Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Eigenwerte, Eigenräume

Die Themen heute sind

- Diagonalisierbarkeitskriterium
- Jordansche Normalform und Nilpotente Matrizen
- Google page rank und Stochastische Matrizen
- Ergodensatz
- Gerschgorin-Kreise

Am Mittwoch hatten wir ein Diagonalisierbarkeitsktriterium formuliert. Heute werden wir es beweisen.

Diagonalisierbarkeitskriterium

Satz. *Eine quadratische Matrix $A \in K^{n\times n}$ ist diagonalisierbar genau dann, wenn die beiden folgenden Bedingungen erfüllt sind:*

1. $\chi_A(t) \in K[t]$ *zerfällt (über K) in Linearfaktoren,*
2. *für jede Nullstelle λ von $\chi_A(t)$ gilt:*

$$m(\chi_A(t), \lambda) = \dim \operatorname{Eig}(A, \lambda).$$

Manchmal nennt man $m(\chi_A(t), \lambda)$ auch **algebraische Vielfachheit** des Eigenwertes λ und $\dim \operatorname{Eig}(A, \lambda)$ seine **geometrische Vielfachheit**.
Mit dieser Terminologie heißt die zweite Bedingung, dass algebraische und geometrische Vielfachheit für alle Eigenwerte übereinstimmen sollen.

Lemma. $A \in K^{n\times n}$. *Seien* $\lambda_1, \dots, \lambda_r$ *paarweise verschiedene Eigenwerte von* A *und* $v_1, \dots, v_r$ *zugehörige Eigenvektoren. Dann sind* $v_1, \dots, v_r$ *linear unabhängig. Genauer gilt: Die Summe*

$$\mathrm{Eig}(A, \lambda_1) \oplus \mathrm{Eig}(A, \lambda_2) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_r) \subseteq K^n$$

ist direkt, also nach Definition:

$$\mathrm{Eig}(A, \lambda_j) \cap \sum_{\substack{\ell=i \\ \ell\neq j}}^{r} \mathrm{Eig}(A, \lambda_\ell) = 0 \qquad \forall\, j \in \{1, \dots, r\}.$$

Beweis. Induktion nach r. Der Fall $r = 1$ ist trivial.
Nehmen wir also an, dass:

$$v_r \in \mathrm{Eig}(A, \lambda_r) \cap \left(\sum_{j=1}^{r-1} \mathrm{Eig}(A, \lambda_j) \right),$$

etwa

$$v_r = w_1 + \cdots + w_{r-1} \qquad \text{für gewisse} \;\; w_j \in \mathrm{Eig}(A, \lambda_j).$$

Da $w_j \in \mathrm{Eig}(A, \lambda_j)$, folgt

$$\lambda_r v_r = A v_r = A w_1 + \cdots + A w_{r-1} = \lambda_1 w_1 + \cdots + \lambda_{r-1} w_{r-1}.$$

Subtraktion des λ_r-fachen von $v_r = w_1 + \cdots + w_{r-1}$ gibt:

$$0 = (\lambda_1 - \lambda_r) w_1 + \cdots + (\lambda_{r-1} - \lambda_r) w_{r-1}.$$

Nach der Induktionsvoraussetzung ist die Summe $\mathrm{Eig}(A, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_{r-1}) \subset K^n$ direkt. Also folgt:

$$(\lambda_j - \lambda_r) w_j = 0 \in \mathrm{Eig}(A, \lambda_j) \quad \forall j.$$

Es gilt aber $\lambda_j \neq \lambda_r$ nach Voraussetzung und somit $w_j = 0 \; \forall j$ und schließlich: $v_r = w_1 + \cdots + w_{r-1} = 0$. $\square$

Beweis des Diagonalisierbarkeitskriteriums. Notwendigkeit der Bedingungen: Ist A diagonalisierbar, dann ist

$$\chi_A(t) = \chi_D(t) = \prod_{j=1}^{n} (\lambda_j - t)$$

zerfallend und

$$\begin{aligned} \dim \mathrm{Eig}(A, \lambda_j) &= \dim \mathrm{Eig}(D, \lambda_j) = \dim \ker \begin{pmatrix} \lambda_1 - \lambda_j & & 0 \\ & \ddots & \\ 0 & & \lambda_n - \lambda_j \end{pmatrix} \\ &= |\{i \in \{1, 2, \ldots, n\} \mid \lambda_i = \lambda_j\}| = m(\chi_A(t), \lambda_j). \end{aligned}$$

Wir haben also noch zu zeigen, dass die Bedingungen auch hinreichend sind: Es sei dazu

$$\chi_A(t) = \prod_{i=1}^{n} (\lambda_i - t) = \prod_{j=1}^{r} (\lambda_j - t)^{m_j},$$

wobei $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte bezeichnen, d.h. $\sum_{j=1}^{r} m_j = n$. Nach dem Lemma gilt:

$$\mathrm{Eig}(A, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_r) \subset K^n$$

Nach der 2. Bedingung ist

$$\sum_{j=1}^{r} \dim \mathrm{Eig}(A, \lambda_j) = \sum_{j=1}^{r} m_j = \deg \chi_A(t) = n = \dim K^n.$$

Dies zeigt

$$\mathrm{Eig}(A, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(A, \lambda_r) = K^n.$$

Wir können eine Basis $\mathcal{A} = \{v_1, \ldots, v_n\}$ von Eigenvektoren wählen, wobei $\{v_1, \ldots, v_{m_1}\}$ eine Basis von $Eig(A, \lambda_1)$ ist, die nächsten m_2 Vektoren eine Basis von $\mathrm{Eig}(A, \lambda_2)$, ..., und die letzten m_r Vektoren eine Basis von $\mathrm{Eig}(A, \lambda_r)$. Bzgl. dieser Basis hat $D = M_{\mathcal{A}}^{\mathcal{A}}(A)$ Diagonalgestalt:

$$D = SAS^{-1},$$

wobei die Matrix S^{-1} die Vekoren $v_1, \ldots v_n$ als Spalten hat. $\square$

Ende Teil 1

Bemerkung. Es gilt immer

$$\dim \mathrm{Eig}(A,\lambda) \leq m(\chi_A(t),\lambda)$$

In der Tat, wenn wir eine Basis $\{v_1,\dots,v_d\}$ von $\mathrm{Eig}(A,\lambda)$ zu einer Basis $\mathcal{B}=\{v_1,\dots,v_n\}$ von K^n ergänzen, dann gilt

$$TAT^{-1} = \begin{pmatrix} D_1 & C \\ 0 & A_2 \end{pmatrix} \text{ für } T = M^{\mathcal{E}}_{\mathcal{B}}(\mathrm{id}_{K^n}),$$

wobei $D_1 = \lambda \mathrm{id}_{K^d}$ eine Diagonalmatrix ist. Also

$$\chi_A(t) = (\lambda - t)^d \chi_{A_2}(t).$$

□

Korollar. *Sei $A \in K^{n\times n}$. Hat $\chi_A(t)$ genau n verschiedene Nullstellen $\lambda_1,\dots,\lambda_n \in K$, dann ist A diagonalisierbar.*

Beweis. Für jedes λ_i gilt

$$1 \leq \dim \mathrm{Eig}(A,\lambda_i) \leq m(\chi_A(t),\lambda_i) \leq 1,$$

da alle Nullstellen einfach sind. □

Fundamentalsatz der Algebra

Für den Körper $\mathbb{C} = \{a + ib \mid a, b \in \mathbb{R}\}$ der komplexen Zahlen (also $i = \sqrt{-1}$ und $i^2 = -1$) ist die erste Bedingung immer erfüllt:

Satz. *Jedes Polynom $P(t) \in \mathbb{C}[t]$ vom Grad $d \geq 1$ hat eine Nullstelle (in $\mathbb{C}$).*

Korollar. *Jedes Polynom $P(t) \in \mathbb{C}[t]$ zerfällt in Linearfaktoren.*

Die zweite Bedingung ist aber nicht immer erfüllt. Zum Beispiel für

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

ist

$$\dim \mathrm{Eig}(A, 0) = 1 < m(\chi_A(t), 0) = 2.$$

Jordansche Normalform (ohne Beweis)

Satz. *Sei* $A \in \mathbb{C}^{n\times n}$ *eine quadratische Matrix mit komplexen Einträgen. Dann existieren* $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$, $k_1, \ldots, k_r \in \mathbb{N}$ *und* $S \in \mathrm{GL}(n, \mathbb{C})$, *so dass:*

$$SAS^{-1} = J = \begin{pmatrix} J(\lambda_1, k_1) & 0 & \cdots & 0 \\ 0 & J(\lambda_2, k_2) & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & J(\lambda_r, k_r) \end{pmatrix},$$

wobei die Eigenwerte $\lambda_1, \ldots, \lambda_r$ *nicht notwendig paarweise verschieden sind, und die*

$$J(\lambda_j, k_j) := \begin{pmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{pmatrix} \in \mathbb{C}^{k_j \times k_j}.$$

sogenannte **Jordankästchen** *der Größe* k_j *zum Eigenwert* λ_j *bezeichnen.*

Nilpotente Matrizen

Definition. Eine **strikte obere Dreiecksmatrix** ist eine obere Dreiecksmatrix mit Nullen auf der Diagonalen.

$$\begin{pmatrix} 0 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ 0 & & & 0 \end{pmatrix}$$

Eine Matrix $N \in K^{n \times n}$ heißt **nilpotent**, wenn ein $k \in \mathbb{N}$ existiert mit $N^k = 0$.

Jede strikte obere Deiecksmatrix $A \in K^{n \times n}$ ist nilpotent: $A^n = 0$.

Satz. *Sei $A \in K^{n\times n}$ eine Matrix mit charakteristischem Polynom*

$$\chi_A(t) = (-1)^n t^n.$$

Dann ist A zu einer strikten oberen Dreiecksmatrix konjugiert.

Beweis. $\lambda = 0$ ist ein Eigenwert von A. Wir ergänzen einen Eigenvektor $v = v_1$ zu einer Basis $\mathcal{B} = \{v_1, \dots v_n\}$ von K^n. Ist $T^{-1} = (v_1 \dots v_n)$, dann gilt

$$TAT^{-1} = \begin{pmatrix} 0 & C \\ 0 & A_2 \end{pmatrix}$$

mit $A_2 \in K^{n-1\times n-1}$. Es gilt: $\chi_A(t) = (-t)\chi_{A_2}(t)$, folglich $\chi_{A_2}(t) = (-t)^{n-1}$. Wir können also mit Induktion nach n voraussetzen, dass A_2 zu einer strikten oberen Dreiecksmatrix konjugiert ist. A ist dies dann ebenfalls. $\square$

Ein wesentlicher Schritt im Beweis des Satzes über die Jordansche Normalform ist es zu zeigen, dass Konjugationsklassen von nilpotenten Matrizen $N \in K^{n\times n}$ in Bijektion zur Menge der Partitionen von n stehen. Ende Teil 2

Google page rank und Randomsurfer

Die grundlegende Idee von Google's page rank ist es, allein aus der Struktur des Internets (und nicht den Inhalten!) die Wichtigkeit der einzelnen Seiten zu ermitteln. Ein Model hierbei ist, einen Randomsurfer zu beobachten, wie er durch das Internet läuft und die Wichtigkeit einer Seiten proportional zur Häufigkeit mit sie besucht wird, zu setzen.

Wir berechnen dies mit einem Wahrscheinlichkeitsmodell:
Die Hyperlinks auf den Seiten geben einen gerichteten Graphen. Befindet sich der Randomsurfer auf Seite j und enthält diese ℓ Links auf andere Seiten, so kann man diese mit einem Klick erreichen. Da keine inhaltliche Information verwendet werden soll, ist die Wahrscheinlichkeit, im nächsten Schritt auf einer dieser ℓ anderen Seiten zu landen, jeweils gleich groß, nämlich $\frac{1}{\ell}$.

Wir setzen

$$p_{ij} := \begin{cases} 0, & \text{falls es keinen Link von der Seite } j \text{ zur Seite } i \text{ gibt} \\ \frac{1}{\ell}, & \text{sonst, wobei } \ell \text{ die Anzahl der Links auf Seite } j \text{ ist.} \end{cases}$$

Dann ist

$$P = (p_{ij}) \in \mathbb{R}^{n \times n},$$

wobei n die Anzahl der Seiten des Netzes ist. P ist eine sogenannte stochastische Matrix.

Definition. Eine $n \times n$ Matrix $P = (p_{ij})$ mit Einträgen in $[0, 1] \subset \mathbb{R}$, die

$$\sum_{i=1}^{n} p_{ij} = 1 \quad \forall j \in \{1, \ldots, n\}$$

erfüllt, heißt (spalten)**stochastische Matrix**.

Ein Netz mit 4 Seiten

Beispiel.

$$P = \begin{pmatrix} 0 & 0 & 1/2 & 0 \\ 1/3 & 0 & 0 & 1 \\ 1/3 & 1/2 & 0 & 0 \\ 1/3 & 1/2 & 1/2 & 0 \end{pmatrix}$$

Berechnen $w_k = P^k w$ für $k = 0, 1, 2, 20, 100$.

$$\begin{pmatrix} .25 \\ .25 \\ .25 \\ .25 \end{pmatrix}, \begin{pmatrix} .125 \\ .333333 \\ .208333 \\ .333333 \end{pmatrix}, \begin{pmatrix} .104167 \\ .375 \\ .208333 \\ .3125 \end{pmatrix}, \begin{pmatrix} .107143 \\ .357142 \\ .214286 \\ .321429 \end{pmatrix}, \begin{pmatrix} .107143 \\ .357143 \\ .214286 \\ .321429 \end{pmatrix}$$

Reihenfolge der Seiten: 2,4,3 und 1. Falls existent, heißt

$$w_\infty = \lim_{k \to \infty} P^k w.$$

eine **stationäre Verteilung** des stochastischen Prozesses.

Hauptsatz über stochastische Matrizen

Satz. *Sei $P \in \mathbb{R}^{n \times n}$ eine stochastische Matrix.*

1. *P hat den Eigenwert $\lambda_1 = 1$.*
2. *Jede stationäre Verteilung w_∞ ist eine Eigenvektor zu $\lambda_1 = 1$.*
3. *Alle komplexen Eigenwerte λ_j von P haben Betrag $|\lambda_j| \leq 1$.*
4. *Sind für ein $k > 0$ alle Einträge von P^k nicht Null, dann ist $\lambda_1 = 1$ der einzige Eigenwert vom Betrag 1 und*

$$\dim \mathrm{Eig}(P, 1) = 1 = m(\chi_P(t), 1).$$

5. *(Ergodensatz) Unter der Voraussetzung von 4. gilt*

$$\lim_{k \to \infty} P^k = (w_\infty w_\infty \dots w_\infty) \in \mathbb{R}^{n \times n}$$

Ende Teil 3

Bemerkung. Für Google's page rank ändert man die Matrix noch etwas ab. Statt P wie oben eingeführt, betrachtet man

$$Q = \varepsilon P_{all} + (1-\varepsilon)P$$

für $0 < \varepsilon \ll 1$, wobei $P_{all} = \frac{1}{n}\begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix}$.

Auch eine neue Adresse ins Addressfeld eintragen, führt mit geringer Wahrscheinlichkeit zu einer neuen Seite. Q ist eine stochastische Matrix, die den Voraussetzungen des Ergodensatzes genügt.

Ändert sich das World-Wide-Web, so muss die Matrix P um die neuen Seiten vergrößert werden. Da die stationäre Verteilung des alten Netzes eine gute Approximation der neuen stationären Verteilung ist, kann das neue Ranking mit wenigen Iterationen berechnet werden.

Zum Beweis des Hauptsatzes über stochastische Matrizen

Zu 1.: Da eine Matrix und ihre Transponierte das gleiche charakteristische Polynom haben, genügt es zu zeigen, dass P^t den Eigenwert $\lambda_1 = 1$ hat. Es gilt

$$(1 \dots 1)P = (1 \dots 1),$$

da P eine stochastische Matrix ist. Also ist $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 1 von P^t.

Zu 2.:

$$w_\infty = \lim_{k\to\infty} P^k w = \lim_{k\to\infty} P^{k+1} w = P \lim_{k\to\infty} P^k w = P w_\infty$$

3. folgt aus dem Satz von Gerschgorin.

Satz über die Gerschgorin-Kreise

Satz. *Sei* $A = (a_{kj}) \in \mathbb{C}^{n\times n}$ *eine quadratische komplexe Matrix. Es bezeichne*

$$D_j = \{z \in \mathbb{C} \mid |z - a_{jj}| \le \sum_{k\neq j} |a_{kj}|\}.$$

Dann liegen alle Eigenwerte λ *von* A *in der Vereinigung* $\bigcup D_j$*, wobei jede zusammenhängende Komponente der Vereinigung genauso viele Eigenwerte enthält wie Kreise an deren Vereinigung beteiligt sind.*

Beispiel. $A = \begin{pmatrix} 2 & 1 & 0.5 \\ 0.2 & 5 & 0.7 \\ 1 & 0 & 6 \end{pmatrix}$

Eigenwerte

≈ 1.8692

≈ 4.87303

≈ 6.25777

Anwendung auf stochastische Matrizen

Im Fall einer stochastischen Matrix $P = (p_{ij})$ hat der Kreis

$$D_j = \{z \in \mathbb{C} \mid |z - p_{jj}| \leq \sum_{k \neq j} |p_{kj}|\}.$$

Mittelpunkt in $p_{jj} \in [0,1]$, und 1 liegt auf dem Rand des Kreises, da

$$1 - p_{jj} = \sum_{k \neq j} p_{kj} \Leftrightarrow \sum_{k=1}^{n} p_{kj} = 1$$

Alle Gerschgorin-Kreise sind also im Einheitskreis $\{z \in \mathbb{C} \mid |z| \leq 1\}$ enthalten, und, falls $p_{jj} > 0$, ist $1 \in D_j$ der einzige Punkt $z \in D_j$ mit $|z| = 1$.

Eine stochastische Matrix mit $\zeta_k = e^{\frac{2\pi i}{k}}$ als Eigenwert

Die Permutationsmatrix $Q = (e_{\sigma(1)} \dots e_{\sigma(k)}) \in \mathbb{R}^{k \times k}$ zu $\sigma = (1\ 2 \dots\ k)$ hat das charakteristische Polynom

$$\chi_Q(t) = (-1)^k (t^k - 1).$$

Ende Teil 4