Sylvesters kriterium
– Nej, ikke mit kriterium

Rasmus Sylvester Bryder

Inspireret af en statistikers manglende råd om hvornår en kvadratisk matrix er positivt definit uden at skulle ud i at bestemme dens egenværdier, ledte jeg efter en betingelse for, at en matrix har denne egenskab, der kunne være nem at vise i praksis. Jeg skulle ikke lede længe efter et: Sylvesters kriterium, som vises sidst i denne artikel, er et tilpas simpelt og samtidigt overraskende resultat (kort sagt, det kilder en lille smule, når man læser det). For at komme frem til sætningen vil vi undervejs bevæge os forbi en række virkelig interessante sætninger om komplekse matricer, som jeg ikke selv har set før forarbejdet til denne artikel, og forhåbentlig vil man kunne drage nytte af disse. For at alle kan følge med, har jeg sørget for, at jeg kun bevæger mig inden for territoriet trådt under LinAlg-kurset, og jeg vil henvise til alles førsteårs-LinAlg-bog [2] i det følgende i form af sætningsangivelser sat i parenteser.

Først noget notation. Vi definerer $\mathbb{M}_{m,n}$ som værende vektorrummet af $m \times n$-matricer med indgange i $\mathbb{C}$. $\mathbb{M}_n$ betegner rummet $\mathbb{M}_{n,n}$, og $I_n \in \mathbb{M}_n$ betegner enhedsmatricen (bestående af ettaller i diagonalen og nuller alle andre steder). For enhver matrix $A = (a_{ij})$ i $\mathbb{M}_{m,n}$ (forstået på den måde, at a_{ij} er indgangen i skæringen mellem i'te række og j'te søjle) defineres den *adjungerede* matrix $A^* = (b_{ij}) \in \mathbb{M}_{n,m}$ ved $b_{ij} = \overline{a_{ji}}$; for eksempel vil vi have for

$$A = \begin{pmatrix} 1 & i \\ -i & 0 \\ 4 & 2+3i \end{pmatrix},$$

at dens adjungerede er givet ved

$$A^* = \begin{pmatrix} 1 & i & 4 \\ -i & 0 & 2-3i \end{pmatrix}.$$

Regneregler for adjungering er såre simple og ganske vigtige: der gælder, at

$$(AB)^* = B^*A^*, \quad (\lambda A_1 + \mu A_2)^* = \overline{\lambda} A_1^* + \overline{\mu} A_2^*, \quad (A^*)^* = A$$

for alle $A_1, A_2 \in \mathbb{M}_{m,n}$, $B \in \mathbb{M}_{n,p}$ og $\lambda, \mu \in \mathbb{C}$. En kompleks $n \times n$-matrix A kaldes *selvadjungeret* eller *hermitisk* hvis $A^* = A$.

I tilfælde af, at nogen har glemt et ekstremt relevant koncept fra LinAlg, bliver det kort genopfrisket her: ved en *egenvektor* for en matrix $A \in \mathbb{M}_n$ forstås en vektor $x \in \mathbb{C}^n$, så x *ikke* er nulvektoren, og $Ax = \lambda x$ for et $\lambda \in \mathbb{C}$. λ kaldes i dette fald den til x hørende *egenværdi* for A. A kaldes *diagonaliserbar*, hvis der findes en basis for $\mathbb{C}^n$ bestående af egenvektorer for A. Et centralt og ganske uomstændigt resultat fra lineær algebra lyder, at egenværdierne for A netop er rødderne til det såkaldte karakteristiske polynomium defineret ved $\lambda \mapsto \det(A - \lambda I_n)$ (6.1.8), så ved Algebraens Fundamentalsætning har enhver kompleks matrix A i alt n egenværdier talt med multiplicitet (summa summarum: de behøver ikke at være forskellige).

I det følgende opfattes enhver vektor $x \in \mathbb{C}^n$ endvidere automatisk som en $n \times 1$-matrix og enhver 1×1-matrix som et komplekst tal. Vi definerer et indre produkt på $\mathbb{C}^n$ ved

$$\langle x, y \rangle = y^* x = \sum_{i=1}^{n} x_i \overline{y_i}$$

for vektorer $x = (x_1, \dots, x_n)$ og $y = (y_1, \dots, y_n)$ med komplekse indgange.[1]

Lemma 1 *Alle egenværdier for en selvadjungeret matrix A er reelle.*

Bevis. Antag, at λ er en egenværdi for A med en tilhørende egenvektor x. Da vil

$$\begin{aligned}\lambda\|x\|^2 &= \langle \lambda x, x\rangle = \langle Ax, x\rangle \\ &= x^*Ax = x^*A^*x = (Ax)^*x \\ &= \langle x, Ax\rangle = \langle x, \lambda x\rangle \\ &= \overline{\lambda}\|x\|^2,\end{aligned}$$

så $\lambda = \overline{\lambda}$; altså vil $\lambda \in \mathbb{R}$. □

Nu til det vigtige.

Definition 2 En matrix $U \in \mathbb{M}_n$ kaldes *unitær*, hvis $U^*U = UU^* = I_n$.

Vi husker, at en $n \times n$-matrix er unitær, hvis og kun hvis dens søjler udgør en ortonormalbasis i $\mathbb{C}^n$ jf. (7.2.3); vi kommer til at bruge dette en masse i det følgende. Vi starter først med en ordentlig moppedreng.

Sætning 3 (Schurs triangulariseringssætning, 1909) *For hver kompleks matrix $A \in \mathbb{M}_n$ findes en unitær matrix $U \in \mathbb{M}_n$ således at U^*AU er en øvre trekantsmatrix.*

[1] Nogle studerende er måske mere vant til notationen $x \cdot y$ frem for $\langle x, y\rangle$. Under alle omstændigheder gælder regnereglerne (7.1.1) selvfølgelig stadig.

Bevis. Vi viser dette ved induktion efter n; i tilfældet $n = 1$ er der intet at vise (sæt $U = I_1$). Antag derfor, at sætningen gælder for komplekse $n \times n$-matricer; vi viser ud fra dette, at sætningen holder for $(n+1) \times (n+1)$-matricer. Lad derfor $A \in \mathbb{M}_{n+1}$, og lad λ være en egenværdi for A med en tilhørende egenvektor $x \in \mathbb{C}^{n+1}$. Da vil

$$x_1 = \frac{1}{\|x\|} x$$

også være en egenvektor for A hørende til λ. Mængden $\{x_1\}$ kan nu udvides til en basis $\{x_1, \ldots, x_{n+1}\}$ for $\mathbb{C}^{n+1}$ jf. (4.5.12) og ved Gram-Schmidt-ortogonalisering (7.1.8) opnås en ortonormalbasis $\{z_1, \ldots, z_{n+1}\}$ for $\mathbb{C}^{n+1}$, hvor $z_1 = \frac{1}{\|x_1\|} x_1 = x_1$. Lades U være matricen givet ved

$$U = \begin{pmatrix} z_1 & \cdots & z_{n+1} \end{pmatrix},$$

altså med søjle i lig vektoren z_i, vil U dermed være unitær. Den første søjle i matricen U^*AU er nu produktet U^*AUe_1 jf. (1.3.5), hvor $e_1 = (1, 0, \ldots, 0) \in \mathbb{C}^{n+1}$. Bemærk, at $U^{-1} = U^*$, hvormed $Ue_1 = z_1 = x_1$ og $U^{-1} z_1 = e_1$, således at

$$U^*AUe_1 = U^*A(Ue_1) = U^*Ax_1 = U^*(\lambda x_1) = \lambda(U^{-1}x_1) = \lambda e_1,$$

hvor vi brugte, at x_1 var egenvektor for A. Altså vil

$$U^*AU = \begin{pmatrix} \lambda & B \\ 0_{n,1} & A_1 \end{pmatrix},$$

hvor B er en $1 \times n$-matrix, $0_{n,1}$ er en $n \times 1$-matrix med nuller overalt og A_1 er en $n \times n$-matrix (vi har skrevet U^*AU som en *blokmatrix*). Ved induktionsantagelsen findes en unitær matrix

$U_1 \in \mathbb{M}_n$ så $U_1^* A_1 U_1$ er en øvre trekantsmatrix med diagonalen bestående af egenværdier for A'. Bemærk, at blokmatricen

$$U' = \begin{pmatrix} 1 & 0_{1,n} \\ 0_{n,1} & U_1 \end{pmatrix} \in \mathbb{M}_{n+1}$$

er unitær: dens søjlevektorer har alle længde 1 og er indbyrdes ortogonale, hvorpå de tilmed er lineært uafhængige jf. (7.1.6) og derfor udgør en basis for $\mathbb{C}^{n+1}$ jf. (4.5.7). Bemærk endvidere, at

$$\begin{aligned}(UU')^* C(UU') &= (U')^*(U^*CU)U' \\ &= \begin{pmatrix} 1 & 0_{1,n} \\ 0_{n,1} & U_1^* \end{pmatrix} \begin{pmatrix} \lambda & B \\ 0_{n,1} & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0_{1,n} \\ 0_{n,1} & U_1 \end{pmatrix} \\ &= \begin{pmatrix} \lambda & B' \\ 0_{n,1} & U_1^* A_1 U_1 \end{pmatrix},\end{aligned}$$

hvor B' er en $1 \times n$-matrix[2]. Eftersom $U_1^* A_1 U_1$ er en øvre trekantsmatrix fra induktionsantagelsen, fås at sidst forekommende matrix i ovenstående udregninger også er en øvre trekantsmatrix. Da UU' er unitær (tjek selv!), fås det ønskede for $(n+1) \times (n+1)$-matricer, så induktion giver, at sætningen holder for alle n. □

Bemærk, at ovenstående sætning ikke nødvendigvis for reelle matricer A atter giver en reel unitær matrix U.

Korollar 4 *For $A \in \mathbb{M}_n$ er determinanten af A lig produktet af de n egenværdier for A.*

[2]Her anbefales det at bruge et par minutter på at overbevise sig om, hvorledes blokmatricer af denne form ganges sammen og hvorfor vi får det ovenstående.

Bevis. Tag jf. ovenstående sætning en unitær matrix U således at $T = U^*AU$ er en øvre trekantsmatrix. Da er

$$\det(A - \lambda I_n) = \det(U^*(A - \lambda I_n)U) = \det(T - \lambda I_n)$$

for alle $\lambda \in \mathbb{C}$, således at egenværdierne for A netop er egenværdierne for T. Lades $\lambda_1, \ldots, \lambda_n$ være diagonalelementerne i T, ses hurtigt, at

$$\det(T - \lambda I_n) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n), \quad \lambda \in \mathbb{C},$$

da T er en øvre trekantsmatrix. Altså er egenværdierne for T lig diagonalelementerne i T. Da

$$\det(A) = \det(UTU^*) = \det(T) = \lambda_1 \cdots \lambda_n,$$

følger det ønskede. □

Korollar 5 *For enhver selvadjungeret matrix $A \in \mathbb{M}_n$ findes en ortonormalbasis af egenvektorer for A, og egenværdierne knyttet til de n egenvektorer er netop de n egenværdier for A. Hermed er A diagonaliserbar.*

Bevis. Lad $U \in \mathbb{M}_n$ være en unitær matrix, således at U^*AU er en øvre trekantsmatrix jf. Schurs triangulariseringssætning. Da

$$(U^*AU)^* = U^*A^*(U^*)^* = U^*A^*U = U^*AU,$$

er U^*AU også en nedre trekantsmatrix, idet den er lig $(U^*AU)^*$, hvormed U^*AU er en diagonalmatrix $D \in \mathbb{M}_n$. Søjlevektorerne i U udgør en ortonormalbasis for $\mathbb{C}^n$, da U er unitær. Lad z_i være den i'te søjlevektor i U. Da vil $Ue_i = z_i$, hvor e_i er den i'te basisvektor i $\mathbb{C}^n$ (1.3.5). Betegnes det i'te diagonalelement i D med λ_i, fås nu

$$Az_i = AUe_i = U(U^*AUe_i) = U(De_i) = U(\lambda_i e_i) = \lambda_i(Ue_i) = \lambda_i z_i.$$

Altså er den i'te søjlevektor i U en egenvektor for A med tilhørende egenværdi lig det i'te diagonalelement i D. Ved beviset for Korollar 4 indses, at egenværdierne $\lambda_1, \ldots, \lambda_n$ netop udgør alle egenværdier for A. $\square$

Vi nærmer os nu det ønskede kriterium så småt. Først skal bemærkes, at hvis W_1 og W_2 er underrum af et vektorrum V, da er $W_1 \cap W_2$ og

$$W_1 + W_2 = \{x_1 + x_2 \,|\, x_1 \in W_1,\ x_2 \in W_2\}$$

også underrum af V. Hvis V_1 og V_2 er vektorrum, opnås et nyt vektorrum ved at betragte produktmængden

$$V_1 \times V_2 = \{(x_1, x_2) \,|\, x_1 \in V_1,\ x_2 \in V_2\}$$

og definere addition og skalarmultiplikation på denne koordinatvist, altså ved

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2), \quad \lambda(x_1, y_1) := (\lambda x_1, \lambda x_2)$$

for $x_1, x_2 \in V_1$, $y_1, y_2 \in V_2$ og $\lambda \in \mathbb{C}$. Dette vektorrum kaldes den *direkte sum* af V_1 og V_2 og betegnes $V_1 \oplus V_2$. Det er endvidere let at indse, at hvis $\{x_1, \ldots, x_m\}$ er en basis for V_1 og $\{y_1, \ldots, y_n\}$ er en basis for V_2, opnås en basis for $V_1 \oplus V_2$ ved vektorsættet

$$\{(x_1, 0), \ldots, (x_m, 0), (0, y_1), \ldots, (0, y_n)\}.$$

Altså er $\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2$. Ud fra disse konstruktioner kan vi udlede følgende meget vigtige formel som konsekvens af Dimensionssætningen.

Lemma 6 (Grassmanns formel) *Lad W_1 og W_2 være underrum af et endeligdimensionalt vektorrum V. Da vil*

$$\dim W_1 + \dim W_2 = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

Bevis. Definér en lineær afbildning $f\colon W_1 \oplus W_2 \to W_1 + W_2$ ved $f(x,y) = x + y$. f er surjektiv. Hvis $f(x,y) = 0$ for en vektor $(x,y) \in W_1 \oplus W_2$, vil $x + y = 0$ og dermed $y = -x$ i V. Endvidere vil $x \in W_1$ og $x = -y \in W_2$, så $x \in W_1 \cap W_2$. Dermed vil

$$\ker f \subseteq \{(x,-x) \mid x \in W_1 \cap W_2\}.$$

Omvendt ses, at $f(x,-x) = 0$ for alle $x \in W_1 \cap W_2$, så kernen er *lig* ovenstående mængde.

$\ker f$ er et underrum af $W_1 \oplus W_2$. Defineres nu en afbildning $g\colon \ker f \to W_1 \cap W_2$ ved $g(x,y) = x$, er det let at vise, at g er lineær og en isomorfi. Altså er $\dim \ker f = \dim(W_1 \cap W_2)$, hvormed vi ved Dimensionssætningen (4.7.7) får

$$\begin{aligned}\dim W_1 + \dim W_2 &= \dim(W_1 \oplus W_2)\\ &= \dim(f(W_1 \oplus W_2)) + \dim(\ker f)\\ &= \dim(W_1 + W_2) + \dim(W_1 \cap W_2),\end{aligned}$$

og dermed det ønskede. □

Nu er vi beredte og bevæbnede til tænderne med en kavalkade af evergreens inden for sætninger om matricer. Tally-ho!

Definition 7 En matrix $A \in \mathbb{M}_n$ er *positivt definit*, hvis $x^*Ax > 0$ for alle $x \in \mathbb{C}^n$, hvor x ikke er nulvektoren.

Lemma 8 *En selvadjungeret matrix $A \in \mathbb{M}_n$ er positivt definit, hvis og kun hvis alle dens egenværdier er strengt positive. Specielt er determinanten af A positiv, hvis A er positivt definit.*

Bevis. Hvis A er positivt definit og λ er en egenværdi for A med en tilhørende egenvektor x, vil

$$0 < x^* A x = \langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \|x\|^2,$$

hvormed vi ved division med $\|x\|^2$ får, at $\lambda > 0$.

For at vise den anden implikation, bemærker vi, da A er selvadjungeret, at der findes en ortonormalbasis for A bestående af egenvektorer jf. Korollar 5. Lad $\{z_1, \ldots, z_n\}$ være en sådan, med den til z_i hørende egenværdi betegnet λ_i. For ethvert $x \in \mathbb{C}^n$ findes da entydigt bestemte tal $\mu_1, \ldots, \mu_n \in \mathbb{C}$ så $x = \sum_{i=1}^n \mu_i z_i$, hvormed

$$\begin{aligned}
x^* A x &= \langle Ax, x \rangle \\
&= \left\langle \sum_{i=1}^n \lambda_i \mu_i z_i, \sum_{j=1}^n \mu_j z_j \right\rangle \\
&= \sum_{i,j=1}^n \lambda_i \mu_i \overline{\mu_j} \, \langle z_i, z_j \rangle \\
&= \sum_{i=1}^n \lambda_i \mu_i \overline{\mu_i} \langle z_i, z_i \rangle \\
&= \sum_{i=1}^n \lambda_i |\mu_i|^2,
\end{aligned}$$

da z_i'erne er ortonormale. Hvis alle egenværdier er strengt positive og x er forskellig fra nulvektoren, må der være et $j \in \{1, \ldots, n\}$, så $\mu_j \neq 0$ hvormed $x^* A x \geq \lambda_i |\mu_i|^2 > 0$. Altså er A positivt definit. Det sidste resultat følger af Korollar 4. □

Følgende pudsige begreber er essentielle for det næste (og sidste) resultat.

Definition 9 Lad $A \in \mathbb{M}_n$. For $k = 1, \ldots, n$ kaldes den $k \times k$-matrix, der opnås ved at slette de sidste $n-k$ rækker og søjler i A, for den *k'te ledende undermatrix* af A. Den *k'te ledende underdeterminant* af A er determinanten af den k'te ledende undermatrix af A.

Sætning 10 (Sylvesters kriterium) *Lad $A \in \mathbb{M}_n$ være selvadjungeret. Da er A positivt definit hvis og kun hvis dens ledende underdeterminanter er positive.*

Bevis. Antag først, at A er positivt definit, lad $k \in \{1, \ldots, n\}$ og lad A_k være den k'te ledende undermatrix af A. For enhver vektor $x \in \mathbb{C}^k$ forskellig fra nulvektoren defineres $\tilde{x} \in \mathbb{C}^n$ ved $\tilde{x} = (x_1, \ldots, x_k, 0, \ldots, 0)$, hvor $x = (x_1, \ldots, x_k)$. Specielt er $\tilde{x}$ ikke nulvektoren, og ved direkte udregning indses, at

$$x^* A_k x = \tilde{x}^* A \tilde{x} > 0,$$

således at A_k er positivt definit. Specielt fås ved Lemma 8, at determinanten af A_k er positiv og dermed fås den ene implikation.

Den anden retning vises ved induktion efter n, hvori starttilfældet $n = 1$ er klart. Antag derfor, at det gælder for alle selvadjungerede $n \times n$-matricer, at positivitet af de ledende underdeterminanter medfører, at matricen er positivt definit, og lad A være en selvadjungeret $(n+1) \times (n+1)$-matrix med positive ledende underdeterminanter. Ved Lemma 8 ses, at det er nok at vise, at egenværdierne for A er positive. Grundet induktionsantagelsen vil den n'te ledende undermatrix af A være positivt definit; kald denne for A_n. Eftersom A er selvadjungeret, findes ved Korollar 5 en ortonormalbasis $z_1, \ldots, z_{n+1}$ af egenvektorer for A med tilhørende egenværdier $\lambda_1, \ldots, \lambda_{n+1}$, som udgør alle egenværdier for A. Vi kan ved omarrangering sørge for, at $\lambda_1 \leq \cdots \leq \lambda_{n+1}$, idet egenværdierne er reelle ved Korollar 4. Som i beviset for Lemma 8 har vi nu for ethvert $x \in \mathbb{C}^{n+1}$ på formen $x = \sum_{i=1}^{n+1} \mu_i z_i$, at

$$x^*Ax = \sum_{i=1}^{n} \lambda_i |\mu_i|^2.$$

Bemærk, at μ_i'erne er entydigt bestemt ud fra x, eftersom z_i'erne er en basis. Hvis $x \in \operatorname{span}\{z_1, z_2\}$, vil vi i så fald have

$$\mu_3 = \ldots = \mu_{n+1} = 0,$$

så

$$x^*Ax = \lambda_1|\mu_1|^2 + \lambda_2|\mu_2|^2 \leq \lambda_2(|\mu_1|^2 + |\mu_2|^2) = \lambda_2\|x\|^2,$$

hvoraf sidste lighedstegn indses ved at regne på $\langle x, x\rangle$ direkte.

Lad W være det underrum af vektorer i $\mathbb{C}^{n+1}$, hvis $(n+1)$'te koordinat er lig 0. For enhver vektor $x = (x_1, \ldots, x_n, 0) \in W$ forskellig fra nulvektoren indses nu igen ved direkte udregning, at

$$x^*Ax = (x')^*A_n x' > 0,$$

hvor $x' = (x_1, \ldots, x_n) \in \mathbb{C}^n$, idet A_n er positivt definit. Eftersom $\operatorname{span}\{z_1, z_2\}$ har dimension 2, W har dimension n og de begge er indeholdt i vektorrummet $\mathbb{C}^{n+1}$ af dimension $n+1$, må vi have

$$\dim(W \cap \operatorname{span}\{z_1, z_2\}) > 0$$

og dermed $W \cap \operatorname{span}\{z_1, z_2\} \neq \{0\}$ ved Grassmanns formel. Der findes altså en vektor $x \in W \cap \operatorname{span}\{z_1, z_2\}$ forskellig fra nulvektoren, hvorved vi får

$$\lambda_2\|x\|^2 \geq x^*Ax > 0$$

og dermed $\lambda_{n+1} \geq \cdots \geq \lambda_2 > 0$. Eftersom

$$\lambda_1 \cdots \lambda_{n+1} = \det A > 0,$$

må $\lambda_1 > 0$, så alle egenværdier for A er positive, hvormed A er positivt definit. Induktion giver nu den anden retning for alle n, og altså følger Sylvesters kriterium. $\square$

Litteratur

[1] Kumaresan, S. *Sylvester Criterion for Positive Definiteness*, University of Hyderabad, Department of Mathematics and Statistics, 2004.

[2] Pedersen, Niels Vigand. *Forelæsningsnoter til lineær algebra*, 2. udgave, Københavns Universitet, Matematisk Afdeling, 2009.

[3] Prasolov, Viktor Vasil'evich. *Problems and theorems in linear algebra*, American Mathematical Society, 1994.