33 - SYMMETRISKE MATRISER

I forrige økt avsluttet vi med å se at det er av interesse å vite foran og bak på den kvadratiske formen

$$\mathbf{v}^T A \mathbf{v}$$

der A er en symmetrisk $n \times n$-matrise. (Det er den stort sett alltid når vi er interessert i kvadratiske former.)

La oss først studere et eksempel. Funksjonen gitt ved

$$f(x, y) = x^2 + y^2 + xy$$

har et bunnpunkt i origo, og hessematrise

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$\boxed{1}$ Diagonaliser H og regn ut $v^T A v$ for forskjellige vektorer v og tenk på hva dette forteller om det kritiske punktet til f.

Ok her er fasiten. Vi diagonaliserer matrisen

$$H = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Egenvektorene er

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{og} \quad \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

med respektive egenverdier 3 og 1. Vi definerer

$$V = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

og beregner

$$V^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

og dobbeltsjekker ved å beregne produktet

$$V^{-1}AV = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = D$$

For eksempel er

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = 3$$

og

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

Siden alle vektorer i $\mathbb{R}^2$ åpenbart kan skrives som en lineærkombinasjon av de to egenvektorene til H, må $\mathbf{v}^T A\mathbf{v} > 0$ for alle $\mathbf{v} \in \mathbb{R}^2$.

$\boxed{2}$ Hvorfor det?

La oss nå repetere noe viktig. La A være en $n \times n$-matrise med m lineært uavhengige egenvektorer $\mathbf{v}_1$, $\mathbf{v}_2 \ldots \mathbf{v}_m$ og tilhørende egenverdier λ_1, λ_2,$\cdots$, λ_m, . For hver egenvektor gjelder

$$A\mathbf{v}_k = \lambda_k \mathbf{v}_k.$$

Disse m ligningene kan like gjerne organiseres i en matriseligning

$$AV = VD.$$

$\boxed{3}$ Hvorfor det?

Siden $\mathbf{v}_1$, $\mathbf{v}_2 \ldots \mathbf{v}_m$ er lineært uavhengige, er matrisen V invertibel dersom $m = n$. Vi ganger med V^{-1} fra venstre og får

$$V^{-1}AV = D.$$

Denne operasjonen kalles å *diagonalisere* A. Man kan også gå motsatt vei. Dersom

$$V^{-1}AV = D$$

for en inverterbar $n \times n$-matrise V, kan vi gange fra venstre med V, og få

$$AV = VD$$

Vi ser av denne likningen at kolonnene til V utgjør n lineært uavhengige egenvektorer for A.

$\boxed{4}$ Hvorfor det?

Vi sier at en $n \times n$-matrise med n lineært uavhengige egenvektorer er diagonaliserbar, siden vi kan skrive

$$V^{-1}AV = D$$

hvis og bare hvis A har n lineært uavhengige egenvektorer.

$\boxed{5}$ Er

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

diagonaliserbar?

Matrisen i oppgaven over er ikke diagonaliserbar. Men det er en type matriser som alltid er.

$\boxed{6}$ Diagonaliser matrisen

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

og kikke nøye på egenvektorene.

$\boxed{7}$ Vi vet at lengden til en egenvektor ikke spiller noen rolle. Forklar at man i eksemplet over kan skrive

$$V^T A V = D$$

Dette leder oss over til noe som kalles ortogonale matriser. La oss først repetere litt. Skalarproduktet mellom vektorer i $\mathbb{R}^n$ er:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

og vi sier at $\mathbf{x}$ og $\mathbf{y}$ er ortogonale om $\mathbf{x} \cdot \mathbf{y} = 0$. Det går an å definere lengde på mange fornuftige måter, for eksempel

$$\|\mathbf{x}\|_1 = \sum_{k=1}^{n} |x_k|$$

eller

$$\|\mathbf{x}\|_\infty = \max_k |x_k|$$

men den vanligste er nok den pytagoreiske lengden

$$\|\mathbf{x}\|_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

siden dette er den fysiske lengden til $\mathbf{x}$ for $n = 2$ og $n = 3$. Vi sier at denne lengden er indusert av skalarproduktet, og dersom det ikke er noen subindeks, er det underforstått at det er denne vi mener.

Hvordan kan vi utvide alt dette til vektorer i $\mathbb{C}^n$? I lys av avsnittet over, er det lurt å definere

$$\mathbf{x} \cdot \mathbf{y} = \sum_{k=1}^{n} \overline{x_k} y_k = \overline{x_1} y_1 + \overline{x_2} y_2 + \cdots + \overline{x_n} y_n$$

siden indreproduktet av $\mathbf{x}$ med seg selv da blir en sum av de kvadrerte absoluttverdiene til komponentene til $\mathbf{x}$ og dette virker som et fornuftig mål på lengden til $\mathbb{C}^n$.

$\boxed{8}$ Sjekk at skalarproduktet på $\mathbb{C}^n$ over er et komplekst indreprodukt.

Siden dette er noe vi kan komme til å gjøre ofte, er det funnet opp en spesiell notasjon som tar i bruk matrisemultiplikasjon. 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

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

der radene og kolonnene i A er byttet om. Dersom A er reell, sier vi istedet transponert og skriver A^T. Siden $A^* = \overline{A}$ dersom A er en 1×1-matrise, ser vi at vi nå har to notasjoner for komplekskonjugert, $\overline{z}$ og z^*. Noen fagfelt foretrekker den ene, og noen foretrekker den andre, så det er greit å vite om begge. Vi kan nå skrive skalarproduktet som:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^* \mathbf{y}$$

Vi sier at en matrise A er ortonormal dersom

$$A^* A = I$$

Det er ikke så vanskelig å se at en ortogonal matrise har ortogonale kolonner. Siden ortogonale vektorer åpenbart må være lineært uavhengige, ser vi også at $A^* = A^{-1}$ dersom A er kvadratisk.

$\boxed{9}$ Finn den adjungerte

$$\begin{pmatrix} 1 & 1+i & -i \\ 1-i & 0 & 2-i \\ i & 2+i & 2 \end{pmatrix}$$

og sjekk om kolonnene er ortogonale.

Dersom A har n ortonormale egenvektorer, sier vi at A er ortogonalt diagonaliserbar. Dersom vi setter egenvektorene inn som kolonner i en $n \times n$-matrise V, ser vi at

$$V^* A V = V^* V D = D$$

siden $V^* V = I$. Motsatt ser vi at dersom

$$V^* A V = D,$$

for en ortonormal matrise V, utgjør V sine kolonner n ortonormale egenvektorer for D. Vi kan derfor skrive

$$V^* A V = D$$

hvis og bare hvis A har n ortogonale egenvektorer.

$\boxed{10}$ Diagonaliser

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 6 & 2 \\ 2 & 2 & 6 \end{pmatrix}$$

Er matrisen inverterbar?

En kompleks matrise sies å være hermittisk dersom $A = A^*$. Dersom A er reell er $A^* = A^T$. I litteraturen er det vanlig å reservere begrepet *symmetrisk* for reelle matriser der $A = A^T$. Jeg har aldri sett noen transponere en kompleks matrise uten å komplekskonjugere den, så man kunne i bunn og grunn bare sagt symmetrisk for alle matriser.

$\boxed{11}$ Er

$$\begin{pmatrix} 1 & 1+i & -i \\ 1-i & 0 & 2-i \\ i & 2+i & 2 \end{pmatrix}$$

symmetrisk?

Nå kommer en oppgave som ikke egentlig er vanskelig, men man må ha litt trening i å holde hodet kaldt.

$\boxed{12}$ Vis at dersom $A = A^*$ er $\mathbf{x}^* A \mathbf{x}$ reell.
(Hint: Skriv ut komponentene i matrisemultiplikasjonene.)

Den neste oppgaven følger ganske greit om man kan trikset.

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

Samme med neste.

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

Det er nå mulig å vise at en symmetrisk matrise er ortogonalt diagonaliserbar, men jeg tror det er litt for hardt. Det motsatte er ikke sant, det finnes ortogonalt diagonaliserbare matriser som ikke er symmetriske.

Når vi først er på galeien med å liste opp ting som er sant, men for vanskelige å bevise, kan vi like gjerne introdusere noe som er lett å huske. Vi sier at en matrise er normal dersom $A^*A = AA^*$. En matrise er ortogonalt diagonaliserbar hvis og bare hvis den er normal.

Vi er nå i stand til å gå tilbake til der vi startet, nemlig klassifisering av ekstremalpunkter. Hessematrisen er symmetrisk dersom f er deriverbar, og nå vet vi at hessematrisen derfor er ortogonalt diagonaliserbar. Dersom alle egenverdiene til hessematrisen er positive, sier vi at matrisen er positivt definitt.

$\boxed{16}$ Har vi da et maksimums- eller minimumspunkt?

Tilsvarende gjelder for negativt definitt. Dersom noen egenverdier er positive og noen negative, har vi sadelpunkt.

La oss avslutte med en siste anvendelse av symmetriske matriser og kvadratiske former. Normalfordelingen i flere variable er gitt ved

$$f(x) = \frac{1}{\sqrt{\|2\pi\Sigma\|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

der Σ kovariansmatrisen mellom komponentene i X. Denne matrisen er alltid symmetrisk, og man kan merke seg at innmaten i eksponensialfunksjonen er en kvadratisk form. Derfor er det mulig å skissere nivåkurvene til ormalfordelingen ved å vite litt om geometrien til kvadratiske former.

$\boxed{17}$ Nivåkurvene er visst alltid ellipser. Vis dette.