§ 31 Numerische Hauptachsentransformation und das QR-Verfahren

Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix.
Um die Hauptachsentransformation zu berechnen sind wir bisher wie folgt vorgegangen:

1) Bestimmen eines Eigenwertes λ von A
2) Berechnen eines zugehörigen (normierten) EV v zu λ
3) Berechne $H = v^\perp$ und wenden das Verfahren auf $A|_H$ rekursiv an.

Wie können wir λ numerisch bestimmen für den Fall, dass die Berechnung von χ_A schon sehr aufwendig ist?

Idee: Eigenvektor Iteration
Sei $A \in \mathbb{R}^{n \times n}$ (nicht notwendig) symmetrische Matrix.
Wir starten mit einem beliebigen normierten Vektor w_0 und berechnen rekursiv

$$w_{k+1} = \frac{A w_k}{\| A w_k \|}$$ w_k besser kein EV zum EW $\lambda = 0$

Hoffnung: $\lim_{k \to \infty} w_k = w$ existiert und ist ein EV von A

Satz: Sei $A \overset{\in \mathbb{R}^{n \times n}}{}$ eine symmetrische Matrix
$\lambda_1, \dots, \lambda_n \in \mathbb{R}$ die EW und $\lambda_{\max} = \max\{|\lambda_i|\}$
Sei $W = \mathrm{Eig}(A, \lambda) \oplus \mathrm{Eig}(A, -\lambda)$

Gilt $w_0 \notin W^\perp$ dann konvergiert die Folge (w_{2k}) gegen einen EV w von A^2 zum EW λ^2.

Sind alle λ_i mit $|\lambda_i| = \lambda$ positiv, dann konvergiert schon (w_k) gegen w und w ist ein EV $\underbrace{\ }_{\text{von } A}$ zum Eigenwert λ.

Beweis: Es sei $v_1, \dots, v_n \in \mathbb{R}^n$ eine Orthonormalbasis aus EV von A zu EW $\lambda_1, \dots, \lambda_n$

etwa $w_0 = \sum_{j=1}^{n} \mu_j v_j$ mit $\sum_{j=1}^{n} \mu_j^2 = 1$, da w_0 normiert ist.

$$\Rightarrow \quad A w_0 = \sum_{j=1}^{n} \mu_j A v_j = \sum_{j=1}^{n} \mu_j \lambda_j v_j$$
$$= \lambda \cdot \sum_{j=1}^{n} \mu_j \left(\frac{\lambda_j}{\lambda}\right) \cdot v_j$$

~~und w_1 $\frac{A w_0}{\|A w_0\|} = w_1 = \sum_{j=1}^{n} \mu_j^{(1)} v_j$~~

Wir nehmen an, dass $|\lambda_1| = \lambda_{max}$ und dass $\mu_1 \neq 0$. Letzteres ist erfüllt nach Umnummerierung, da $w_0 \notin W^\perp$.

$$\frac{A w_0}{\|A w_0\|} = \sum_{j=1}^{n} \mu_j^{(1)} v_j \quad \text{und} \quad \frac{\mu_j^{(1)}}{\mu_1} = \frac{\lambda_j}{\lambda_{max}} \frac{\mu_j}{\mu_1}$$

Rekursiv $$\frac{\mu_j^{(k)}}{\mu_1^{(k)}} = \left(\frac{\lambda_j}{\lambda_{max}}\right)^k \frac{\mu_j}{\mu_1}$$

$\Rightarrow$ Der relative Anteil von w_k in die Richtung v_j mit $|\lambda_j| \neq \lambda_{max}$ wird immer kleiner und der GW

$$w = \lim_{k \to \infty} w_{2k} = \sum_{\substack{j \text{ mit} \\ |\lambda_j| = \lambda_{max}}} \mu_j v_j$$

~~$\Rightarrow A^2 w = \sum_{\substack{j \\ |\lambda_j| = \lambda_{max}}} (-1)^{2k} \mu_j v_j = \sum_{\substack{j \\ |\lambda_j| = \lambda_{max}}} \mu_j v_j$~~

~~also die orthogonale Projektion von w_0 auf W~~

$$A^2 w = \sum_{\substack{j \\ |\lambda_j| = \lambda_{max}}} \lambda_j^2 \mu_j v_j = \lambda_{max}^2 \sum_{\substack{j \\ |\lambda_j| = \lambda_{max}}} \mu_j v_j = \lambda_{max}^2 \cdot w$$

$w \in W$ ist die orthogonale Projektion von w_0 auf W und ein EV von A^2 zum EW λ_{max}^2.

Sind alle λ_j mit $|\lambda_j| = \lambda_{max}$ positiv, dann konvergiert schon (w_k) mit Grenzwert ein $w_k \to w$ ein EV zum EW λ_{max}. Ist dies nicht der Fall, dann spannt w und Aw einen 2-dimensionalen Raum auf und A auf diesem 2-dim. Raum zu diagonalisieren ist einfach.

$$W = W^+ \oplus W^- \in \mathrm{Eig}(A, \lambda_{max}) \oplus \mathrm{Eig}(A, -\lambda_{max})$$

$$Aw = \lambda_{max} w^+ - \lambda_{max} w^-.$$

Will man alle EW von $A \in \mathbb{R}^{n \times n}$ symmetrisch gleichzeitig berechnen, so bietet sich das QR-Verfahren an.

QR-Verfahren:

Input: $A \in \mathbb{R}^{n \times n}$ symmetrische Matrix

Output: Folge von Matrizen A_k

1) Setze $A_1 = A$
2) A_k gegeben, berechnen wir die QR-Zerlegung
   $A_k = Q_k R_k$

von A_k mit positiven Einträgen auf der Diagonalen von R_k

Dann sei $A_{k+1} = R_k Q_k$

Satz: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit EW $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > 0$.
Dann konvergiert die Folge (A_k) gegen eine Diagonalmatrix mit Einträgen $\lambda_1, \dots, \lambda_n$ (In der Regel in dieser Reihenfolge)
Ist dies die Reihenfolge, dann konvergieren die Nicht-Diagonalelemente gegen 0 mit Geschwindigkeit

$$a_{ij}^{(k)} = a_{ji}^{(k)} \in O\left(\left|\frac{\lambda_i}{\lambda_j}\right|^k\right) \quad i > j$$

Bemerkung

1) $A_{k+1} = R_k Q_k = \underbrace{{}^tQ_k Q_k}_{= E,\ \text{da } Q_k \text{ orthogonal}} R_k Q_k = {}^tQ_k A_k Q_k$

Also alle Matrizen A_k sind ähnlich zu $A = A_1$

2) Sind zwei der EW λ_j betragsmäßig gleich, etwa $\lambda_j = \lambda_{j+1}$, dann gilt auch hier noch Konvergenz.

Gilt $\lambda_j = -\lambda_{j+1}$ dann kann ein 2x2-Block stehen bleiben.

Beweis: Wir zeigen zunächst:
für die k-te Potenz von A:

$$A^k = \underbrace{Q_1 Q_2 \dots Q_k}_{\substack{P_k \\ \text{(orthogonale} \\ \text{Matrix)}}} \underbrace{R_k R_{k-1} \dots R_1}_{U_k \text{ (obere Dreiecksmatrix)}}$$

mit Induktion nach k

$$A_1 = Q_1 R_1$$

$k \to k+1$

$$A_{k+1} \overset{\text{QR-Zerlegung}}{=} Q_{k+1} R_{k+1} \overset{\text{Def.}}{=} R_k Q_k = {}^tQ_k \underbrace{Q_k R_k}_{A_k} Q_k = {}^tQ_k A_k Q_k$$

$$= {}^tQ_k\, {}^tQ_{k-1} \dots {}^tQ_1 A Q_1 \dots Q_k = {}^tP_k A P_k$$

Für die Potenzen ergibt sich:

$$A^{k+1} = A \cdot A^k \overset{\text{IV}}{=} A P_k U_k = P_k\, {}^tP_k A P_k U_k$$

$$= P_k A_{k+1} U_k = P_k Q_{k+1} R_{k+1} U_k = P_{k+1} U_{k+1}$$

Da alle P_k orthogonal sind und alle U_k obere Dreiecksmatrizen mit positiven Diagonalelementen folgt $A^k = P_k U_k$ ist die QR-Zerlegung von A^k.
Ferner $\Lambda = \operatorname{diag}(\lambda_1, , \lambda_n)$ Diagonalmatrix $S \in O(n)$
$= {}^tSAS \quad , \quad A = SA\,{}^tS$

$\Rightarrow A^k = S \Lambda^k\, {}^tS$

Wir nehmen nun an, dass ${}^tS$ eine LR-Zerlegung hat (Brauchen wir eine Permutationsmatrix in der LR-Zerlegung, dann permutieren die Einträge in der Folge A_k entsprechend)

Also ${}^tS = LR$ mit L untere Δ-Matrix, R obere Δ-Matrix

$$A^k = S \cdot \Lambda^k L R$$

$$= S(\Lambda^k L \Lambda^{-k}) \Lambda^k R$$

Da $L = (l_{ij})$ eine untere Dreiecksmatrix ist, gilt

$$(\Lambda^k L \Lambda^{-k})_{ij} = l_{ij} \left(\frac{\lambda_i^{\,k}}{\lambda_j^{\,k}} \right) \text{ für } i > j$$

$$|\lambda_j| > |\lambda_i|$$

dieser Ausdruck konvergiert also gegen 0 und damit $\Lambda^k L \Lambda^{-k} = E + F_k$ mit $F_k \xrightarrow[k \to \infty]{} 0$.

Eingesetzt liefert dies

$$A^k = S(E + F_k) \Lambda^k R$$

Sei nun $E + F_k = \tilde{Q}_k \tilde{R}_k$ die QR-Zerlegung von $E + F_k$ mit $\tilde{Q}_k \to E$ $\tilde{R}_k \to E$

$$A^k = S \tilde{Q}_k \tilde{R}_k \Lambda^k R = P_k U_k$$

Eindeutigkeit der QR-Zerlegung

$$P_k = S Q_k D_k, \; U_k = D_k \tilde{R}_k \Lambda^k R$$

wobei $D_k = \operatorname{diag}\left((\operatorname{sign} \lambda_1)^k, \ldots, (\operatorname{sign} \lambda_n)^k \right)$

$$\Rightarrow Q_k = {}^t P_{k-1} P_k = D_{k-1} \underbrace{{}^t\tilde{Q}_{k-1} \cdot {}^t S S \tilde{Q}_k}_{\to E} D_k \Longrightarrow D_1 = D_{k-1} D_k$$

$$R_k = U_k U_{k-1}^{-1} = D_k \tilde{R}_k \Lambda^k R R^{-1} \Lambda^{-(k-1)} \tilde{R}_{k-1}^{-1} D_{k-1}^{-1}$$

$$= D_k \tilde{R}_k \Lambda \tilde{R}_{k-1}^{-1} D_{k-1}$$

$$\Rightarrow R_{2k} = \underset{\substack{\shortparallel \\ E}}{D_{2k}} \underbrace{\tilde{R}_{2k}}_{\to E} \Lambda \underbrace{\tilde{R}_{2k-1}}_{E} \underset{\substack{\shortparallel \\ D_1}}{D_{2k-1}} \to \Lambda D_1$$

$$R_{2k+1} = D_1 \tilde{R}_{2k+1} \Lambda \tilde{R}_{2k}^{-1} \to D_1 A = \Lambda D_1$$

$$\Rightarrow A_k = Q_k R_k \to D_1 D_1 \Lambda = \Lambda \qquad \square$$