2 - 9 - KOMPLEKSE INDREPRODUKTROM

For litt siden intoduserte vi transponeringsoperasjonen, som er å bytte om på rader og kolonner i en matrise. Dersom tallene i matrisen er komplekse, viser det seg at det er lurt å komplekskonjugere alt i tillegg til å transponere. La

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

være en $m \times n$-matrise. Den **adjungerte** av A er $n \times m$-matrisen

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

$\boxed{1}$ Finn den adjungerte til

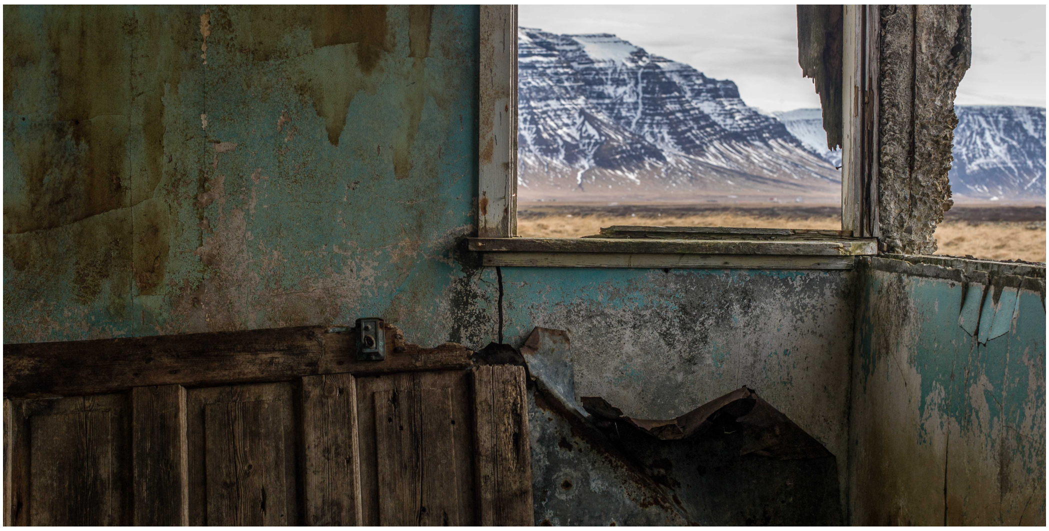
$$\begin{pmatrix} 1+i & 1-i \\ 2 & -i \\ 2-i & 1+i \end{pmatrix}$$

Merk at nå har vi en ny notasjon for $\overline{z}$, nemlig z^*. Du lurer sikkert på hvorfor vi komplekskonjugerer i tillegg til å transponere. Det er fordi de kartesiske koordinatene til z^*w har en mer interessant geometrisk tolking enn de kartesiske koordinatene til zw.

$\boxed{2}$ La z og w være komplekse tall, med $|w| = 1$. Tegn dem inn i det komplekse planet, skriv det komplekse tallet

$$w^*z$$

på kartesisk form, og se veldig nøye på resultatet. Ser du noen projeksjoner?
(Hint: Dette er en slags kompleks mashup av oppgave 3 og 6 i økt 2 - 5.)



Siden $\overline{z}w$ inneholder informasjon om projeksjonen av w på z, og lengden av et komplekst tall er

$$|z| = \sqrt{z^* z},$$

definerer vi **det komplekse skalarproduktet** mellom komplekse kolonnevektorer $\mathbf{z}$ og $\mathbf{w}$ som

$$\mathbf{z}^*\mathbf{w} = \begin{pmatrix} \overline{z}_1 & \overline{z}_2 & \cdots & \overline{z}_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \overline{z}_1 w_1 + \overline{z}_2 w_2 + \cdots + \overline{z}_n w_n$$

Vi sier at vektorer i $\mathbb{C}^n$ er ortogonale dersom det komplekse skalarproduktet mellom dem er null, og lengden til en kompleks vektor $\mathbf{z}$ defineres som

$$\|\mathbf{z}\| = \sqrt{\mathbf{z}^*\mathbf{z}}$$

Det komplekse skalarproduktet er noe mer kompliserte å forholde seg til enn det reelle. Den første forskjellen er at Pytagoras' setning kun har enveis implikasjon:

hvis $\mathbf{z}$ og $\mathbf{w}$ er ortogonale, er $\|\mathbf{z}+\mathbf{w}\|^2 = \|\mathbf{z}\|^2 + \|\mathbf{w}\|^2$

Beregningen

$$\|\mathbf{z}+\mathbf{w}\|^2 = (\mathbf{z}+\mathbf{w}) \cdot (\mathbf{z}+\mathbf{w}) = \|\mathbf{z}\|^2 + \mathbf{z}^*\mathbf{w} + \mathbf{w}^*\mathbf{z} + \|\mathbf{w}\|^2.$$

er fremdeles gyldig, men siden vi ikke uten videre kan bytte plass på faktorene i det komplekse skalarproduktet, kan vi ikke slå sammen de to midterste leddene, og disse kan kansellere på andre måter enn at $\mathbf{z}$ og $\mathbf{w}$ er ortogonale.

$\boxed{3}$ Finn et eksempel på $\mathbf{z}$ og $\mathbf{w}$ slik at

$$\mathbf{z}^*\mathbf{w} + \mathbf{w}^*\mathbf{z} = 0$$

uten at $\mathbf{z}^*\mathbf{w} = 0$.

Dersom $\mathbf{z}^*\mathbf{w} = 0$ er også $\mathbf{w}^*\mathbf{z} = 0$,

$\boxed{4}$ vis dette

og dette betyr at dersom $\mathbf{z}^*\mathbf{w} = 0$, får vi

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

$\boxed{5}$ Denne kommer på eksamen.

Cauchy-Schwarz og trekantulikheten gjelder også i det komplekse tilfellet, men jeg skal ikke plage deg med bevisene. Det jeg derimot må plage deg med, er

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{int} \qquad \text{og} \qquad c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}\,dt$$

Du lurte forhåpentligvis på hvor minusen i eksponensialfunksjonen kom fra. Forhåpentligvis skjønner du også nå at formelen for fourierkoeffisientene c_n er et indreprodukt, men den er altså et **komplekst indreprodukt**. Her er aksiomene. Et komplekst indreproduktet er konjugert symmetrisk:

$$(\mathbf{z}, \mathbf{w}) = \overline{(\mathbf{w}, \mathbf{z})}$$

positivt definitt:

$$(\mathbf{z}, \mathbf{z}) > 0 \quad \text{dersom} \quad \mathbf{z} \neq \mathbf{0},$$

og lineært i enten første eller andre faktor, altså enten

$$(a\mathbf{z} + b\mathbf{v}, \mathbf{w}) = a(\mathbf{z}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w}) \qquad \text{eller} \qquad (\mathbf{z}, a\mathbf{v} + b\mathbf{w}) = a(\mathbf{z}, \mathbf{v}) + b(\mathbf{z}, \mathbf{w})$$

$\boxed{6}$ Sjekk at det komplekse skalarproduktet er et indreprodukt, lineært i andre faktor.

$\boxed{7}$ Sjekk at

$$(x, y) = \int_{-\pi}^{\pi} x(t)\overline{y(t)}\,dt$$

er det indreprodukt, lineært i første faktor.

$\boxed{8}$ Sjekk at dersom et komplekst indreprodukt er lineært i den ene faktoren, er det **antilineært** i den andre. Altså at dersom for eksempel

$$(a\mathbf{z} + b\mathbf{v}, \mathbf{w}) = a(\mathbf{z}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w}) \qquad \text{er} \qquad (\mathbf{z}, a\mathbf{v} + b\mathbf{w}) = \overline{a}(\mathbf{z}, \mathbf{v}) + \overline{b}(\mathbf{z}, \mathbf{w})$$

Tidligere i ditt liv har du "utledet" koeffisentene

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} x(t)e^{-int}\,dt$$

ved å observere at de komplekse eksponensialfunksjonene er ortogonale, og så bruke essensielt det samme resonnementet som du har gjort så fryktelig mange ganger dette semesteret.

Ved første øyekast virker det kronglete med ubestemmeligheten angående hvorvidt indreproduktet skal være lineært i første eller andre faktor. Men det finnes mange indreprodukter, og i forskjellige fagfelt opererer man med forskjellige konvensjoner, så det er like greit å venne seg til begge varianter. Komplekse vektorer kan være ortonormale på samme måte som reelle vektorer, men en matrise med ortonormale kolonner kalles **unitær**, ikke ortogonal.

$\boxed{10}$ Vis at for en unitær matrise Q er

$$Q^*Q = I = QQ^*$$

og sjekk at denne matrisen er unitær:

$$\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}$$

Det er et viktig tilfelle der egenvektorer automatisk er ortogonale. De tre neste oppgavene leder frem til dette. En kompleks matrise er **hermittisk** dersom $A = A^*$. [1]

$\boxed{11}$ Er matrisen over hermittisk?

Uttrykket

$$\mathbf{x}^*A\mathbf{x}$$

kalles en **kvadratisk form**. Disse ser ikke så spennende ut, men dukker stadig opp i anvendelser, for eksempel i statistikk.

$\boxed{12}$ Vis at dersom A er hermittisk, er $\mathbf{x}^*A\mathbf{x}$ reell.

$\boxed{13}$ Vis at en hermittisk matrise har reelle egenverdier.
(Hint: La $\mathbf{v}$ være en normalisert egenvektor og se på $\mathbf{v}^*A\mathbf{v}$.)

$\boxed{14}$ Vis at egenvektorene til to distinkte egenverdier er ortogonale for hermittiske matriser.

Dersom T er en abstrakt lineæroperator og $(\cdot, \cdot)$ er et komplekst indreprodukt, defineres den **adjungerte operatoren** T^* som noe som tilfredsstiller

$$(x, Ty) = (T^*x, y),$$

og en hermittisk operator som noe som tilfredsstiller

$$(x, Ty) = (Tx, y).$$

Dette er kvantefysikk. Den kvadratiske formen (x, Tx) kalles **forventningen** til T i tilstanden x.

$\boxed{15}$ Jeg vet ikke hvordan du gjorde oppgave 12-14 over, men om du skrev ut alt på komponentform, kan du prøve å ta tar utgangspunkt i abstrakte indreprodukter og hermittiske lineæroperatorer istedet. Det blir mye enklere.
(Hint: Ta utgangspunkt i likningen

$$Tx = \lambda x$$

og ta indreproduktet med x og bruk at $T = T^*$ samt det andre aksiomet for indreprodukt.)

[1] Men man kunne like gjerne sagt “symmetrisk”. Det er ingen som noensinne transponerer en kompleks matrise uten å komplekskonjugere innmaten.

Parsevals sats gjelder også i komplekse indreproduktrom, og beviset er så og si identisk.

$\boxed{16}$ Vis at dersom $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ er en ortonormal vektormengde og

$$\mathbf{z} = \sum_{k=1}^{n} c_k \mathbf{v}_k \qquad \text{er} \qquad (\mathbf{z}, \mathbf{z}) = \sum_{k=1}^{n} |c_k|^2 = \sum_{k=1}^{n} c_k \overline{c_k}.$$

Faktisk kan vi klare enda bedre:

$\boxed{17}$ Vis at dersom $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ er en ortonormal vektormengde og

$$\mathbf{z} = \sum_{k=1}^{n} c_k \mathbf{v}_k \qquad \text{og} \qquad \mathbf{w} = \sum_{k=1}^{n} d_k \mathbf{v}_k \qquad \text{er} \qquad (\mathbf{z}, \mathbf{w}) = \sum_{k=1}^{n} c_k \overline{d_k}.$$

Parsevals sats gjelder også for fourierrekker, men akkurat som for utledningen av formelen for koeffisientene, må vi nøye oss med heuristisk utledning. Den er basert på samme triks som i oppgave 10, men vi må gjøre noen bytter av uendelig sum og integraltegn som vi strengt tatt ikke vet om går bra, og derfor skal vi nøye oss med selve resultatet. Dersom et signal kan skrives

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n t/T} \qquad \text{med} \qquad c_n = \frac{1}{T}\int_{-T/2}^{T/2} x(t) e^{-int}\, dt$$

er

$$\frac{1}{T}\int_{-T/2}^{T/2} x^2(t)\, dt = \sum_{n=-\infty}^{\infty} |c_n|^2$$

Vi trenger faktisk ikke noe mer enn at x er riemannintegrerbar for at dette skal være sant.

$\boxed{18}$ Jeg kunne bedt deg skrive ut detaljene i resonnementet, men det blir på en måte litt teit, for det hadde blitt å gjøre oppgave 10 om igjen, bare med ∞ istedet for n oppå summen.

UKENS NØTTER

Ekte mannfolk bruker kvaternioner når de skal holde styr på rotasjoner:
`https://en.wikipedia.org/wiki/Quaternions_and_spatial_rotation`
og dette er standardmåten å gjøre det på i moderne robotstyringsalgoritmer. Dersom du ønsker å rotere $x \in \mathbb{R}^3$ om y der $\|y\| = 1$, kan du lage

$$x = x_1 i + x_2 j + x_3 k$$

og

$$y = \cos\left(\frac{\theta}{2}\right) + (y_1 i + y_2 j + y_3 k)\sin\left(\frac{\theta}{2}\right)$$

og beregne yxy^{-1}; imaginærdelene til denne gir deg den roterte vektoren.

$\boxed{1}$ Vis dette.
(Brett opp ermene og regn i vei. Tok meg om lag en uke, men jeg er heller ikke spesielt smart.)

Og nå til noe helt annet. La oss anta du har n realiseringer av en p-dimensjonal tilfeldig variabel. Det vanlig å sette disse opp i en $n \times p$-matrise:

$$X = \left(\begin{array}{c|c|c|c} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_p \end{array} \right)$$

Tenk at du måler lengde, bredde og høyde på alle plankene i en plankestabel, og så **sentrerer** du ved å ta gjennomsnitt i hver variabel og trekke denne fra hver enkelt måling. Til slutt setter du alt opp i matrisen X på en slik måte at rad k er de sentrerte målene til planke k.

$\boxed{2}$ Vis at elementene i $\frac{1}{n-1}X^TX$ er den empiriske kovariansen mellom kolonnene i X.

Vi kaller $\frac{1}{n-1}X^TX$ den **empiriske kovariansmatrisen**, og den er en estimator for den antatte bakenforliggende kovariansmatrisen. Vi sier at en matrise A er **positivt definitt** dersom

$$\mathbf{x}^TA\mathbf{x} > 0$$

for alle $\mathbf{x}$, og **positivt semidefinitt** dersom

$$\mathbf{x}^TA\mathbf{x} \geq 0$$

for alle $\mathbf{x}$.

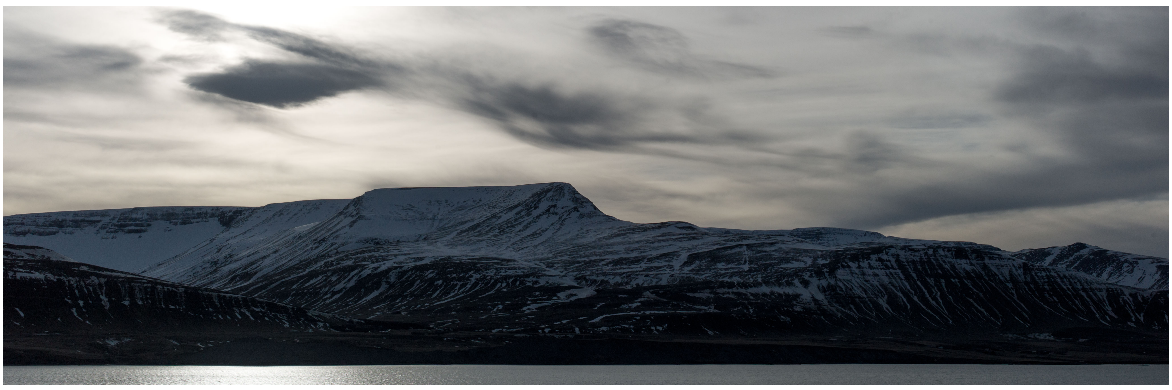
$\boxed{3}$ Vis at den empiriske kovariansmatrisen er positivt semidefinitt.

I flervariabel statistikk finnes det noe som kalles **prinsipalkomponentanalyse**.[2] Dette er en teknikk for å lete opp akser i datamengden der samvariasjonen er sterk. Teknikken baserer seg på å projisere X inn på en retning $\mathbf{v}$ slik at den empiriske variansen til de resulterende punktene blir maksimert. Dette får man til ved å lete opp egenvektorene til de største egenverdiene til kovariansmatrisen.

$\boxed{4}$ I 2-6 er det flere tovariable datamengder. Plott kolonnene mot hverandre i python, regn ut egenvektorene, og sjekk samvariasjon.

$\boxed{5}$ Er det samvariasjon mellom høyde og bredde i kaibordene fra sommeren 2023?
`https://folk.ntnu.no/mortano/python/kai/kai.csv`

$\boxed{6}$ Forklar hvorfor egenvektorene til X^TX forteller om samvariasjonen mellom kolonnene i X. (Hint: Finn ut hva en "Rayleigh quotient" er.[3])



[2] https://en.wikipedia.org/wiki/Principal_component_analysis
[3] https://en.wikipedia.org/wiki/Rayleigh_quotient