

# Lineare Algebra II - Eigenwertprobleme

-1-

In der heutigen Vorlesung wollen wir uns noch einmal dem Themenblock der numerischen Verfahren der Linearen Algebra zuwenden. Nachdem wir vergangene Woche die Lösung von linearen Gleichungssystemen  $A\vec{x} = \vec{b}$  anhand des Gauß'schen Eliminationsverfahrens diskutiert haben, wollen wir uns heute der Berechnung von Eigenwerten und Eigenvektoren einer gegebenen Matrix  $A$  widmen.

Wir suchen also Lösungen der Gleichung

$$A\vec{x} = \lambda\vec{x}$$

wobei  $N \times N$  Matrix  $A$  gegeben sei.  
gesucht sind Eigenvektoren  $\vec{x}_i$  mit Eigenwerten  $\lambda_i$

Aus der Linearen Algebra kennen wir die Bedingung für die Existenz nicht-trivialer Eigenvektoren (d.h.  $\vec{x} \neq 0$ ):

$$\det(A - \lambda \mathbb{1}) = 0$$

Beispiel:

$$A = \begin{pmatrix} 1 & 2 \\ 4 & -1 \end{pmatrix} \rightarrow A - \lambda \mathbb{1} = \begin{pmatrix} 1-\lambda & 2 \\ 4 & -1-\lambda \end{pmatrix}$$

$$\rightarrow \det(A - \lambda \mathbb{1}) = \lambda^2 - 1 - 8 = \lambda^2 - 9 = 0 = f(\lambda)$$

charakteristisches Polynom

$$\rightarrow \lambda_{1/2} = \pm 3$$

Gesucht sind also die Nullstellen des charakteristischen Polynoms.

Im allgemeinen ist eine derartige Nullstellensuche zur Bestimmung der Eigenwerte  $\lambda_i$  allerdings ineffizient, da die Nullstellen (im allgemeinen) sehr dicht liegen können.

Ist die Bedingung  $\det(A - \lambda \mathbb{1}) = 0$  erfüllt, gelten die folgenden Zusammenhänge:

-2-

- orthogonale Matrizen

Orthogonale Matrizen sind reelle  $N \times N$  Matrizen mit

$$A \cdot A^t = A^t \cdot A = \mathbb{1} \quad (\text{Transposition } a_{ij}^t = a_{ji})$$

Zu jeder orthogonalen Matrix  $A$  gibt es eine unitäre Matrix  $S$ , so daß  $S^{-1}AS = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$  ist, und  $S$  aus den Eigenvektoren besteht.

- unitäre Matrizen

Unitäre Matrizen sind komplexe  $N \times N$  Matrizen mit

$$A \cdot A^* = A^* A = \mathbb{1} \quad (\text{Adjungation } a_{ij}^* = \overline{a_{ji}})$$

Wie für die orthogonalen Matrizen gibt es für jede unitäre Matrix  $A$  eine unitäre Matrix  $S$ , so daß

$$\underline{S^{-1}AS = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)}$$

und  $S$  aus den Eigenvektoren besteht.

- Symmetrische Matrizen

Für eine symmetrische Matrix gilt  $A^t = A$ .

Ist eine symmetrische Matrix zusätzlich noch reell, so ist sie normal, und damit diagonalisierbar. Die Eigenvektoren sind orthogonal.

- hermitesche Matrizen

Für eine hermitesche Matrix gilt  $A^* = A$  (oder  $A^t = \overline{A}$ ).

Eine hermitesche Matrix läßt sich diagonalisieren. Alle Eigenwerte sind reell.

Kommentar: Für die oberen Dreiecksmatrizen, welche wir in der letzten Vorlesung besprochen haben, lassen sich die Eigenwerte sofort ablesen: Es sind die Einträge auf der Hauptdiagonalen. Allerdings gilt für das Gauß'sche Eliminationsverfahren nicht, daß die Eigenwerte der Matrizen  $A^{(0)}, A^{(1)}, \dots, A^{(n-1)}$  übereinstimmen, denn das Eliminationsverfahren beschreibt keine unitäre Transformation im obigen Sinne!

Das Lanczos-Verfahren ist ein iterativer Algorithmus zur Berechnung der Eigenwerte und -vektoren einer quadratischen, hermiteschen Matrix. Das Verfahren ist eine Adaption der sogenannten Power-Methode zur Berechnung des größten Eigenwerts (und entsprechendem Eigenvektor). Es ist besonders geeignet für sehr große, dünnbesetzte Matrizen.

### Power-Methode

Die Power-Methode ist ein iteratives Verfahren zur Bestimmung des größten (dominanten) Eigenwerts einer quadratischen  $n \times n$  Matrix  $A$  mit  $n$  linear unabhängigen Eigenvektoren und der Annahme eines dominanten Eigenwerts. Die Methode ist sehr speichereffizient, da nur zwei Vektoren gespeichert werden müssen.

Sei  $\vec{x}_0$  ein beliebiger Vektor, welchen wir in der Basis der  $n$  linear unabhängigen Eigenvektoren  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$  schreiben können als

$$\vec{x}_0 = c_1 \cdot \vec{v}_1 + c_2 \cdot \vec{v}_2 + \dots + c_n \cdot \vec{v}_n$$

Wobei die Koeffizienten  $\{c_1, c_2, \dots, c_n\}$  zunächst unbekannt sind.

Multipliziere diese Gleichung mit der Matrix  $A$

$$A \vec{x}_0 = c_1 \cdot \lambda_1 \cdot \vec{v}_1 + c_2 \cdot \lambda_2 \cdot \vec{v}_2 + \dots + c_n \cdot \lambda_n \cdot \vec{v}_n$$

$\downarrow$  Eigenwert

$$A^2 \vec{x}_0 = c_1 \cdot \lambda_1^2 \cdot \vec{v}_1 + c_2 \cdot \lambda_2^2 \cdot \vec{v}_2 + \dots + c_n \cdot \lambda_n^2 \cdot \vec{v}_n$$

$\vdots$

$$A^m \vec{x}_0 = c_1 \cdot \lambda_1^m \cdot \vec{v}_1 + c_2 \cdot \lambda_2^m \cdot \vec{v}_2 + \dots + c_n \cdot \lambda_n^m \cdot \vec{v}_n \quad (\text{I})$$

$$A^{m+1} \vec{x}_0 = c_1 \cdot \lambda_1^{m+1} \cdot \vec{v}_1 + c_2 \cdot \lambda_2^{m+1} \cdot \vec{v}_2 + \dots + c_n \cdot \lambda_n^{m+1} \cdot \vec{v}_n \quad (\text{II})$$

Dividiere Gleichung (I) durch  $\lambda_1^m$ :

-4-

$$\frac{1}{\lambda_1^m} \cdot A^m \vec{x}_0 = c_1 \cdot \vec{v}_1 + c_2 \cdot \left(\frac{\lambda_2}{\lambda_1}\right)^m \vec{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1}\right)^m \vec{v}_n \quad (I)$$

Für  $\lambda_1 \gg \lambda_2, \lambda_3, \dots, \lambda_n$  (dominanter Eigenwert) gehen die Brüche  $(\lambda_i/\lambda_1)^m$  gegen Null für  $m \rightarrow \infty$ . Damit erhalten wir

$$\frac{1}{\lambda_1^m} A^m \vec{x}_0 = c_1 \cdot \vec{v}_1$$

$$\frac{1}{\lambda_1^{m+1}} A^{m+1} \vec{x}_0 = c_1 \cdot \vec{v}_1$$

Bilde nun das Skalarprodukt mit einem beliebigen, zu  $\vec{v}_1$  nicht-orthogonalem Vektor  $\vec{y}$

$$\frac{1}{\lambda_1^m} A^m \vec{x}_0 \cdot \vec{y} = c_1 \cdot \vec{v}_1 \cdot \vec{y} = \frac{1}{\lambda_1^{m+1}} A^{m+1} \vec{x}_0 \cdot \vec{y}$$

$$\rightarrow \lambda_1 = \frac{A^{m+1} \vec{x}_0 \cdot \vec{y}}{A^m \vec{x}_0 \cdot \vec{y}}$$

$$\vec{x}_1 = A^m \vec{x}_0$$

### Krylov - Raum

Wir können die Power-Methode, anstelle einer Sequenz von Matrizen-Multiplikationen, auch als eine Sequenz von Vektoren paraphrasieren

$$\vec{u}_1 = \vec{x}_0, \vec{u}_2 = A \vec{u}_0, \vec{u}_3 = A^2 \vec{u}_0, \dots, \vec{u}_m = A^{m-1} \vec{u}_0$$

Diese Vektoren spannen den sogenannten Krylov-Unterraum  $K_m$  auf

$$K_m = \text{Span} \{ \vec{u}_1, \vec{u}_2, \vec{u}_3, \dots, \vec{u}_m \}$$

In jedem Schritt des Lanczos-Verfahrens werden wir nun eine orthogonale Basis für diesen Krylov-Unterraum konstruieren

$$\text{basis} = \{ \vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \}$$

Diese Vektoren sind nicht identisch zu den Eigenvektoren von A.

Aus diesen Vektoren können wir spaltenweise eine Transformationsmatrix  $V_m$  konstruieren der Form

$$V_m = \underbrace{\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \end{pmatrix}}_{m \text{ Spalten}} \left. \vphantom{\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \end{pmatrix}} \right\} n \text{ Zeilen}$$

welche unsere ursprüngliche  $n \times n$  Matrix A in eine tridiagonale Form bringt

$$T_{mm} = V_m^* A V_m$$

$m \times m \quad m \times n \quad n \times n \quad n \times m$

$$m \begin{matrix} \square \\ \square \\ \square \end{matrix}^m = \begin{matrix} \square \\ \square \\ \square \end{matrix}^n_m \begin{matrix} \square \\ \square \\ \square \end{matrix}^n_n \begin{matrix} \square \\ \square \\ \square \end{matrix}^m_n$$

mit

$$T_{mm} = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \beta_3 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_m \\ & & & \beta_m & \alpha_m \end{pmatrix}$$

Für  $m=n$  sind die Matrizen  $T_{nn}$  und A ähnlich.

### Lanczos-Algorithmus

Der eigentliche Lanczos-Algorithmus zur Bestimmung der Vektoren  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$  und der Einträge  $\alpha_1, \alpha_2, \dots, \alpha_m$  sowie  $\beta_2, \beta_3, \dots, \beta_m$  der tridiagonalen Matrix sieht nun wie folgt aus

$$\text{Initialisierung} \left[ \begin{array}{l} \beta_1 = 0 \\ \vec{v}_0 = 0 \\ \vec{v}_1 = \text{zufälliger Vektor der Dimension } n, \text{ normiert} \end{array} \right.$$

Lanczos-Iteration

für  $j = 1, 2, \dots, M$  tue {

$$\vec{\omega}_j = A \cdot \vec{v}_j - \beta_j \vec{v}_{j-1}$$

$$\alpha_j = \vec{v}_j^T \cdot \vec{\omega}_j$$

$$\vec{\omega}_j = \vec{\omega}_j - \alpha_j \cdot \vec{v}_j$$

$$\beta_{j+1} = \|\vec{\omega}_j\|$$

$$\vec{v}_{j+1} = \vec{\omega}_j / \beta_{j+1} \}$$

Wir benötigen nur 3 Vektoren der Größe  $n$  im Speicher!

Die so aufgestellte tridiagonale Matrix  $T_{mm}$  kann einfach diagonalisiert werden, was uns die Eigenwerte

$$\{\tau_1, \tau_2, \dots, \tau_m\}$$

gibt. Diese sind extrem gute Annäherungen an die Eigenwerte von  $A$ .  
 (Die Konvergenz ist sehr schnell und es benötigt  $m \ll n$  Schritte.)

### Vorteile

- Näherung des extremen Eigenwerts schon nach kleiner Iterationszahl sehr genau
- $\dim_{\text{Lanczos-Matrix}} \ll \dim_A$ , d.h.  $m \ll n$
- Speicher-effizient

### Nachteile

- Konvergenz zu höheren Zuständen kann unregelmäßig sein
- "Geister"-Eigenwerte: Verlust an Orthogonalität der Lanczos-Vektoren aufgrund der endlichen numerischen Präzision.