Sylvesters treghetslov

Harald Hanche-Olsen

9. mai 1995

Sammendrag

Notatet gir et alternativt bevis for Sylvesters treghetslov ("law of inertia"). Beviset i Strang er i og for seg korrekt, men krever mer teori for å kunne kalles rigorøst. Beviset som presenteres her er mer likt det man vanligvis finner i lærebøker, men er mer geometrisk av natur.

I dette notatet vil A være en symmetrisk, reell $n \times n$-matrise. Husk at A er ortogonalt diagonaliserbar, slik at det finnes ortonormale vektorer $v_1, \ldots, v_n$ som danner en basis for $\mathbf{R}^n$ og slik at $Av_i - \lambda_i v_i$ for $i = 1, \ldots, n$.

1 Definisjon La $\mathrm{d}_+(A)$ være antall positive egenverdier for A, $\mathrm{d}_-(A)$ antall negative egenverdier, og $\mathrm{d}_0(A)$ være antall egenverdier som er lik 0. Vi teller med multiplisitet, slik at $\mathrm{d}_+(A) + \mathrm{d}_-(A) + \mathrm{d}_0(A) = n$ (med andre ord, $\mathrm{d}_+(A)$ er antall i slik at $\lambda_i > 0$ og så videre).

Fordi det ikke er noen enkel sammenheng mellom egenverdier og egenvektorer for A og de for $C^{\mathrm{T}}AC$, må beviset for Sylvesters teorem (se nedenfor!) gå veien om en annen karakterisering av $\mathrm{d}_+(A)$.

2 Definisjon Et underrom $V \subseteq \mathbf{R}^n$ kalles A-*positivt* dersom $x^{\mathrm{T}}Ax > 0$ for enhver vektor $x \in V$ med $x \neq 0$.

Dette er en nyttig definisjon fordi $CV = \{Cv : v \in V\}$ er A-positivt hvis og bare hvis V er $C^{\mathrm{T}}AC$-positivt. Beviset for denne påstanden overlater vi til leseren.

3 Lemma $\mathrm{d}_+(A)$ *er den største mulige dimensjonen for et* A*-positivt underrom av* $\mathbf{R}^n$.

Bevis: Rommet utspent av egenvektorene med positiv egenverdi har dimensjon $\mathrm{d}_+(A)$. Mer presist, om vi sorterer egenverdiene i avtagende rekkefølge får vi $\lambda_i > 0$ for $i = 1, \ldots, \mathrm{d}_+(A)$ mens $\lambda_i \leq 0$ når $i > \mathrm{d}_+(A)$. En vilkårlig vektor $x \in \mathbf{R}^n$ skrives $x = \sum_{i=1}^n c_i v_i$, og vi finner

$$x^{\mathrm{T}}Ax = \sum_{i=1}^{n} c_i^2 \lambda_i$$

fordi vektorene v_i er ortogonale egenvektorer for A, slik at $v_i^{\mathrm{T}}Av_j = \lambda_j v_i^{\mathrm{T}} v_j = 0$ om $i \neq j$, mens $v_i^{\mathrm{T}}Av_i = \lambda_i$ siden vi har regnet med $\|v_i\| = 1$.

Av formelen ovenfor følger umiddelbart at $x^{\mathrm{T}}Ax > 0$ når $x \neq 0$ er en lineærkombinasjon av $v_1, \ldots, v_{\mathrm{d}_+(A)}$, slik at rommet V utspent av disse vektorene er et A-positivt underrom. Åpenbart har dette rommet dimensjon $\mathrm{d}_+(A)$.

Vi ser også at $x^{\mathrm{T}}Ax \leq 0$ for $x \in V^{\perp}$, siden $V^{\perp}$ er utspent av vektorene v_i med $i > \mathrm{d}_+(A)$, dvs. $\lambda_i \leq 0$.

La nå U være et annet A-positivt underrom av $\mathbf{R}^n$. Vi lar P være projeksjonsmatrisen på V, det vil si $P = \sum_{i=1}^{\mathrm{d}_+(A)} v_i v_i^{\mathrm{T}}$.

Om $x \in U$ og $Px = 0$ er altså $x \in U \cap V^{\perp}$. Men $x \in V^{\perp}$ impliserer $x^{\mathrm{T}}Ax \leq 0$ mens $x \in U$ impliserer enten $x^{\mathrm{T}}Aa > 0$ eller $x = 0$. Eneste mulighet er da $x = 0$, og projeksjonsmatrisen definerer derfor en en-til-en lineæravbildning fra U til V. Det er bare mulig dersom $\dim(U) \leq \dim(V) = \mathrm{d}_+(A)$, slik at $\mathrm{d}_+(A)$ er den største mulige dimensjon til et A-positivt underrom som påstått. □

Som et *eksempel* kan vi betrakte matrisen

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

Det A-positive rommet V konstruert i lemmaet er xy-planet, mens et rom generelt er A-positivt hvis og bare hvis det ikke treffer dobbeltkjeglen $z^2 \geq x^2 + y^2$ utenfor origo. Ved å tegne noen figurer kan du lett overbevise deg om at ortogonalprojeksjon fra et A-positivt underrom til xy-planet er en-til-en. Det var dette som motiverte konstruksjonen av beviset i dette notatet.

4 Teorem (Sylvester) *Om A er en symmetrisk, reell $n \times n$-matrise og C er en invertibel $n \times n$-matrise, er $\mathrm{d}_+(A) = \mathrm{d}_+(C^{\mathrm{T}}AC)$, $\mathrm{d}_-(A) = \mathrm{d}_-(C^{\mathrm{T}}AC)$ og $\mathrm{d}_0(A) = \mathrm{d}_0(C^{\mathrm{T}}AC)$.*

Bevis: At $\mathrm{d}_+(A) = \mathrm{d}_+(C^{\mathrm{T}}AC)$ følger direkte fra lemmaet og bemerkningen foran lemmaet.

Dermed følger likheten $\mathrm{d}_-(A) = \mathrm{d}_-(C^{\mathrm{T}}AC)$ fordi $\mathrm{d}_-(A) = \mathrm{d}_+(-A)$.

At $\mathrm{d}_0(A) = \mathrm{d}_0(C^{\mathrm{T}}AC)$ følger av at $\mathrm{d}_0(A) = \dim \mathcal{N}(A) = n - r$ der r er rangen til A, og at A og $C^{\mathrm{T}}AC$ har samme rang som A siden multiplikasjon med en invertibel matrise på den ene eller andre siden ikke endrer rangen.

Alternativt kan man bruke identiteten $\mathrm{d}_+(A) + \mathrm{d}_-(A) + \mathrm{d}_0(A) = n$. □