Egenvärden och egenvektorer

Definition 1 *Antag att A är en $n \times n$-matris. En n-vektor $\mathbf{v} \neq \mathbf{0}$ som är sådan att A verkar som multiplikation med ett tal λ på $\mathbf{v}$, d v s*

$$A\mathbf{v} = \lambda\mathbf{v},\ \lambda \in \mathbb{R},$$

kallas för en **egenvektor** *till A. Talet λ kallas för ett* **egenvärde** *till A.*

Exempel 1
1. *Låt $A = s \cdot I$ vara matrisen som skalar med talet s. Då är alla vektorer egenvektorer med egenvärdet s.*

2. *Mer allmänt så låt A vara en diagonalmatris $A = diag(a_1, a_2, \ldots, a_n)$. Om $\mathbf{e}_i$ är standardbasvektor nummer i så gäller att*

$$A\mathbf{e}_i = a_i\mathbf{e}_i,$$

så $\mathbf{e}_i$ är alltså en egenvektor till A med egenvärdet a_i.

3. *Låt $A = \left(\begin{smallmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{smallmatrix}\right)$ vara rotation kring origo vinkeln t moturs. Om $t \notin \{0, \pi\}$ så saknar A egenvektorer, ty då är aldrig $\mathbf{v}$ och $A\mathbf{v}$ parallella.*

4. *Låt A vara projektionen i $\mathbb{R}^3$ på xy-planet, d v s*

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

Då gäller att

$$A \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

så $(0\ 0\ 1)^t$ är en egenvektor med egenvärdet 0. Om $\mathbf{v} = (x\ y\ 0)^t$ är en vektor i xy-planet så är $A\mathbf{v} = \mathbf{v}$ så alla dessa är egenvektorer med egenvärdet 1. För alla andra vektorer är $\mathbf{v}$ och $A\mathbf{v}$ ej parallella så dessa är de enda egenvektorerna.

Anmärkning 1 *Om $\mathbf{v}$ är en egenvektor till A med egenvärde λ och $c \neq 0$. Då gäller att*

$$A(c\mathbf{v}) = c(A\mathbf{v}) = c(\lambda\mathbf{v}) = \lambda(c\mathbf{v}),$$

så $c\mathbf{v}$ är också en egenvektor med samma egenvärde. Speciellt kan man alltid välja en egenvektor med given längd, t ex kan man välja en enhetsvektor.

Hur kan man beräkna egenvärde och egenvektorer när man inte har enkla fall då man kan lösa det geometriskt? Följande metod fungerar i princip, men i praktiken endast för små matriser. Vi har

$$A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow A\mathbf{x} - \lambda\mathbf{x} = 0 \Leftrightarrow (A - \lambda I)\mathbf{x} = 0$$

så $(A - \lambda I)\mathbf{x} = 0$ har alltså en icke-trivial lösning. Detta är ekvivalent med att

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

(Vi måste ha minst en nollrad efter Gauss-elimination.) Ekvationen med determinanten kallas för den **karakteristiska ekvationen** och är en polynomekvation av grad n.

Exempel 2 *Bestäm egenvärden och egenvektorer till matrisen*

$$A = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}.$$

Den karakteristiska ekvationen blir

$$\begin{aligned} 0 &= |A - \lambda I| = \left| \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| \\ &= \begin{vmatrix} 3/4 - \lambda & 1/4 \\ 1/4 & 3/4 - \lambda \end{vmatrix} = (3/4 - \lambda)^2 - 1/16 \end{aligned}$$

vilket är ekvivalent med

$$(3/4 - \lambda)^2 = 1/16 \Leftrightarrow \lambda = \frac{3}{4} \pm \sqrt{\frac{1}{16}} = \frac{3}{4} \pm \frac{1}{4} = \begin{cases} 1 \\ \frac{1}{2} \end{cases}$$

Egenvektorerna är lösningar till $(A - \lambda I)\mathbf{x} = \mathbf{0}$ *för* $\lambda = 1$ *och* $\lambda = 1/2$.
$\lambda = 1$*:*

$$(A - 1 \cdot I) = \begin{pmatrix} -1/4 & 1/4 \\ 1/4 & -1/4 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1/4 & 1/4 \\ 0 & 0 \end{pmatrix}$$

Vi får $\mathbf{x} = \left(\begin{smallmatrix} t \\ t \end{smallmatrix}\right)$*, så t ex* $\mathbf{x} = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ *är en egenvektor till egenvärdet* 1.
$\lambda = 1/2$*:*

$$(A - \frac{1}{2} \cdot I) = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1/4 & 1/4 \\ 0 & 0 \end{pmatrix}$$

Vi får $\mathbf{x} = \left(\begin{smallmatrix} t \\ -t \end{smallmatrix}\right)$*, så t ex* $\mathbf{x} = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ *är en egenvektor till egenvärdet* 1/2.

Vi såg i exemplen ovan att antalet egenvärden till en matris kan variera. Eftersom karakteristiska ekvationen för en $n \times n$-matris är ett polynom av grad n så kan det ha allt från 0 till n (reella) egenvärden. Vad som är viktigare än antalet egenvärden är antalet linjärt oberoende egenvektorer. Vi såg t ex att en skalningsmatris $A = s \cdot I$ bara hade ett egenvärde s, men att alla vektorer var egenvektorer så att den har n linjärt oberoende egenvektorer. Dessa kan dessutom väljas så att de blir ortogonala. Detta är sant för symmetriska matriser i allmänhet och vi har följande så kallade **spektralsats** för symmetriska matriser.

Sats 1 *Antag att* A *är en* $n \times n$*-matris. Då har* A n *stycken ortogonala egenvektorer om och endast om* A *är symmetrisk.*

Det fullständiga beviset för denna sats är komplicerat, men följande viktiga delresultat är inte så svårt att bevisa:

Sats 2 *Antag att* A *är en symmetrisk* $n \times n$*-matris. Då är egenvektorer* $\mathbf{u}$ *respektive* $\mathbf{v}$ *ortogonala om de hör till olika egenvärden* μ *respektive* λ*.*

Bevis. Antag att A är symmetrisk, d v s att $A = A^t$, samt att $\mathbf{u}$ respektive $\mathbf{v}$ är egenvektorer till A med motsvarande egenvärden μ respektive λ där $\mu \neq \lambda$. Med andra ord så gäller att $A\mathbf{u} = \mu\mathbf{u}$ och $A\mathbf{v} = \lambda\mathbf{v}$. Det gäller att bevisa att $\mathbf{u} \cdot \mathbf{v} = 0$.

Genom att utnyttja att $A = A^t$ och att $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^t\mathbf{y}$ så får vi att

$$(A\mathbf{x}) \cdot \mathbf{y} = (A\mathbf{x})^t\mathbf{y} = (\mathbf{x}^t A^t)\mathbf{y} = \mathbf{x}^t(A^t\mathbf{y}) = \mathbf{x}^t(A\mathbf{y}) = \mathbf{x} \cdot (A\mathbf{y}),$$

för godtyckliga vektorer $\mathbf{x}$ och $\mathbf{y}$. Vi har att $A\mathbf{u} = \mu\mathbf{u}$ och $A\mathbf{v} = \lambda\mathbf{v}$ vilket ger

$$(A\mathbf{u}) \cdot \mathbf{v} = (\mu\mathbf{u}) \cdot \mathbf{v} = \mu(\mathbf{u} \cdot \mathbf{v}) \text{ och } \mathbf{u} \cdot (A\mathbf{v}) = \mathbf{u} \cdot (\lambda\mathbf{v}) = \lambda(\mathbf{u} \cdot \mathbf{v}).$$

Men enligt ovan är $(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A\mathbf{v})$ så $\mu(\mathbf{u} \cdot \mathbf{v}) = \lambda(\mathbf{u} \cdot \mathbf{v})$, d v s

$$(\mu - \lambda)(\mathbf{u} \cdot \mathbf{v}) = 0.$$

Eftersom $\mu \neq \lambda$ så följer det att $\mathbf{u} \cdot \mathbf{v} = 0$ och beviset är klart. $\square$

För icke-symmetriska matriser kan i stort sett vad som helst inträffa. Man vet dock att antalet linjärt oberoende egenvektorer inte är mindre än antalet olika egenvärden:

Sats 3 *Egenvektorer till olika egenvärden är linjärt oberoende.*

Bevis. Låt $\mathbf{v}_1, \dots, \mathbf{v}_n$ vara n stycken egenvektorer till A med motsvarande olika egenvärden $\lambda_1, \dots, \lambda_n$. Med andra ord så är $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ och $\lambda_i \neq \lambda_j$ om $i \neq j$.
För att visa att egenvektorerna är linjärt oberoende så gäller det per definition att visa att

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \tag{1}$$

bara har den triviala lösningen $c_1 = \cdots = c_n = 0$.
Vi gör ett bevis med hjälp av induktion över n. Om $n = 1$ så är det trivialt sant, ty egenvektorer är skilda från nollvektorn så

$$\mathbf{0} = c_1\mathbf{v}_1$$

har bara trivial lösning.
Antag nu att det alltid gäller om man man har $n-1$ olika egenvektorer och egenvärden för något givet $n > 1$. Vi ska visa att i så fall gäller det också om man har n stycken. Vi multiplicerar båda sidor i (1) med matrisen A (från vänster) och får

$$\mathbf{0} = A(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n = c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n.$$

Om vi multiplicerar (1) med λ_1 och subtraherar de två ekvationerna så får vi att

$$\mathbf{0} = c_2(\lambda_1 - \lambda_2)\mathbf{v}_2 + \cdots + c_n(\lambda_1 - \lambda_n)\mathbf{v}_n.$$

Enligt induktionsantagandet så gäller att $\mathbf{v}_2, \dots, \mathbf{v}_n$ är linjärt oberoende så eftersom $\lambda_1 - \lambda_i \neq 0$ om $i > 1$ så har denna ekvation bara den triviala lösningen $c_2 = \cdots = c_n = 0$. Därmed reduceras (1) till $\mathbf{0} = c_1\mathbf{v}_1$ och därmed är $c_1 = 0$ också. Vi har alltså visat att det bara finns den triviala lösningen till (1) och alltså att $\mathbf{v}_1, \dots, \mathbf{v}_n$ är linjärt oberoende,
Därmed kan vi med hjälp av induktionsaxiomet dra slutsaten att egenvektorer till olika egenvärden är linjärt oberoende.

□

Från denna sats får vi direkt följande:

Sats 4 *En (kvadratisk) matris med k olika egenvärden har åtminstone k linjärt oberoende egenvektorer. Speciellt gäller att en $n\times n$-matris med n olika egenvärden har n stycken linjärt oberoende egenvektorer som alltså utgör en bas för $\mathbb{R}^n$.*

Ett viktigt resultat får man av följande enkla räkningar där vi utnyttjar att multiplikation av tal med matris är kommutativ. Antag att $A\mathbf{v} = \lambda\mathbf{v}$, där $\lambda \in \mathbb{R}$. Då gäller

$$\begin{aligned} A^2\mathbf{v} &= A(A\mathbf{v}) = A(\lambda\mathbf{v}) = \lambda(A\mathbf{v}) = \lambda\cdot\lambda\mathbf{v} = \lambda^2\mathbf{v} \\ \mathbf{v} &= (A^{-1}A)\mathbf{v} = A^{-1}(A\mathbf{v}) = A^{-1}(\lambda\mathbf{v}) = \lambda(A^{-1}\mathbf{v}), \end{aligned}$$

så $A^{-1}\mathbf{v} = (1/\lambda)\mathbf{v}$. Med hjälp av en enkel induktion så generaliserar vi detta till:

Sats 5 *Antag att $\mathbf{v}$ är en egenvektor till A med egenvärdet λ, d v s $A\mathbf{v} = \lambda\mathbf{v}$. Då gäller att $\mathbf{v}$ är en egenvektor till A^m för alla $m \in \mathbb{Z}$ med egenvärdet λ^m, d v s*

$$A^m\mathbf{v} = \lambda^m\mathbf{v}.$$

Denna sats kan man bl a använda till att avgöra vad som händer med A^n när $n \to \infty$ och den ger också en grund för numeriska metoder att beräkna egenvärden.

Antag nu att A är en $n \times n$-matris med n linjärt oberoende egenvektorer $\mathbf{v}_1, \dots, \mathbf{v}_n$ med motsvarande egenvärden $\lambda_1, \dots, \lambda_n$. Vi antar att vi har ordnat dem så att $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Då utgör egenvektorerna en bas för $\mathbb{R}^n$ och därmed kan en godtycklig vektor $\mathbf{x}$ skrivas som

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

där $c_1, \dots, c_n$ är $\mathbf{x}$ koefficienter i basen $\mathbf{v}_1, \dots, \mathbf{v}_n$. Vi får att

$$A\mathbf{x} = A(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n = c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n,$$

och mer allmänt får man om man använder satsen ovan att

$$A^m\mathbf{x} = c_1\lambda_1^m\mathbf{v}_1 + \cdots + c_n\lambda_n^m\mathbf{v}_n.$$

Vi "normerar" genom att multiplicera med $1/\lambda_1^m$, där ju λ_1 är det egenvärde med störst belopp. Då får vi

$$(1/\lambda_1^m)A^m\mathbf{x} = c_1\mathbf{v}_1 + c_2\frac{\lambda_2^m}{\lambda_1^m}\mathbf{v}_2 + \cdots + c_n\frac{\lambda_n^m}{\lambda_1^m}\mathbf{v}_n.$$

Eftersom λ_i^m/λ_1^m går mot 0 då m växer mot oändligheten så kommer (en lämplig multipel) av $A^m\mathbf{x}$ att närma sig egenvektorn $\mathbf{v}_1$ när m blir större och större (i alla fall om $c_1 \neq 0$). Vi får alltså en egenvektor till det till beloppet största egenvärdet. Detta är den så kallade **potensmetoden** och den fungerar bara om det finns ett största (till beloppet) egenvärde.

Exempel 3 *Vi ska titta på den flyttningsmodell som vi jobbade med på gruppövningen som är uppgift 2.58 i boken. Vi hade där att vektorn* $\mathbf{x}_n$ *var en 3-vektor som beskrev hur många som bodde i storstad, tätort respektive på landsbygden efter* n *år. Dessa var relaterade genom*

$$\mathbf{x}_{n+1} = A\mathbf{x}_n,$$

där $A=$

$$\begin{pmatrix} 0.96 & 0.01 & 0.015 \\ 0.03 & 0.98 & 0.005 \\ 0.01 & 0.01 & 0.98 \end{pmatrix}.$$

Om $\mathbf{x}_0$ *är den ursprungliga befolkningsfördelningen så blir alltså*

$$\mathbf{x}_n = A^n\mathbf{x}_0.$$

Om man räknar ut egenvärdena för A *så får man att de är* 1, 0.97 *respektive* 0.95. *Vi ser att det finns tre olika egenvärden och därmed* 3 *linjärt oberoende egenvektorer* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *som man lätt kan beräkna när man har egenvärdena. Nu kan vi uttrycka den ursprungliga vektorn* $\mathbf{x}_0$ *i basen* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ *genom*

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3,$$

för några tal c_1, c_2 *och* c_3. *Om vi utnyttjar diskussionen ovan om potenser och egenvärden så får vi*

$$\begin{aligned} \mathbf{x}_n &= A^n\mathbf{x}_0 = A^n(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) \\ &= c_1\lambda_1^n\mathbf{v}_1 + c_2\lambda_2^n\mathbf{v}_2 + c_3\lambda_3^n\mathbf{v}_3 = c_1 \cdot 1^n\mathbf{v}_1 + c_2 \cdot 0.97^n\mathbf{v}_2 + c_3 \cdot 0.95^n\mathbf{v}_3. \end{aligned}$$

När n *går mot oändligheten så går* 0.97^n *och* 0.95^n *båda mot* 0 *så* $\mathbf{x}_n$ *närmar sig* $c_1\mathbf{v}_1$. *Fördelningen kommer alltså att närma sig (en multipel av) egenvektorn* $\mathbf{v}_1$. *Om man räknar ut och normerar* $\mathbf{v}_1$ *så att summan av elementen är* 1 *så får man*

$$\mathbf{v}_1 = \begin{pmatrix} 7/30 \\ 13/30 \\ 1/3 \end{pmatrix} \approx \begin{pmatrix} 0.2333 \\ 0.4333 \\ 0.3333 \end{pmatrix},$$

vilket man möjligen känner igen från gruppövningen. Detta betyder t ex att efter "lång tid" så kommer en tredjedel av befolkningen att bo på landsbygden. Observera att denna fördelning är oberoende av startvektorn. (Vän av ordning påpekar givetvis att vi måste ha $c_1 \neq 0$, *men så kommer alltid att vara fallet då vi har en startvektor med positiva koordinater.)*
Observera att det inte var en slump att 1 *var ett egenvärde och att alla andra egenvärden var mindre än* 1. *Det följer av problemets natur vilket vi återkommer till senare.*