_2$, ...,$\mathbf{u}_n$ er ortogonal. Vi har da sett tidligere i kapitlet at den automatisk er lineært uavhengig. Det er klart at $\mathbf{u}_1$ er en ortogonal basis for $\operatorname{Span}(\mathbf{v}_1)$. 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 $\langle \mathbf{u}_j, \mathbf{u}_k \rangle = 0$.

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

□

Eksempel 9.20. 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. △

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
$$[P_U] = \begin{bmatrix} P_U(\mathbf{e}_1) & P_U(\mathbf{e}_2) & \ldots & P_U(\mathbf{e}_n) \end{bmatrix},$$
og da kan man, for alle $\mathbf{x}$ i V, beregne den ortogonale projeksjonen $P_U(\mathbf{x}) = [P_U]\mathbf{x}$.

Eksempel 9.21. 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 9.20:
$$\mathbf{u}_1 = \begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}1\\0\\-2\\1\end{bmatrix} \text{ og } \mathbf{u}_3 = \begin{bmatrix}-1\\2\\0\\1\end{bmatrix}.$$
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

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

$\triangle$

Når U er et underrom av $\mathbb{R}^n$, har vi en ofte raskere metode for å finne standardmatrisen til projeksjonen ned på U når vi kjenner en ortonormal basis for U.

Teorem 9.22. *La U være et underrom av $\mathbb{R}^n$ med ortonormal basis $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$, hvor $m \leq n$, og la P_U være lineærtransformasjonen gitt ved projeksjonen på U. Vi lar $[P_U]$ betegne standardmatrisen til P_U, og lar $A_U = \begin{bmatrix} \mathbf{u}_1 & \dots & \mathbf{u}_m \end{bmatrix}$ være matrisen med vektorene $\mathbf{u}_1, \dots, \mathbf{u}_m$ som kolonner. Da har vi at*

$$[P_U] = A_U A_U^T.$$

Bevis. Det holder å vise at $P_U(\mathbf{x}) = A_U A_U^T \mathbf{x}$, for enhver $\mathbf{x} \in \mathbb{R}^n$; for hvis dette stemmer, vil

$$\begin{aligned} [P_U] &= \begin{bmatrix} P_U(\mathbf{e}_1) & \dots & P_U(\mathbf{e}_n) \end{bmatrix} \\ &= \begin{bmatrix} A_U A_U^T \mathbf{e}_1 & \dots & A_U A_U^T \mathbf{e}_n \end{bmatrix} = A_U A_U^T. \end{aligned}$$

Merk videre at enhver $\mathbf{x} \in \mathbb{R}^n$ kan skrives som $\mathbf{x} = (\mathbf{x} - P_U(\mathbf{x})) + P_U(\mathbf{x})$, hvor $P_U(\mathbf{x}) \in U$ og $\mathbf{x} - P_U(\mathbf{x}) \in U^\perp$. Hvis da $P_U(\mathbf{v}) = A_U A_U^T \mathbf{v}$ for alle $\mathbf{v}$ i U og $U^\perp$, vil

$$\begin{aligned} P_U(\mathbf{x}) &= P_U((\mathbf{x} - P_U(\mathbf{x})) + P_U(\mathbf{x})) \\ &= P_U(\mathbf{x} - P_U(\mathbf{x})) + P_U(P_U(\mathbf{x})) \\ &= A_U A_U^T(\mathbf{x} - P_U(\mathbf{x})) + A_U A_U^T P_U(\mathbf{x}) \\ &= A_U A_U^T \mathbf{x}. \end{aligned}$$

Derfor holder det å vise at $P_U(\mathbf{x}) = A_U A_U^T \mathbf{x}$, for alle $\mathbf{x}$ i U og $U^\perp$. La først $\mathbf{x} \in U$. Siden $\mathbf{u}_1, \dots, \mathbf{u}_m$ er en basis for U, kan vi skrive

$$\mathbf{x} = c_1 \mathbf{u}_1 + \dots + c_m \mathbf{u}_m,$$

for noen unike koeffisienter $c_1, \dots, c_m$. Derfor vil

$$A_U A_U^T \mathbf{x} = c_1 A_U A_U^T \mathbf{u}_1 + \dots + c_m A_U A_U^T \mathbf{u}_m.$$

Hvis $\mathbf{u}_i$, for $1 \leq i \leq m$, er en av basisvektorene, vil

$$A_U^T \mathbf{u}_i = \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix} \mathbf{u}_i = \begin{bmatrix} \mathbf{u}_1^T \mathbf{u}_i \\ \vdots \\ \mathbf{u}_m^T \mathbf{u}_i \end{bmatrix} = \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{u}_i \rangle \\ \vdots \\ \langle \mathbf{u}_m \mathbf{u}_i \rangle \end{bmatrix} = \mathbf{e}_i,$$

siden $\{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ er en ortonormal basis; dermed vil

$$A_U A_U^T \mathbf{u}_i = A_U \mathbf{e}_i = \mathbf{u}_i,$$

og vi kan konkludere med at

$$\begin{aligned} A_U A_U^T \mathbf{x} &= c_1 A_U A_U^T \mathbf{u}_1 + \dots + c_m A_U A_U^T \mathbf{u}_m \\ &= c_1 \mathbf{u}_1 + \dots + c_m \mathbf{u}_m = \mathbf{x}. \end{aligned}$$

Vi vet fra før at $P_U(\mathbf{x}) = \mathbf{x}$ dersom $\mathbf{x} \in U$, så $P_U(\mathbf{x}) = A_U A_U^T \mathbf{x}$, for alle $\mathbf{x} \in U$. La så $\mathbf{x} \in U^\perp$. Da vil

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

for alle $1 \leq i \leq m$, og vi får at

$$\begin{aligned} A_U A_U^T \mathbf{x} &= A_U \begin{bmatrix} \mathbf{u}_1^T \mathbf{x} \\ \vdots \\ \mathbf{u}_m^T \mathbf{x} \end{bmatrix} \\ &= A_U \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{u}_m, \mathbf{x} \rangle \end{bmatrix} = A_U \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0}. \end{aligned}$$

Vi vet fra før at $P_U(\mathbf{x}) = \mathbf{0}$, når $\mathbf{x} \in U^\perp$, slik at

$$P_U(\mathbf{x}) = \mathbf{0} = A_U A_U^T \mathbf{x},$$

for alle $\mathbf{x} \in U^\perp$. Med dette og hva vi har merket oss i begynnelsen, kan vi da konkludere med at $[P_U] = A_U A_U^T$. $\square$

Eksempel 9.23. Vi ønsker igjen å finne standardmatrisen til P_U, for U som i Eksempel 9.21, men nå ved å bruke Teorem 9.22. En ortogonal basis ble oppgitt som

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

men vi trenger en ortonormal basis for å kunne bruke Teorem 9.22. Vektorene i basisen over er allerede ortogonale, så for å få en ortonormal basis, deler vi hver vektor i basisen med dens lengde. En ortonormale basis er da gitt ved

$$\mathbf{u}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} \text{ og } \mathbf{u}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix},$$

Matrisen A_U er gitt som

$$A_U = \begin{bmatrix} \mathbf{u}_1 \dots \mathbf{u}_m \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \\ 1 & -2 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

og dermed er

$$A_U^T = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 0 & -2 & 1 \\ -1 & 2 & 0 & 1 \end{bmatrix}.$$

Vi beregner

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

$\triangle$

Indreprodukt mellom funksjoner

Denne seksjonen inneholder et eksempel på et nytt indreprodukt. Vi skal definere et indreprodukt av kontinuerlige funksjoner. Med 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 – altså 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 kravene for et indreprodukt holder.

Teorem 9.24. *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

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

Positivitet:

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

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

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

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

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

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

Lengdene er

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

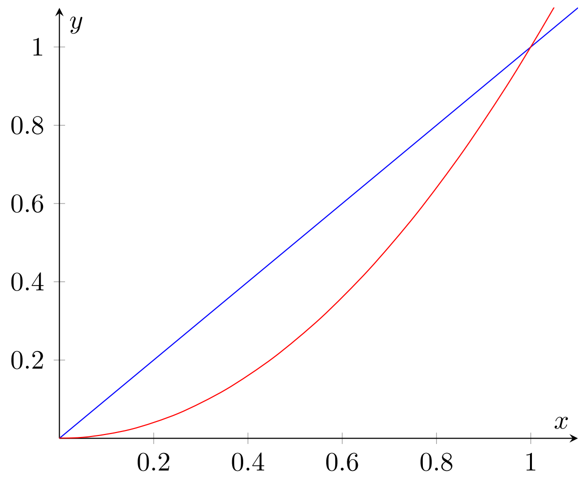
Hva er vinkelen mellom dem?

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

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:

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

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?

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

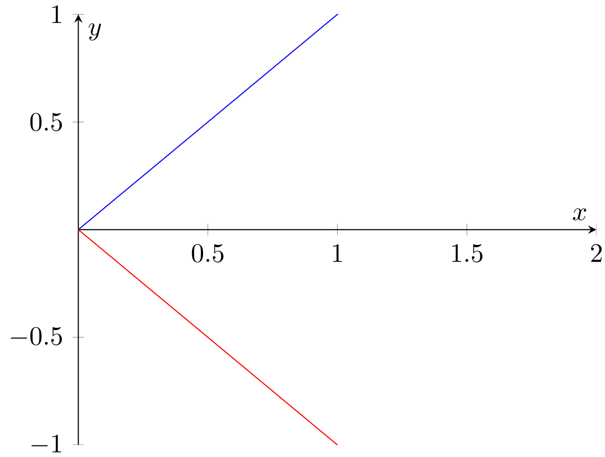
Vi sjekker at $P_x(x^2)$ og $x^2 - P_x(x^2)$ faktisk er ortogonale:

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

△

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 9.26. 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 forståelsen 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.

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

Dette betyr at θ er 180 grader. △

Eksempel 9.27. 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 9.25. 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:

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

△

Eksempel 9.28. 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 $\cos(2x) = 2\cos^2 x - 1$. Da blir

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

På lignende vis regner vi ut

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

Vektorene 1 og $\cos x$ er ortogonale:

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

Tilsvarende regning gir

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

Bruk substitusjon for å regne ut det siste integralet:

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

△

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 9.28.

Teorem 9.29. *Vektorene* 1, $\cos(nx)$ *og* $\sin(mx)$, *hvor* $n, m = 1, 2, 3, \ldots$, *er parvis ortogonale i* $\mathcal{C}_s([-\pi, \pi])$.