8. Eigenwerte und Eigenvektoren

8.1. Einführung, Definition, Grundlagen

Eigenwertproblem
Gegeben ist eine $n \times n$-Matrix A. Gesucht sind

- ein vom Nullvektor verschiedener Vektor $\mathbf{x}$ und
- ein Skalar λ (auch $\lambda = 0$ ist erlaubt),

welche die Gleichung

$$A\mathbf{x} = \lambda\mathbf{x} \tag{10}$$

erfüllen.
Ein solches λ heißt *Eigenwert* von A, ein passendes $\mathbf{x}$ heißt *Eigenvektor* von A zum Eigenwert λ.

Die Situation „Matrix mal Eigenvektor ist Null mal Vektor“, also $A\mathbf{x} = 0\mathbf{x}$, kann durchaus auftreten. In so einem Fall ist $\lambda = 0$ ein Eigenwert von A.

Wir verlangen aber ausdrücklich, dass $\mathbf{x}$ nicht der Nullvektor 0 sein darf, weil die Gleichung $A \cdot 0 = \lambda 0$ immer und für beliebiges λ erfüllt ist. (Das wäre dann keine sinnvolle Definition von λ).

Anschauliche Interpretation

Der Hauptberuf einer Matrix ist, Vektoren zu multiplizieren. Wenn sie das macht, hat der Ergebnisvektor gewöhnlich eine andere Länge und zeigt in eine andere Richtung als der Ausgangsvektor. Jeder Matrix hat aber ganz spezielle „eigene“ Vektoren, bei denen sie zwar die Länge ändert, die Richtung aber gleich lässt (falls $\lambda > 0$) oder genau umkehrt (falls $\lambda < 0$). Es kann auch passieren (falls $\lambda = 0$), dass ein Eigenvektor von der Matrix zum Nullvektor gemacht wird. Reguläre Matrizen tun das aber nicht, nur singuläre Matrizen sind so fies.

Beispiel:
Die Matrix $A = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}$ macht aus $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$A\mathbf{u} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \end{bmatrix},$$

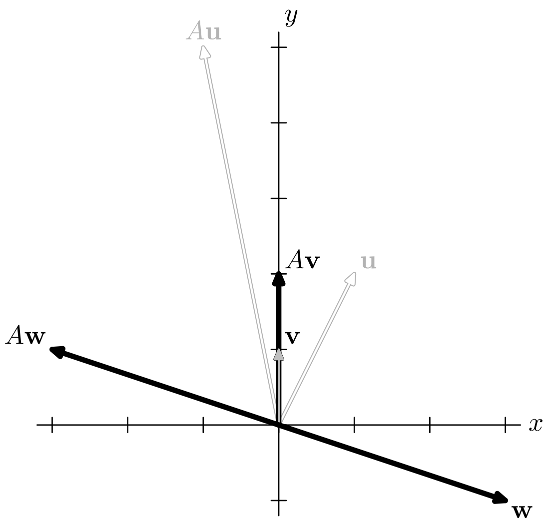
andere Länge und Richtung – kein Eigenvektor.
Der Vektor $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ wird hingegen zu

$$A\mathbf{v} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

gleiche Richtung, doppelte Länge –
Eigenvektor zum Eigenwert 2.
Der Vektor $\mathbf{w} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ wird zu

$$A\mathbf{w} = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

gleiche Länge, umgekehrte Richtung –
Eigenvektor zum Eigenwert -1.



Zu einem Eigenwert gibt es viele Eigenvektoren

Ist $\mathbf{x}$ ein Eigenvektor von A zum Eigenwert λ , so ist auch jedes Vielfache $\mu\mathbf{x}$, $\mu \neq 0$, ein Eigenvektor von A zum selben Eigenwert λ, denn da A, λ und $\mathbf{x}$ die Gleichung (10) erfüllen, gilt auch

$$A(\mu\mathbf{x}) = \mu(A\mathbf{x}) = \mu(\lambda\mathbf{x}) = \lambda(\mu\mathbf{x}).$$

Eigenvektoren sind also nur bis auf einen skalaren Faktor eindeutig bestimmt.

Eine $n \times n$-Matrix hat n Eigenwerte

Durch Umformung von (10) erhalten wir

$$\begin{aligned} A\mathbf{x} &= \lambda\mathbf{x} \\ A\mathbf{x} - \lambda\mathbf{x} &= 0 \\ (A - \lambda I)\mathbf{x} &= 0 \end{aligned} \tag{11}$$

wobei I die Einheitsmatrix ist. Der Vektor $\mathbf{x} = 0$ (der Nullvektor) ist immer eine Lösung von Gleichung (11), aber der interessiert uns nicht. Wir wollen, dass es noch andere, nicht triviale Lösungen gibt. Das ist nur genau dann der Fall, wenn

$$\det(A - \lambda I) = 0. \tag{12}$$

Entwickeln der Determinante liefert ein Polynom n-ten Grades in λ. Dieses Polynom heißt das *charakteristische Polynom* von A. Jedes Polynom n-ten Grades hat genau n reelle oder komplexe Nullstellen (sagt der Fundamentalsatz der Algebra; mehrfache Nullstellen zählt er dabei entsprechend ihrer Vielfachheit). Daraus folgt, dass jede $n \times n$-Matrix genau n (reelle oder komplexe, unter Umständen mehrfach gezählte) Eigenwerte hat.

Die Eigenwerte einer Matrix A lassen sich als Wurzeln (Nullstellen) ihres charakteristischen Polynoms berechnen. Dies ist ein klassische Verfahren, aber im allgemeinen nur für kleine Matrizen (2×2,3×3) sinnvoll.

Rechenbeispiel: Wir berechnen die Eigenwerte der Matrix A vom vorigen Beispiel (Seite 66).

$$\begin{aligned}\det(A - \lambda I) &= \det\left(\begin{bmatrix}-1 & 0\\ 1 & 2\end{bmatrix} - \lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right) = \det\begin{bmatrix}-1-\lambda & 0\\ 1 & 2-\lambda\end{bmatrix},\\ &= (-1-\lambda)\cdot(2-\lambda) - 0\cdot 1 = \lambda^2 - \lambda - 2\end{aligned}$$

Die Nullstellen des charakteristischen Polynoms $p(\lambda) = \lambda^2 - \lambda - 2$ und daher die Eigenwerte von A sind $\lambda_1 = -1$, $\lambda_2 = 2$.

Noch ein Beispiel: Wir berechnen die Eigenwerte der Matrix $A = \begin{bmatrix}8 & 1 & 6\\ 3 & 5 & 7\\ 4 & 9 & 2\end{bmatrix}$ als Wurzeln des charakteristischen Polynoms.

$$\det(A - \lambda I) = \det\left(\begin{bmatrix}8 & 1 & 6\\ 3 & 5 & 7\\ 4 & 9 & 2\end{bmatrix} - \lambda\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\right) = \det\begin{bmatrix}8-\lambda & 1 & 6\\ 3 & 5-\lambda & 7\\ 4 & 9 & 2-\lambda\end{bmatrix}$$

Berechnen der Determinante, z.B. mit der Regel von Sarrus, liefert

$$\begin{aligned}\det(A-\lambda I) &= (8-\lambda)(5-\lambda)(2-\lambda) + 28 + 162 - 24(5-\lambda) - 63(8-\lambda) - 3(2-\lambda)\\ &= -360 + 24\,\lambda + 15\,\lambda^2 - \lambda^3\end{aligned}$$

Das charakteristische Polynom von A lautet daher $p(\lambda) = -360 + 24\,\lambda + 15\,\lambda^2 - \lambda^3$, und seine Wurzeln (Nullstellen) können mit Verfahren des Kapitels 1 gefunden werden. Die drei Wurzeln, und damit die drei Eigenwerte von A sind $\lambda_1 = 15, \lambda_{2,3} = \pm 4{,}89898$.
Für größere Matrizen ist dieses Rechenverfahren zu umständlich und zu anfällig gegenüber Rundungsfehlern.

Kennt man einen Eigenwert λ, dann findet man einen dazu passenden Eigenvektor als Lösung des Gleichungssystems (11).

Für $\lambda = 15$ im obigen Beispiel suchen wir also eine Lösung des Systems

$$\begin{bmatrix}-7 & 1 & 6\\ 3 & -10 & 7\\ 4 & 9 & -13\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

In diesem System ergibt die Summe aus erster und zweiter Gleichung genau das Negative der dritten. Wenn Sie sich nun wundern, wie Sie aus nur zwei linear unabhängigen Gleichungen die drei Unbekannten x_1, x_2, x_3 bestimmen sollen, haben Sie die Pointe nicht verstanden: Das Gleichungssystem *darf nicht eindeutig lösbar sein!*. Wir haben mit der Determinantenbedingung (12) eigens dafür gesorgt, dass es außer der trivialen Lösung $x_1 = x_2 = x_3 = 0$ noch andere, sinvolle Lösungen gibt. Wir können zum Beispiel $x_1 = 1$ frei wählen und in die ersten beiden Gleichungen einsetzen, dann lassen sich x_2 und x_3 bestimmen:

$$\begin{array}{rcrcl} x_2 & + & 6x_3 & = & 7\\ -10x_2 & + & 7x_3 & = & 1\end{array} \qquad \text{Lösung: } x_2 = 1,\ x_3 = 1$$

8.2. Eigenwertaufgabe $A\mathbf{x} = \lambda\mathbf{x}$: weitere Eigenschaften

Matrizen in physikalischen Anwendungen sind zumeist symmetrisch (Steifigkeitsmatrix, Spannungstensor, Trägheitstensor...). Die zugehörigen Eigenwerte entsprechen physikalischen Größen (Schwingungsfrequenzen, Hauptspannungen, Trägheitsmomenten...), von denen man erwartet, dass sie reell sind. Da ist es beruhigend zu wissen, dass die Mathematik garantiert:

Alle Eigenwerte einer symmetrischen Matrix sind reell.

Für eine symmetrische Matrix A vereinfacht sich die Eigenwertaufgabe beträchtlich. Dazu gibt es eine reichhaltige und elegante mathematische Theorie und eine Fülle von Verfahren.

Die Eigenvektoren einer symmetrischen Matrix bilden ein orthogonales System.

Sind die Eigenvektoren einer symmetrischen Matrix A zusätzlich auch auf Länge 1 normiert, dann lassen sie sich in einer orthogonalen Matrix Q zusammenfassen. Dann gilt:

$$Q^T AQ = D \quad \text{mit } D \text{ Diagonalmatrix aus Eigenwerten} .$$

Eigenwerte lassen sich „verschieben“
Ist λ ein Eigenwert von A und $\mathbf{x}$ ein zugehöriger Eigenvektor, dann ist für beliebiges $s \in \mathbb{R}$ der Wert $\lambda + s$ Eigenwert von $A + sI$ und dasselbe $\mathbf{x}$ ist zugehöriger Eigenvektor.

Eigenwerte der Inversen Matrix sind Inverse der Eigenwerte
Ist λ ein Eigenwert von A mit zugehorigem Eigenvektor $\mathbf{x}$, dann ist $1/\lambda$ ein Eigenwert von A^{-1} mit demselben Eigenvektor $\mathbf{x}$.

Ähnlichkeitstransformation
Ist X eine invertierbare Matrix[5] , dann haben A und $X^{-1}AX$ die gleichen Eigenwerte. Die Transformation von A zu $X^{-1}AX$ heißt *Ähnlichkeitstransformation* .

Moderne Rechenverfahren nützen Ähnlichkeitstransformationen, die Verschiebung von Eigenwerten und das Berechnen inverser Eigenwerte geschickt aus. Ohne solche Tricks wären die Verfahren deutlich langsamer.

Diagonalisieren
Hat eine $n \times n$-Matrix A genau n linear unabhängige Eigenvektoren, dann kann man sie als Spaltenvektoren zu einer $n \times n$-Matrix V zusammenfassen. Die Ähnlichkeitstransformation mit V erzeugt eine Diagonalmatrix D; in deren Haupdiagonale stehen die Eigenwerte.

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

Beispiel: die Matrix $A = \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix}$ vom Beispiel auf Seite 66 hat Eigenvektoren $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ und $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$.
Hier sind also

$$V = \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix} \text{ und } V^{-1} = \begin{bmatrix} 1/3 & 1 \\ 1/3 & 0 \end{bmatrix}, \quad V^{-1}AV = \begin{bmatrix} 1/3 & 1 \\ 1/3 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 3 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

Der MATLAB-Befehl zur Berechnung von Eigenwerten und -vektoren lautet `[V,D] = eig(A)`. Das im Hintergrund laufende Rechenverfahren führt iterativ Ähnlichkeitstransformationen durch, so dass sich die transformierte Matrix schrittweise einer Diagonalmatrix annähert. Als Endergebnis liefert es die Diagonalmatrix D und die Matrix V der Ähnlichkeitstransformation. Dann gilt: in der Hauptdiagonale von D stehen die Eigenwerten von A, die zugehörigen Eigenvektoren stehen in den Spalten von V.

Es kann sein, dass eine $n \times n$-Matrix A nicht n linear unabhängige Eigenvektoren hat. Das verkompliziert die Situation. Zum Glück treten – wie schon erwähnt – in wichtigen Anwendungen symmetrische Matrizen auf. Für sie gibt es immer eine Transformationsmatrix aus linear

[5] gleichbedeutend sind: nichtsinguläre Matrix, reguläre Matrix

unabhängigen Eigenvektoren. Als zusätzlicher Bonus gilt: die Transformationsmatrix lässt sich als *orthogonale* Matrix Q wählen.

Kurzfassung: Symmetrische Matrizen sind durch orthogonale Transformationen diagonalisierbar, $Q^T AQ = D$.

8.3. Vektoriteration

Die Berechnung von Eigenwerten und -vektoren ist in vielen Anwendungen von Bedeutung, beginnend mit der Mechanik (Eigenschwingungen) bis zur Ökonomie (Durchschnittspreise mit maximalem Gewinn) und zur Bewertung der Wichtigkeit von Internetseiten (PageRank einer Homepage als Eigenvektor der Google-Matrix, ein einfaches Beispiel findet sich in den Übungsaufgaben).

Dabei ist es nicht immer notwendig, alle Eigenwerte zu kennen, oft genügt es, den größten Eigenwert oder jenen mit größten Absolutbetrag zu berechnen.

Es kann in Ausnahmefällen mehrere Eigenwerte mit gleichem Absolutbetrag geben. Gibt es aber unter allen Eigenwerten *genau einen* mit maximalem Betrag, so nennen wir ihn *dominanten* Eigenwert. Ein solcher Eigenwert λ und ein zugehöriger Eigenvektor lassen sich durch ein einfaches iteratives Verfahren als Fixpunkt des Systems

$$\mathbf{x} = \frac{1}{\lambda} A\mathbf{x}$$

finden: die sogenannte Vektoriteration (englisch: *power iteration*).

Die Vektoriteration geht auf Richard von Mises[6] und Hilda Geiringer[7] zurück. Kurzfassung: Man nehme irgendeinen Vektor $\mathbf{x}^{(0)}$ und multipliziere ihn mit A. Das Ergebnis skaliere man geeignet durch Division mit einem Faktor $\lambda^{(1)}$ und nenne es $\mathbf{x}^{(1)}$. Dies Verfahren setze man fort, bis sich die $\mathbf{x}^{(k)}$ nur mehr vernachlässigbar ändern. Dann ist $\lambda^{(k)}$ der Eigenwert mit größtem Absolutbetrag und $\mathbf{x}^{(k)}$ ein zugehöriger Eigenvektor.

Vektoriteration für den betragsgrößten Eigenwert
(Potenzmethode nach v. Mises und Geiringer, *power iteration*)
Gegeben eine $n \times n$-Matrix A, ein Startvektor $\mathbf{x}^{(0)} \neq 0$ und ein fix gewähltes $i \in \{1, \dots, n\}$

Iteriere für $k = 1,2,\dots$
berechne $\mathbf{y}^{(k)} = A\mathbf{x}^{(k-1)}$
setze $\lambda^{(k)} = y_i^{(k)}$ (i-te Komponente von $\mathbf{y}^{(k)}$)
setze $\mathbf{x}^{(k)} = \mathbf{y}^{(k)}/\lambda^{(k)}$ (Skalierung)

[6] Richard von Mises, 1883 (Lemberg, Österreich; heute Lwiw, Ukraine) –1953 (Boston, USA). Mathematiker, studiert in Wien, ist Professor in Straßburg, Dresden, Berlin, Istanbul und ab 1939 an der Harvard University; richtungsweisende Arbeiten auf fast allen Gebieten der angewandten Mathematik, sowie der Strömungslehre, speziell unter Berücksichtigung des Flugzeugbaus (hält 1913 die erste Vorlesung über Motorflug!).

[7] Hilda Geiringer (Wien, 1893 – Santa Barbara, Kalifornien, 1973), studiert in Wien, arbeitet in Berlin als Assistentin von v. Mises am Inst. f. Angewandte Mathematik über Wahrscheinlichkeitstheorie und Statistik, emigriert 1933, lehrt in Brüssel, Istanbul und verschiedenen amerikanischen Universaitäten

Mit Papier und Bleistift ist die Multiplikation $A\mathbf{x}$ von tödlicher Kompliziertheit; Computer hingegen finden diese Methode recht einfach. Sie konvergiert, wenn es einen dominanten Eigenwert gibt und der Startvektor eine Komponente in Richtung des zugehörigen Eigenvektors hat.

Normaler Weise ist es irrelevant, welche Komponente von $\mathbf{x}^{(k)}$ zur Skalierung gewählt wird, außer der Eigenvektor hat genau für diese Komponente den Wert 0. Eine alternative Skalierunsvorschrift vermeidet die Festlegung auf eine bestimmte Komponente; das ist für Rechenprogramme einfacher:

Finde die betragsgrößte Komponente $y_i^{(k)}$ und setze $\lambda^{(k)} = y_i^{(k)}$.

Durch Anwenden der Potenzmethode auf die Matrix A^{-1} kann man den *betragskleinsten* Eigenwert ermitteln. Allgemein kann man jenen Eigenwert, welcher einem Wert μ am nächsten liegt, durch die Potenzmethode, angewandt auf $(A - \mu I)^{-1}$ finden (*inverse Iteration*). Dabei werden A^{-1} oder $(A - \mu I)^{-1}$ allerdings nicht explizit berechnet. Vielmehr wird der Iterationsschritt $\mathbf{y}^{(k)} = A^{-1}\mathbf{x}^{(k-1)}$ umformuliert zu $A\mathbf{y}^{(k)} = \mathbf{x}^{(k-1)}$ und $\mathbf{y}^{(k)}$ als Lösung dieses Systems bestimmt. Falls μ eine gute Näherung an ein λ_i ist, konvergiert das Verfahren rasch.

Für große, schwach besetzte symmetrische Matrizen gibt es ein besonders effizientes Verfahren, wenn nur der größte oder einige von den größten (oder kleinsten) Eigenwerten gesucht sind, das Lanczos-Verfahren. Im Kern stecken dort, wie bei der Potenzmethode, Iterationen der Form $\mathbf{y}^{(k)} = A\mathbf{x}^{(k-1)}$.

Auch dazu gibt es MATLAB-Befehle. Zum Beispiel liefert `d = eigs(A)` für schwach besetzte Matrizen die sechs betragsgrößten Eigenwerte und zugehörige Eigenvektoren.

8.4. Einfache Formen

Viele wichtige Verfahren zur Eigenwertberechnung bringen eine Matrix A durch eine Folge von Ähnlichkeitstransformationen $X^{-1}AX$ auf eine einfache Form und berechnen dann die Eigenwerte. (Die Eigenwerte ändern sich bei Ähnlichkeitstransformationen nicht.) Standard-Verfahren ist der sogenannte QR-Algorithmus.

Einfache Formen

- Diagonalmatrix
- Dreiecksmatrix
- Tridiagonalmatrix

Für Diagonal- und Dreiecksmatrizen sind die Eigenwerte genau die Diagonalelemente.

Für eine Tridiagonalmatrix lässt sich das charakteristische Polynom leicht auswerten, ohne dass man es explizit anschreiben müsste. Für eine symmetrische Tridiagonalmatrix T,

$$T = \begin{bmatrix} a_1 & b_1 & & \cdots & 0 \\ b_1 & a_2 & b_2 & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & \ddots & b_{n-1} \\ 0 & \cdots & & b_{n-1} & a_n \end{bmatrix},$$

zeigt Entwickeln der Determinante $\det(T - xI)$, dass der Wert p(x) des charakteristischen Polynoms rekursiv so bestimmt werden kann:

setze $p_0(x) = 1, p_{-1}(x) = 0$.
Für $k = 1, \ldots, n$
$p_k(x) = (a_k - x)p_{k-1}(x) - b_{k-1}^2 p_{k-2}(x)$
Ergebnis $p(x) = p_n(x)$.

Ein Verfahren zur Nullstellenbestimmung, zum Beispiel Intervallhalbierung, liefert mit dieser Methode die Eigenwerte als Nullstellen des charakteristischen Polynoms.