Egenverdier for 2×2 matriser

(Bearbeidet versjon av tidligere notat på nett-sidene til MA1201 - Lineær algebra og geometri. Versjon oppdatert med referanser til 10.utg av læreboken.)

Egenvektorer og egenverdier er introdusert på s. 295, Anton/Rorres (10.utg.).

Definisjon 1. Hvis $T\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ er en lineær operator, kalles et tall λ som er slik at

$$T\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

for en $\mathbf{x} \neq \mathbf{0}$ i $\mathbb{R}^n$ en **egenverdi** til T.

En $\mathbf{x} \neq \mathbf{0}$ som oppfyller (1) kalles **egenvektor** svarende til egenverdien λ.

Det er mange anvendelser der disse begrepene er nyttige. Vi skal senere se på et par slike anvendelser for å motivere innføringen av egenverdi og egenvektor.

Vi vet at for en lineær operator T finnes det en $n \times n$-matrise A slik at (1) kan representeres på formen:

$$A\mathbf{x} = \lambda\mathbf{x} \tag{2}$$

hvor A er en $n \times n$-matrise, $\mathbf{x}$ er en $n \times 1$-matrise og $A\mathbf{x}$ står for vanlig matrise-multiplikasjon.

Vi kan da omskrive (2) til formen:

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{3}$$

Dette er altså et homogent likningssystem med n likninger og n ukjente der matrisen har formen

$$(A - \lambda I)\mathbf{x} = \begin{bmatrix} (a_{11} - \lambda) & a_{12} & \cdots & a_{1n} \\ a_{21} & (a_{22} - \lambda) & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & (a_{nn} - \lambda) \end{bmatrix}. \tag{4}$$

Siden vi bare er interessert i løsninger av likningssystemet (3) som er $\neq \mathbf{0}$, må vi først bestemme λ slik at

$$\det(A - \lambda I) = 0.$$

Vi vet fra teorien at systemet da har uendelig mange løsninger - og da spesielt løsninger $\neq \mathbf{0}$.

Hvorfor er vi ikke interessert i løsninger $\mathbf{x} = \mathbf{0}$ av (1)? For det første ser vi jo straks at $\mathbf{x} = \mathbf{0}$ er en løsning av (1). Men når vi ser nærmere på de anvendelser vi har av egenverdier/vektorer senere, ser vi straks at den trivielle null-løsningen er helt uinteressant! Derfor må vi starte med å bestemme de λ-verdiene som gjør at $\det(A - \lambda I) = 0$. Vi innser lett fra (4) at vi da får å løse en likning av grad n:

$$\lambda^n + \alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \cdots \alpha_{n-1}\lambda + \alpha_n = 0$$

Vi skal i dette notatet kun holde oss til tilfellet $n = 2$.

Eksempel 1. Vi skal bestemme egenverdier og egenvektorer for matrisen:

$$A = \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix}.$$

Vi må da bestemme røttene i likningen:

$$\begin{vmatrix} (5-\lambda) & -6 \\ 2 & (-2-\lambda) \end{vmatrix} = 0.$$

Utregning gir:

$$\lambda^2 - 3\lambda + 2 = 0$$

som kan faktoriseres til:

$$(\lambda - 1)(\lambda - 2) = 0$$

Vi får da egenverdier $\lambda_1 = 1$ og $\lambda_2 = 2$.

$\underline{\lambda_1 = 1:}$

Vi vender tilbake til likningssystemet i (3) for å finne egenvektoren tilhørende denne λ-verdien:

$$\begin{bmatrix} (5-1) & -6 \\ 2 & (-2-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

som ved Gauss-Jordan blir:

$$\left[\begin{array}{rr|r} 2 & -3 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

som gir løsningene:

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} t \quad ; \quad t \neq 0$$

$\underline{\lambda_2 = 2:}$

For denne verdi av λ blir likningssystemet:

$$\begin{bmatrix} (5-2) & -6 \\ 2 & (-2-2) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

som ved Gauss-Jordan blir:

$$\left[\begin{array}{rr|r} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

som gir løsningene:

$$\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} t \quad ; \quad t \neq 0$$

Merknad. Vi observerer straks at dersom λ_0 er en egenverdi med egenvektor $\mathbf{v}_0$, så er også $k\mathbf{v}_0$ en egenvektor for denne samme egenverdi for hver $k \neq 0$ i $\mathbb{R}$. Vi har nemlig:

$$A(k\mathbf{v}_0) = k(A\mathbf{v}_0) = k(\lambda_0 \mathbf{v}_0) = \lambda_0(k\mathbf{v}_0).$$

Derfor kan vi f.eks. gjerne bestemme t for våre to egenvektorer $\mathbf{v}_1$ og $\mathbf{v}_2$ i eksemplet ovenfor slik at begge har lengde (norm) lik 1:

$$\mathbf{v}_1 = \frac{1}{\sqrt{13}} \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Oppgave 1. Kontroller ved utregning at $A\mathbf{v}_1 = \mathbf{v}_1$ og $A\mathbf{v}_2 = 2\mathbf{v}_2$ i eksemplet ovenfor.

Vi trenger nå følgende:

Definisjon 2. Vi sier at en $n \times n$-matrise A er **diagonaliserbar** dersom det finnes en invertibel $n \times n$-matrise P som er slik at $P^{-1}AP = D$ der D er en diagonalmatrise.

Vi skal illustrere dette vha. vårt eksempel ovenfor. Vi velger nå

$$P = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}.$$

Vi benytter formelen fra boken (Teorem 1.4.5, s. 43) for å bestemme P^{-1}.

$$P^{-1} = \frac{1}{3-4} \begin{bmatrix} 1 & -2 \\ -2 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix}$$

Vi har da:

$$P^{-1}AP = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} 5 & -6 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Altså har vi diagonalisert matrisen A. Det som er et spørsmål her er hvordan man fant en "riktig" matrise P som var slik at

$$P^{-1}AP = D$$

der D er en diagonalmatrise. Vi merker oss videre at D har formen

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

der $\lambda_1 = 1$ og $\lambda_2 = 2$ – altså egenverdiene til matrisen vi startet med. Er dette en ren tilfeldighet? Vi kan også observere at $P = \begin{bmatrix} \mathbf{v}_1 \mid \mathbf{v}_2 \end{bmatrix}$ der $\mathbf{v}_1$ og $\mathbf{v}_2$ er egenvektorer til egenverdiene $\lambda_1 = 1$ og $\lambda_2 = 2$. Er dette tilfeldigheter eller kan vi bevise at vi alltid har en slik sammenheng?

Oppgave 2. La A være en 2×2-matrisen gitt ved:

$$A = \begin{bmatrix} 9 & 2 \\ 2 & 6 \end{bmatrix}$$

Bestem egenverdiene λ_1 og λ_2, og vis at tilhørende egenvektorer er:

$$\mathbf{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

La P være gitt som:

$$P = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$$

Vis at da må $P^{-1}AP = D$, der

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

og der λ_1 og λ_2 er egenverdiene til A.

Teorem 1. *La A være en 2×2-matrise med to reelle egenverdier λ_1 og λ_2 der $\lambda_1 \neq \lambda_2$. La $\mathbf{v}_1 = \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix}$ og $\mathbf{v}_2 = \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix}$ være to tilhørende egenvektorer. La videre*

$$P = \begin{bmatrix} \mathbf{v}_1 \mid \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$$

Da vil: $P^{-1}AP = D$ der D er diagonalmatrisen:

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

(Merknad: Dette teorem viser at det vi observerte i eksemplet og oppgaven ovenfor ikke var tilfeldig!)

Bevis: Vi beviser først at matrisen P definert som ovenfor nødvendigvis er invertibel. Vi har fra boken at

$$P \text{ er invertibel hvis og bare hvis } \det P \neq 0.$$

Det gjelder her å bevise at

$$\begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} \neq 0.$$

Anta at det motsatte, dvs:

$$\begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} = 0.$$

Siden vi har:

$$\begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix} = \begin{vmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{vmatrix}$$

følger det fra Teorem 3.5.4 (a), s. 166, at denne determinanten $= 0$ hvis og bare hvis arealet av parallellogrammet utspent av vektorene $\mathbf{v}_1$ og $\mathbf{v}_2$ blir 0. Siden både $\mathbf{v}_1$ og $\mathbf{v}_2$ er $\neq \mathbf{0}$, betyr dette at $\mathbf{v}_1 = k\mathbf{v}_2$ for en skalar $k \neq 0$. Ut fra tidligere merknad betyr dette at:

$$A\mathbf{v}_1 = \lambda_2 \mathbf{v}_1,$$

mao. at også $\mathbf{v}_1$ er en egenvektor svarende til egenverdien λ_2. Dette betyr at $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ og $A\mathbf{v}_1 = \lambda_2 \mathbf{v}_1$ som gir $\lambda_1 \mathbf{v}_1 = \lambda_2 \mathbf{v}_1$ eller

$$(\lambda_1 - \lambda_2)\mathbf{v}_1 = \mathbf{0}.$$

Siden $\mathbf{v}_1 \neq \mathbf{0}$ betyr dette at $\lambda_1 = \lambda_2$, i strid med antagelsen i teoremet. Altså er P invertibel. Det gjenstår da å bevise at:

$$AP = PD.$$

Vi har da

$$AP = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \left[\, A\mathbf{v}_1 \mid A\mathbf{v}_2 \,\right]$$

siden

$$A\mathbf{v}_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = \begin{bmatrix} a_{11}v_{11} + a_{12}v_{21} \\ a_{21}v_{11} + a_{22}v_{21} \end{bmatrix}$$

og

$$A\mathbf{v}_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = \begin{bmatrix} a_{11}v_{21} + a_{12}v_{22} \\ a_{21}v_{21} + a_{22}v_{22} \end{bmatrix}$$

Videre har vi:

$$PD = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{bmatrix} = \left[\begin{array}{c|c} \lambda_1 \mathbf{v}_1 & \lambda_2 \mathbf{v}_2 \end{array} \right].$$

Siden $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ og $A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$ har vi dermed bevist at:

$$AP = PD$$

eller ekvivalent, siden P er invertibel:

$$P^{-1}AP = D. \qquad \square$$

Korollar 2. *For hvert naturlig tall n har vi*

$$A^n = PD^nP^{-1}.$$

Bevis: Fra $P^{-1}AP = D$ følger $A = PDP^{-1}$ ut fra regneregler for matriseprodukt. Altså holder likheten for $n = 1$. Anta så at

$$A^k = PD^kP^{-1}.$$

Vi har da:

$$\begin{aligned} A^{k+1} = A^k A &= (PD^kP^{-1})(PDP^{-1}) \\ &= PD^k(P^{-1}P)DP^{-1} = PD^kDP^{-1} = PD^{k+1}P^{-1}. \end{aligned}$$

Ut fra induksjonsprinsippet holder derfor $A^n = PD^nP^{-1}$ for alle $n \in \mathbb{N}$. $\square$

Anvendelse. På en øy lever det rever og kaniner. Ved tiden $t = 0$ er antallet rever lik $R_0 = 100$ og antallet kaniner $K_0 = 100$. Vi betegner antall rever etter n måneder med R_n og antall kaniner etter n måneder med K_n. Vi stiller opp følgende sammenheng:

$$R_{n+1} = 0.4R_n + 0.3K_n \tag{5}$$

$$K_{n+1} = -0.4R_n + 1.2K_n \tag{6}$$

Hva er idéen bak denne matematiske modellen for utviklingen av disse to dyrebestander? Leddet $0.4R_n$ gir at dersom kaninbestanden etter n måneder er lik 0, vil bare 40% av revene overleve til neste måned. Leddet $0.3K_n$ i likning (5) ovenfor betyr at revebestanden øker avhengig av hvor mange kaniner som finnes. I likning (6) har vi antatt at kaninbestanden øker med 20% på en måned dersom ingen rever finnes, mens leddet $-0.4R_n$ representerer revenes innvirkning på kaninbestanden avhengig av antall rever etter n måneder.

Vi har da to lineære likninger som angir utviklingen i bestanden fra måned n til måned $n+1$. Hvorvidt denne modellen er realistisk eller ikke kan vel diskuteres. Men den samme type matematiske modell benyttes ved forskjellige studier innenfor populasjonsdynamikk.

Vi kan stille opp ovenstående likning på matriseform:

$$\begin{bmatrix} R_{n+1} \\ K_{n+1} \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 4 & 3 \\ -4 & 12 \end{bmatrix}\begin{bmatrix} R_n \\ K_n \end{bmatrix}.$$

Vi ønsker å finne ut hvordan de to bestandene endrer seg etter f.eks 10 måneder – og hvordan utviklingen endrer seg i det lange løp. Vil f.eks. en eller kanskje begge bestandene dø ut fordi revene spiser så mange kaniner at de selv lider sultedøden?

Vi får da å regne ut

$$\begin{bmatrix} R_{n+1} \\ K_{n+1} \end{bmatrix} = \frac{1}{10^n}\begin{bmatrix} 4 & 3 \\ -4 & 12 \end{bmatrix}^n \begin{bmatrix} R_0 \\ K_0 \end{bmatrix} = A^n \begin{bmatrix} R_0 \\ K_0 \end{bmatrix}$$

der A er den angitte matrise. Etter en relativt omstendig regning kommer man fram til at

$$A^{10} \approx \begin{bmatrix} -0.491 & 0.745 \\ -0.994 & 1.497 \end{bmatrix}.$$

Ytterligere regning gir indikasjon på at

$$A^n \underset{n\to\infty}{\longrightarrow} \begin{bmatrix} -1/2 & 3/4 \\ -1 & 3/2 \end{bmatrix}$$

Det gir

$$A^n \begin{bmatrix} 100 \\ 100 \end{bmatrix} \underset{n\to\infty}{\longrightarrow} \begin{bmatrix} 25 \\ 50 \end{bmatrix}.$$

Vi prøver å studere det samme problemet ved å diagonalisere A. Vi må da bestemme egenverdiene.

$$|A - \lambda I| = \begin{vmatrix} (2/5-\lambda) & 3/10 \\ -4/10 & (12/10-\lambda) \end{vmatrix} = 0$$

som gir likningen

$$(2/5-\lambda)(12/10-\lambda) + 12/100 = 0$$

eller

$$(\lambda - 1)(\lambda - 3/5) = 0$$

som gir egenverdiene

$$\lambda_1 = 1 \quad ; \quad \lambda_2 = 3/5$$

Bestemmelse av egenvektorene gir:
$\underline{\lambda_1 = 1:}$

$$\begin{bmatrix} -3/5 & 3/10 \\ -4/10 & 1/5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad ; \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} t \; ; \; t \neq 0.$$

$\underline{\lambda_2 = 3/5:}$

$$\begin{bmatrix} -1/5 & 3/10 \\ -4/10 & 3/5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad ; \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} t \; ; \; t \neq 0.$$

Dermed settes

$$P = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

og vi får

$$P^{-1} = -\frac{1}{4} \begin{bmatrix} 2 & -3 \\ -2 & 1 \end{bmatrix}.$$

Derfor har vi

$$P^{-1}AP = D = \begin{bmatrix} 1 & 0 \\ 0 & \frac{3}{5} \end{bmatrix}$$

som gir

$$\begin{aligned} A^n = PD^nP^{-1} &= -\frac{1}{4} \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{3}{5} \end{bmatrix}^n \begin{bmatrix} 2 & -3 \\ -2 & 1 \end{bmatrix} \\ &= -\frac{1}{4} \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & (\frac{3}{5})^n \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -2 & 1 \end{bmatrix} \\ &= -\frac{1}{4} \begin{bmatrix} 1 & 3(\frac{3}{5})^n \\ 2 & 2(\frac{3}{5})^n \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -2 & 1 \end{bmatrix} \\ &= -\frac{1}{4} \begin{bmatrix} 2 - 6(\frac{3}{5})^n & 3(\frac{3}{5})^n - 3 \\ 4 - 4(\frac{3}{5})^n & 2(\frac{3}{5})^n - 6 \end{bmatrix} \end{aligned}$$

Dette gir så

$$\begin{aligned} A^n \begin{bmatrix} R_0 \\ K_0 \end{bmatrix} &= -\frac{1}{4} \begin{bmatrix} 2 - 6(\frac{3}{5})^n & 3(\frac{3}{5})^n - 3 \\ 4 - 4(\frac{3}{5})^n & 2(\frac{3}{5})^n - 6 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \end{bmatrix} \\ &= \begin{bmatrix} 25 + 75(\frac{3}{5})^n \\ 50 + 50(\frac{3}{5})^n \end{bmatrix} \underset{n \to \infty}{\longrightarrow} \begin{bmatrix} 25 \\ 50 \end{bmatrix}. \end{aligned}$$

Når n vokser vil altså bestanden stabilisere seg på 25 rever og 50 kaniner på øyen.

Definisjon 3. En $n \times n$-matrise A sies å være **ortogonal** dersom $A^T = A^{-1}$.

Merknad. Det innsees lett at A er ortogonal hvis og bare hvis $A^T A = I$: Hvis $A^T = A^{-1}$ vet vi at $A^T A = A^{-1} A = I$. Dersom $A^T A = I$ er A^T venstreinvers til A, og det følger fra Teorem 1.6.3(a) at $A^T = A^{-1}$.

Eksempel 2. Dersom $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ blir $A^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. Vi har da

$$A^T A = \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Altså er A ortogonal.

Eksempel 3. $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ er ortogonal ut fra eksempel 2 siden $\cos\frac{\pi}{4} = \sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}$.

Definisjon 4. La $\mathbf{v}_1$ og $\mathbf{v}_2$ være to vektorer i $\mathbb{R}^2$. Vi sier at $\{\mathbf{v}_1, \mathbf{v}_2\}$ er en **ortogonal mengde** dersom $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. $\{\mathbf{v}_1, \mathbf{v}_2\}$ sies å være en **ortonormal mengde** dersom man i tillegg har at $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$.

Teorem 3. *En 2×2-matrise A er ortogonal hvis og bare hvis søylevektorene $\mathbf{v}_1$ og $\mathbf{v}_2$ utgjør en ortonormal mengde.*

Bevis: Hvis $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, er $\mathbf{v}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}$ og $\mathbf{v}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$. Vi har da $A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$ som gir

$$A^T A = \begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{bmatrix}$$

ut fra vanlige regler for matrisemultiplikasjon og definisjon av skalarprodukt.(Kontroller denne påstanden på egen hånd!)

Dersom A er ortogonal , skal pr. def. $A^T A = I$ eller ekvivalent:

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_1 &= \|\mathbf{v}_1\|^2 = 1 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 &= 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_2 &= \|\mathbf{v}_2\|^2 = 1 \end{aligned}$$

Altså utgjør søylevektorene i A en ortonormal mengde.

Dersom søylevektorene i A utgjør en ortonormal mengde, blir $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$ og $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Altså følger det av det ovenstående at $A^T A = I$, mao. A er en ortogonal matrise. $\square$

Definisjon 5. Vi sier at A er **ortogonalt diagonaliserbar** dersom det finnes en ortogonal matrise P som er slik at $P^T AP = D$ der D er en diagonalmatrise.

Teorem 4. *A ortogonalt diagonaliserbar hvis og bare hvis A er symmetrisk.*

(Vi minner om at A er symmetrisk dersom $A = A^T$.)

Bevis: Vi antar først at A er ortogonalt diagonaliserbar. Altså at det finnes en ortogonal matrise P slik at:

$$P^T AP = D$$

der D er en diagonalmatrise. Dette gir:

$$A = PDP^T.$$

siden $P^T = P^{-1}$. Av dette følger så:

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PD^T P^T = PDP^T = A$$

ut fra Teorem 1.4.8 og det faktum at en diagonalmatrise er symmetrisk. Altså er A symmetrisk.

Vi antar så at $A^T = A$. Vi skal først bevise at da har A enten to reelle egenverdier λ_1 og λ_2 der $\lambda_1 \neq \lambda_2$ eller A har formen:

$$A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

I det siste tilfellet er A allerede en diagonalmatrise og diagonalisering er derfor ikke nødvendig. En annen sak er at $I^T AI = A$, slik at A formelt er ortogonalt diagonaliserbar. La oss anta at $A = A^T$, dvs. at

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

Vi studerer som vanlig likningen:

$$|A - \lambda I| = \begin{vmatrix} (a-\lambda) & b \\ b & (c-\lambda) \end{vmatrix} = 0$$

Dette gir

$$\lambda^2 - (a+c)\lambda + ac - b^2 = 0$$

Vi ser på diskrimanten Δ i formelen for røttene i 2.gradslikningen:

$$\begin{aligned}\Delta = B^2 - 4AC &= (a+c)^2 - 4 \cdot 1(ac - b^2) \\ &= a^2 + 2ac + c^2 - 4ac + 4b^2 \\ &= (a-c)^2 + (2b)^2\end{aligned}$$

Dette uttrykket er aldri negativt og lik 0 hvis og bare hvis $a = c$ og $b = 0$, mao. når

$$A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

I alle andre tilfeller har derfor A to relle og innbyrdes ulike egenverdier. Vi betegner egenvektorene tilhørende egenverdiene λ_1 og λ_2 henholdsvis $\mathbf{v}_1$ og $\mathbf{v}_2$. Vi har da:

$$A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \quad ; \quad A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$$

Ut fra tidligere bemerkning kan vi anta uten tap av generalitet at $\|\mathbf{v}_1\| = 1$ og $\|\mathbf{v}_2\| = 1$. Vi minner nå om at skalarproduktet av to vektorer kan uttrykkes vha. produktet av to matriser i det man oppfatter en 1×1-matrise som et reelt tall:

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

(Se s. 139.) Videre er det klart at

$$B^T = B$$

når B er en 1×1-matrise. Dette gir:

$$\begin{aligned}\lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2) &= \mathbf{v}_1 \cdot (\lambda_2 \mathbf{v}_2) = \mathbf{v}_1 \cdot (A\mathbf{v}_2) = \mathbf{v}_1^T (A\mathbf{v}_2) \\ &= (\mathbf{v}_1^T A \mathbf{v}_2)^T = \mathbf{v}_2^T A^T (\mathbf{v}_1^T)^T = \mathbf{v}_2^T A \mathbf{v}_1 \\ &= \mathbf{v}_2^T (\lambda_1 \mathbf{v}_1) = \mathbf{v}_2 \cdot (\lambda_1 \mathbf{v}_1) = \lambda_1 (\mathbf{v}_2 \cdot \mathbf{v}_1)\end{aligned} \tag{7}$$

Vi har mao.

$$\lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_2) = \lambda_2(\mathbf{v}_1 \cdot \mathbf{v}_2)$$

Siden $\lambda_1 \neq \lambda_2$, må derfor $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Altså er $\{\mathbf{v}_1, \mathbf{v}_2\}$ en ortonormal mengde. Hvis vi som tidligere definerer P ved $P = \left[\begin{array}{c|c} \mathbf{v}_1 & \mathbf{v}_2 \end{array}\right]$ har vi foran bevist at P er ortogonal og

$$P^{-1}AP = D$$

der D er diagonalmatrisen

$$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \qquad \square$$

Oppgave 3. Kontroller riktigheten av alle likhetene i (7).

Fra ovenstående bevis har vi også:
(a) Dersom A er symmetrisk og ikke er en diagonalmatrise, har A to reelle egenverdier, $\lambda_1 \neq \lambda_2$, med tilhørende egenvektorer $\mathbf{v}_1$ og $\mathbf{v}_2$ som danner en ortonormal mengde. Da er $P = \left[\begin{array}{c|c} \mathbf{v}_1 & \mathbf{v}_2 \end{array} \right]$ en ortogonal matrise og $A = PDP^T$.
(b) Hvis A allerede er en diagonalmatrise er diagonalisering unødvendig.

Merknad. Hvis A ikke er symmetrisk har den ikke nødvendigvis to rellle egenverdier. Vi har for

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

at likningen $|A - \lambda I| = 0$ blir $\lambda^2 + 1 = 0$ som **ikke** har reelle røtter.