4 - 5 - MER LINEÆRALGEBRA

La oss begynne med litt geometrisk fikling.

$\boxed{1}$ Skriv en rutine som tar inn $A \in \mathbb{R}^{2\times 2}$ og plotter $A\mathbf{x}$ for alle $\mathbf{x}$ på enhetssirkelen.

Fikk du til dette, vil du oppdage at koden din alltid produserer ellipser, ihvertfall så lenge A har linært uavhengige kolonner. La oss finne ut mer om disse ellipsoidene. Trikset er å studere A^TA og AA^T. Vi lar $n > p$, slik at A er høyt og tynn og A^T er lang og flat. Dette er ikke noen essensiell restriksjon, men vi gjør den fordi det hjelper litt på visualiseringen.

$\boxed{2}$ Både A^TA og AA^T er ortogonalt diagonaliserbare. Hvorfor?

Siden A^TA er ortogonalt diagonaliserbar, finnes det et ortonormalt sett med egenvektorer. La oss sette disse opp i en ortogonal matrise V. Det samme gjelder for AA^T; la oss sette disse egenvektoren opp i en ortonormal matrise U.

$\boxed{3}$ Hva er dimensjonene til A^TA, AA^T, V og U?
Hvor mange av egenverdiene til AA^T kan være forskjellige fra null?

Her er enda en observasjon.

$\boxed{4}$ Vis at dersom $\mathbf{v}$ er en egenvektor til A^TA med egenverdi λ, er $A\mathbf{v}$ en egenvektor til AA^T. Hva er egenverdien? Kan du si noe om fortegnet?

Hvis du skjønte forrige oppgave, skjønner du forhåpentligvis at A^TA og AA^T har de samme egenverdiene så lenge du ser bort fra multiplisiteten til de egenverdiene som er null. La oss nå anta at $\mathbf{v}$ er en egenvektor til A med lengde 1.

$\boxed{5}$ Hva er lengden til $A\mathbf{v}$?

Lengden til $A\mathbf{v}$ kalles **singulærverdi**, og vi bruker bokstaven σ_k. Kolonnene i U og V kalles henholdsvis de **venstre og høyre singulærvektorene**. Hvis du tenker nøye på alt du har gjort til nå, vil du se at det går an å skrive

$$A = U\Sigma V^T$$

der σ-ene sitter i den rektangulære diagonalmatrisen Σ. Dette kalles **SVD-faktoriseringen**[1] til A. Denne finnes alltid, og dersom du sorterer singulærverdiene i synkende rekkefølge (σ_1 er størst) og velger alle positive, er Σ entydig. De unitære matrisene U og V vil ikke nødvendigvis være entydige, men de er det i noen tilfeller.

$\boxed{6}$ Hvordan ser Σ ut? Finn svd-faktoriseringen til

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Singulærverdidekomposisjonen er en av de viktigste matrisefaktoriseringene. Har du den har du mye, og pythonkommandoen heter numpy.linalg.svd. Nå skal vi se litt på hva svd kan brukes til. Det første er kompresjon. Vi kan skrive A slik:

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^* + \sigma_3 \mathbf{u}_3 \mathbf{v}_3^* + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^*$$

der r er antallet singulærverdier som ikke er null. Vi ser nå at dersom det er bare er to av singulærverdiene som er over for eksempel maskinpresisjon, er all informasjon i matrisen inneholdt i uttrykket

$$A = \mathbf{u}_1 \sigma_1 \mathbf{v}_1^* + \mathbf{u}_2 \sigma_2 \mathbf{v}_2^*.$$

$\boxed{7}$ La oss lage et bilde av en katt. Lag deg en matrise A i python med tilfeldige tall mellom -1 og 1. Dette er bakgrunnen. Sett så ett element til 10 eller 100 eller noe sånt. Dette er katten. Kjør svd på A og skriv ut

$$A - \mathbf{u}_1 \sigma_1 \mathbf{v}_1^* \qquad \text{og} \qquad A - \left(\mathbf{u}_1 \sigma_1 \mathbf{v}_1^* + \mathbf{u}_2 \sigma_2 \mathbf{v}_2^*\right).$$

[1] https://en.wikipedia.org/wiki/Singular_value_decomposition

Å jobbe med lineæralgebra til maskinpresisjon på laptopen din kan være ganske forskjellig fra å jobbe med penn og papir og platonsk kjærlighet. Nå skal vi se på dette. Du husker forhåpentligvis indreproduktet $(\cdot,\cdot)$. Vi sier at lengden

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x},\mathbf{x})}$$

er **indusert** av indreproduktet $(\cdot,\cdot)$. Men det finnes også lengder som ikke er indusert av et indreprodukt. Derfor har vi egne aksiomer for lengde:

- $\|\mathbf{x}\| > 0$ dersom $\mathbf{x} \neq \mathbf{0}$
- $\|a\mathbf{x}\| = |a|\|\mathbf{x}\|$
- $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

På fagspråket sier man **norm** istedet for lengde. Det finnes mange lengder som ikke er indusert av indreprodukt

$\boxed{8}$ Vis at $\|\mathbf{x}\|_2 = \sqrt{(\mathbf{x},\mathbf{x})}$ er en lengde.

$\boxed{9}$ Vis at $\|\mathbf{x}\|_\infty = \max_k |x_k|$ og $\|\mathbf{x}\|_1 = \sum_k |x_k|$ er lengder når $\mathbf{x} \in \mathbb{C}^n$. Forklar subskriptene.

$\boxed{10}$ Alle lengdene over er spesialtilfeller av p-normen

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_k |x_k|^p}.$$

Skisser enhetssirkelen i $\mathbb{R}^2$ for forskjellige valg av p.
Fun fact: Fontenen på Sergelplassen i Stockholm har form som enhetssirkelen for $p = 4$.

Du kan også slenge inn en vektmatrise, og skrive $\|\mathbf{x}\| = \|A\mathbf{x}\|_p$.

$\boxed{11}$ Finn en A slik at enhetssirkelen blir en ellipse.

Dette leder oss inn på **matrisenormen indusert av** $\|\cdot\|$:

$$\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

$\boxed{12}$ Vis at dette er en lengde.

Alt etter hva slags norm du putter inn i definisjonen over, får du forskjellige matrisenormer. Heldigvis trenger vi ikke stappe inn masse vektorer i A og finne ut hvilken som gjør $A\mathbf{x}$ lengst, for noen har gjort det for oss. For eksempel er

$$\|A\|_1 = \max_{1\leq k\leq n} \|\mathbf{a}_k\| \qquad (\mathbf{a}_k \text{ er kolonne } k \text{ i } A).$$

For ∞-normen gjelder det samme, men med rad k istedet for kolonne k. Det finnes også matriselengder som ikke er indusert av indreprodukt. Det viktigste eksemplet er kanskje **frobeniuslengden**

$$\|A\|_F = \sqrt{\left(\sum_{k=1}^{n}\sum_{l=1}^{p} a_{kl}^2\right)}.$$

Denne kan også skrives

$$\|A\|_F = \mathrm{tr}(A^TA) = \mathrm{tr}(AA^T).$$

der tr er noe som kalles **sporet**. Sporet til en matrise er summen av diagonalelementene.

Den viktigste matrisenormen er normen indusert av det vanlige skalarproduktet, altså 2-normen. Velger vi denne, blir $\|A\| = \sigma_1$, altså den største singulærverdien.

$\boxed{13}$ Vis dette.

Det er en fun fact of life som ikke er helt triviell å bevise, nemlig at antall lineært uavhengige rader og kolonner i en matrise alltid er det samme.

$\boxed{14}$ Dobbeltsjekk dette for matrisen

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

Men nå må vi huske på at lineær avhengighet er en platonsk abstraksjon. Du finner aldri tre dønn lineært avhengige vektorer i en anvendelse der det er målefeil og unøyaktigheter og fuglene vet hva. Det som i praksis kan skje, er derimot at vektorer er *nesten lineært avhengige*, og dette skaper all slags problemer i praksis.

Så hva vil det si at vektorer er nesten lineært avhengige? Dette kan skje på to måter.

$\boxed{15}$ Gå tilbake til TMA4101 og repeter definisjonen av lineær uavhengighet og finn ut hvordan vektorer kan være nesten lineært uavhengige.

Tommelfingerregel er at du får problemer når kolonner peker i nesten samme retning eller har veldig varierende lengde. Alt i alt definerer vi derfor noe som kalles kondisjonstall:

$$\kappa(A) = \frac{\sigma_1}{\sigma_r}$$

der σ_1 og σ_r er den største og minste singulærverdien.

$\boxed{16}$ Vis at dersom A er inverterbar, er

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2.$$

For å få en følelse for kondisjonstallet, er det best å fikle litt med noen matriser. I python heter det np.linalg.cond. Problemet med store kondisjonstall, er at *små endringer i matrisen gir store endringer i annet, for eksempel egenverdier.* Dette gjør at modeller kan bli ubrukelige i praksis, siden små målefeil kan gi store svingninger i hva slags prediksjoner modellen spyr ut.

$\boxed{17}$ Regn ut kondisjonstallet til

$$\begin{pmatrix} 1 & a \\ b & 1 \end{pmatrix}$$

for forskjellige verdier for a og b. For en illustrasjon av hvordan ting går skeis, kan du prøve å regne ut egenverdier og egenvektorer når $a = 1000$ og $b = 0$ eller $b = .0001$.

Det finnes et viktig matrisemåletall, nemlig **spektralradien**

$$\rho(A) = \max_k |\lambda_k|.$$

$\boxed{18}$ Vis at dette ikke er en lengde.