Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Lücken

Die Themen heute sind

- der Fundamentalsatz der Algebra
- die Jordansche Normalform

Im Beweis des Hauptsatzes über stochastische Matrizen blieb der Beweis des Fundamentalsatzes der Algebra und der Beweis des Satzes über die Jordansche Normalform offen. Heute wollen wir diese Lücken schließen.

Fundamentalsatz der Algebra

Satz. *Sei* $p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in \mathbb{C}[t]$ *ein normiertes Polynom mit komplexen Koeffizienten* $a_k \in \mathbb{C}$ *vom Grad* $n \geqslant 1$*. Dann hat* $p(t)$ *ein Nullstelle* $z \in \mathbb{C}$*.*

Korollar. *Ein normiertes Polynom vom Grad* $n \geqslant 1$

$$p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in \mathbb{C}[t]$$

faktorisiert vollständig in Linearfaktoren, d.h., es existieren Nullstellen $z_1, \ldots, z_n \in \mathbb{C}$ *von* $p(t)$*, so dass folgendes gilt:*

$$p(t) = \prod_{k=1}^{n} (t - z_k).$$

Beweis des Korollars mit dem Satz

Nach dem Satz hat $p(t)$ eine Nullstelle z_n. Division mit Rest von $p(t)$ nach $t - z_n$ liefert einen Ausdruck

$$p(t) = q(t)(t - z_n) + r$$

mit r ein Polynom vom Grad $< 1 = \deg(t - z_n)$, also ist r eine Konstante. Da $p(z_n) = 0$, folgt $r = 0$. Also

$$p(t) = q(t)(t - z_n),$$

wobei $q(t)$ ein normiertes Polynom vom Grad $\deg q(t) = n - 1$ ist. Mit Induktion nach n können wir annehmen, dass

$$q(t) = (t - z_1) \cdot \ldots \cdot (t - z_{n-1})$$

in $n - 1$ Linearfaktoren zerfällt, und die Behauptung folgt. $\square$

Beweis des Fundamentalsatzes

Wir können $a_0 \neq 0$ annehmen, da andernfalls $z = 0$ eine Nullstelle ist. Wir betrachten für einen Radius r den geschlossenen Weg (Kurve)

$$\gamma_r \colon [0, 2\pi] \to \mathbb{C}, \alpha \mapsto \gamma_r(\alpha) = p(re^{i\alpha})$$

in der komplexen Zahlenebene.

Wir stellen uns dabei vor, dass $\alpha \mapsto r^n e^{in\alpha}$ den Spaziergang eines Herren und $\gamma_r \colon \alpha \mapsto p(r^n e^{in\alpha})$ den Weg eines Hundes beschreibt, den der Herr an der Leine führt.

Es handelt sich dabei um eine Leine mit flexibler Länge

$$\ell(\alpha) = |a_{n-1} r^{n-1} e^{i(n-1)\alpha} + \ldots + a_1 r e^{i\alpha} + a_0|$$

1. Ist $r \gg 1$, dann ist $\ell(\alpha) < r^n$, und der Hund muß genauso oft um den Nullpunkt herumlaufen wie der Herr, also n-mal. In der Tat

$$\frac{\ell(\alpha)}{r^n} = \Big| \sum_{k=1}^{n} a_{n-k} \frac{1}{r^k} e^{i(n-k)\alpha} \Big| \leqslant \sum_{k=1}^{n} \frac{|a_{n-k}|}{r^k} \xrightarrow[r\to\infty]{} 0.$$

2. Ist $1 \gg r > 0$, dann läuft der Hund in einer Umgebung von $a_0 \in \mathbb{C}$ herum. Nur der Herr läuft n-mal um den Nullpunkt herum.

3. Hat $p(t)$ keine Nullstelle z mit $r_1 \leqslant |z| \leqslant r_2$, dann können wir für jedes r mit $r_1 \leqslant r \leqslant r_2$ genau sagen, wie oft der Hund bei dem Weg γ_r um den Nullpunkt herumgelaufen ist. Die Umlaufzahl hängt einerseits stetig von $r \in [r_1, r_2]$ ab. Andererseits hat die Umlaufzahl Werte in $\mathbb{Z}$, da der Hund am Ende des Weges wieder am Startpunkt ist: $\gamma_r(0) = \gamma_r(2\pi)$. Die Umlaufzahlen für die Wege γ_r mit $r \in [r_1, r_2]$ sind folglich konstant.
4. Angenommen $p(t)$ hat gar keine Nullstellen. Dann sind nach 3. die Umlaufzahlen für γ_r mit $r \gg 0$ und γ_r für $1 \gg r > 0$ gleich. Dies ist aber nach 1. und 2. nicht der Fall, ein Widerspruch. Es muss also eine Nullstelle geben.

Ende Teil 1

Jordansche Normalform

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*

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

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

Beweisschritte

Schritt 1. *Sei $A \in \mathbb{C}^{n \times n}$ eine quadratische Matrix mit komplexen Einträgen. Dann existieren $S \in \mathrm{GL}(n, \mathbb{C})$, so dass*

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ 0 & & & \lambda_n \end{pmatrix}$$

eine obere Dreiecksmatrix ist.

Beweis. Das charakteristische Polynom $\chi_A(t)$ hat nach dem Fundamentalsatz der Algebra eine Nullstelle λ_1. Sei $v_1 \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Wir ergänzen v_1 zu einer Basis

$$\mathcal{B} = \{v_1, \ldots, v_n\} \text{ von } \mathbb{C}^n.$$

Bzgl. dieser Basis hat A die Gestalt

$$A' = \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix} = M_{\mathcal{B}}^{\mathcal{B}}(A) = S_1 A S_1^{-1},$$

wobei die Spalten von S_1^{-1} die Vektoren $v_1, \ldots, v_n$ sind. Es gilt:

$$\chi_A(t) = \chi_{A'}(t) = (\lambda_1 - t)\chi_{\tilde{A}}(t).$$

Mit Induktion können wir annehmen, dass eine Matrix $\tilde{S} \in \mathrm{GL}(n-1, \mathbb{C})$ existiert, so dass $\tilde{S}\tilde{A}\tilde{S}^{-1}$ eine obere Dreiecksmatrix ist.

Die Matrix

$$S = S_2 S_1 \text{ mit } S_2 = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S} \end{pmatrix}$$

überführt A in eine obere Dreiecksmatrix:

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S} \end{pmatrix} \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}^{-1} \end{pmatrix} &= \begin{pmatrix} \lambda_1 & * \\ 0 & \tilde{S}\tilde{A} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \tilde{S}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & *\tilde{S}^{-1} \\ 0 & \tilde{S}\tilde{A}\tilde{S}^{-1} \end{pmatrix}. \quad \square \end{aligned}$$

Ende Teil 2

Schritt 2: Zerlegung in Haupträume

Definition. Sei $A \in \mathbb{C}^{n \times n}$ eine quadratische Matrix und $\lambda \in \mathbb{C}$ ein Eigenwert von A. Dann heisst

$$\mathrm{Hau}(A, \lambda) = \ker((A - \lambda E)^n) \subset \mathbb{C}^n$$

der **Hauptraum** von A zum Eigenwert λ.

Bemerkung. Die Räume

$$\begin{aligned} 0 \subsetneq \mathrm{Eig}(A, \lambda) &= \ker(A - \lambda E) \subset \ker((A - \lambda E)^2) \subset \dots \\ \dots &\subset \ker((A - \lambda E)^m) \subset \dots \subset \mathrm{Hau}(A, \lambda) \subset \mathbb{C}^n \end{aligned}$$

bilden eine aufsteigende Kette von Untervektorräumen, die, wie wir gleich sehen werden, nach spätestens n Schritten nicht mehr grösser wird.

Satz. *Sei $A \in \mathbb{C}^{n \times n}$ eine quadratische Matrix und $\lambda \in \mathbb{C}$ ein Eigenwert von A. Sei $m = m(\chi_A(t), \lambda)$ die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms. Dann gilt*

$$\dim \mathrm{Hau}(A, \lambda) = m.$$

Beweis. Die Nullstellen $\lambda_1, \ldots, \lambda_n$ nummerieren wir so durch, dass $\lambda = \lambda_1 = \ldots = \lambda_m$ und die verbleibenden Eigenwerte $\lambda_{m+1}, \ldots, \lambda_n$ ungleich λ sind. Dann hat in unserer Basis aus Schritt 1. die obere Dreiecksmatrix die Gestalt

$$S(A - \lambda E)S^{-1} = \left(\begin{array}{ccc|ccc} 0 & \cdots & * & & & \\ & \ddots & \vdots & & B & \\ & & 0 & & & \\ \hline & & & \lambda_{m+1} - \lambda & \ldots & * \\ & 0 & & & \ddots & \vdots \\ & & & & & \lambda_n - \lambda \end{array}\right)$$

Der rechte obere $m \times m$ Block

$$N = \begin{pmatrix} 0 & \cdots & * \\ & \ddots & \vdots \\ & & 0 \end{pmatrix}$$

ist eine strikte obere Dreiecksmatrix, also spätestens $N^m = 0$. Es folgt

$$S(A - \lambda E)^k S^{-1} = \left(\begin{array}{c|ccc} 0 & & \tilde{B} & \\ \hline & (\lambda_{m+1} - \lambda)^k & \dots & \tilde{*} \\ 0 & & \ddots & \vdots \\ & & & (\lambda_n - \lambda)^k \end{array} \right)$$

und deshalb

$$m = \dim \ker((A - \lambda E)^k) = \dim \mathrm{Hau}(A, \lambda) \quad \forall k \geqslant m. \qquad \square$$

Korollar.

1. *A bildet die Haupträume in sich ab.*
2. *Bezeichnen $\lambda_1, \ldots, \lambda_s$ die paarweise verschiedenen Eigenwerte von A, dann gilt*

$$\mathrm{Hau}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Hau}(A, \lambda_s) = \mathbb{C}^n.$$

Beweis. 1.

$$SAS^{-1} = \left(\begin{array}{ccc|ccc} \lambda_1 & \cdots & * & & & \\ & \ddots & \vdots & & B & \\ & & \lambda_1 & & & \\ \hline & & & \lambda_{m+1} & \cdots & * \\ & 0 & & & \ddots & \vdots \\ & & & & & \lambda_n \end{array}\right)$$

zeigt, das der Hauptraum

$$\mathrm{Hau}(SAS^{-1}, \lambda_1) = \mathrm{Spann}(e_1, \ldots, e_m)$$

von SAS^{-1} in sich abgebildet wird.

Der Hauptraum

$$\mathrm{Hau}(A, \lambda_1) = \mathrm{Spann}(S^{-1}e_1, \ldots, S^{-1}e_m)$$

wird also von $A = S^{-1}SA$ in sich abgebildet. Das gleiche trifft auf alle anderen Haupträume zu.

2. Es bezeichne von nun an $\lambda_1, \ldots, \lambda_s$ die paarweise verschiedenen Eigenwerte von A. Angenommen

$$z = z_2 + \ldots + z_s \in \mathrm{Hau}(A, \lambda_1) \cap (\sum_{\ell=2}^{s} \mathrm{Hau}(A, \lambda_\ell) \subset \mathbb{C}^n.$$

Dann wird z von einer Potenz von $(A - \lambda_1 E)$ auf Null abgebildet. $z_\ell \in \mathrm{Hau}(A, \lambda_\ell)$ wird durch Potenzen von $A - \lambda_1 E$ nicht auf Null abgebildet, da $\lambda_\ell \neq \lambda_1$, es sei denn, $z_\ell = 0$.

Mit Induktion nach s dürfen wir annehmen, dass die Summe

$$\sum_{\ell=2}^{s} \mathrm{Hau}(A, \lambda_\ell) = \mathrm{Hau}(A, \lambda_2) \oplus \ldots \oplus \mathrm{Hau}(A, \lambda_s).$$

direkt ist.

Es folgt

$$\operatorname{Hau}(A,\lambda_1)\cap\Big(\sum_{\ell=2}^{s}\operatorname{Hau}(A,\lambda_\ell)\Big)=\{0\},$$

und damit ist auch

$$\sum_{\ell=1}^{s}\operatorname{Hau}(A,\lambda_\ell)=\operatorname{Hau}(A,\lambda_1)\oplus\operatorname{Hau}(A,\lambda_2)\oplus\ldots\oplus\operatorname{Hau}(A,\lambda_s)$$

direkt.

$$\bigoplus_{\ell=1}^{s}\operatorname{Hau}(A,\lambda_\ell)=\mathbb{C}^n$$

folgt nun aus Dimensionsgründen:

$$\chi_A(t)=-(1)^n\prod_{\ell=1}^{s}(t-\lambda_\ell)^{m_\ell}$$

gibt

$$\dim\mathbb{C}^n=n=\deg\chi_A(t)=\sum_{\ell=1}^{s}m_\ell\overset{1.}{=}\sum_{\ell=1}^{s}\dim\operatorname{Hau}(A,\lambda_\ell).\qquad\square$$

Ende Teil 3

Zusammenfassung von Schritt 2.

Wir wissen nun, dass A konjugiert zu einer Blockmatrix der Gestalt

$$\begin{pmatrix} \lambda_1 E_{m_1} + N_1 & & \\ & \ddots & \\ & & \lambda_s E_{m_s} + N_s \end{pmatrix}$$

ist, wobei die N_ℓ strikte obere Dreiecksmatrizen vom Format $m_\ell \times m_\ell$ sind.

Schritt 3. Klassifikation von nilpotenten Matrizen.

Satz. *Sei $N \in \mathbb{C}^{m \times m}$ eine nilpotente Matrix. Dann ist N konjugiert zu einer Matrix*

$$J = \begin{pmatrix} J(0, r_1) & & \\ & \ddots & \\ & & J(0, r_k) \end{pmatrix}.$$

Wenn wir die Blöcke

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

der Größe nach sortieren

$$r_1 \geqslant r_2 \geqslant \ldots \geqslant r_k > 0,$$

dann ist $r_1 + \ldots + r_k = m$ eine Partition von m.

Bemerkung. Die Anzahl k der Jordankästchen von N ist

$$k = \dim \operatorname{Eig}(N, 0).$$

Der Trick beim Beweis besteht darin, statt der Eigenvektoren die Vektoren zu betrachten, die am längsten durch Potenzen von N nicht auf Null abgebildet werden. Für einen Jordanblock

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

ist dies z.B. der Vektor $e_r \in \mathbb{C}^r$. Es gilt

$$J(0,r)e_r = e_{r-1}, J(0,r)e_{r-1} = e_{r-2}, \ldots, J(0,r)e_2 = e_1$$

bis wir schließlich den Eigenvektor e_1 erreichen, der auf Null abgebildet wird

$$J(0,r)e_1 = 0.$$

Hilfssatz. *Es seien Vektoren* $v_1, \dots, v_\ell \in \ker N^k$ *für ein* $k \geqslant 2$ *gegeben. Dann gilt*

1. $Nv_1, \dots Nv_\ell \in \ker N^{k-1}$.
2. *Falls* $\dim(\ker N^{k-1} + \mathrm{Spann}(v_1, \dots, v_\ell)) = \dim \ker N^{k-1} + \ell$ *gilt, dann gilt auch*

$$\dim(\ker N^{k-2} + \mathrm{Spann}(Nv_1, \dots, Nv_\ell)) = \dim \ker N^{k-2} + \ell.$$

Beweis. Die erste Aussage ist klar. Zur zweiten Aussage: Angenommen

$$\dim(\ker N^{k-2} + \mathrm{Spann}(Nv_1, \dots, Nv_\ell)) < \dim \ker N^{k-2} + \ell.$$

Dann existieren Skalare $\alpha_1, \dots, \alpha_\ell$, die nicht alle Null sind, so dass

$$\sum_{j=1}^{\ell} \alpha_j N v_j \in \ker N^{k-2}$$

gilt. Es folgt

$$N^{k-1}(\sum_{j=1}^{\ell} \alpha_j v_j) = N^{k-2}(\sum_{j=1}^{\ell} \alpha_j N v_j) = 0, \text{ also } \sum_{j=1}^{\ell} \alpha_j v_j \in \ker N^{k-1},$$

was der Voraussetzung widerspricht.

Korollar. *Für $k \geqslant 2$ gilt*

$$\dim \ker N^k - \dim \ker N^{k-1} \leqslant \dim \ker N^{k-1} - \dim \ker N^{k-2}. \qquad \square$$

Wir betrachten nun die minimale Potenz p, so dass $N^p = 0$, also

$$\ker N^{p-1} \subsetneq \ker N^p = \mathbb{C}^m \text{ und } \ell = \dim \ker N^p - \dim \ker N^{p-1} > 0.$$

Wir ergänzen eine Basis von $\ker N^{p-1}$ durch Vektoren $v_1, \dots, v_\ell$ zu einer Basis von $\ker N^p$.
Als nächstes betrachten wir die Vektoren

$$Nv_1, \dots, Nv_\ell \in \ker N^{p-1}.$$

Nach dem Hilfssatz können wir diese Vektoren zusammen mit einer Basis von $\ker N^{p-2}$ durch zusätzliche Vektoren $v_{\ell+1}, \dots, v_{\ell_1}$ zu einer Basis von $\ker N^{p-1}$ ergänzen. Im nächsten Schritt betrachten wir die Vektoren

$$N^2 v_1, \dots, N^2 v_\ell, Nv_{\ell+1}, \dots, Nv_{\ell_1},$$

die wir gegebenfalls durch weitere Vektoren $v_{\ell_1+1}, \dots, v_{\ell_2}$ ergänzen.

Insgesamt erhalten wir ein System von m Vektoren

$$
\begin{array}{ccccccccc}
v_1 & \dots & v_\ell & & & & & & \\
Nv_1 & \dots & Nv_\ell & v_{\ell+1} & \dots & v_{\ell_1} & & & \\
\vdots & & \vdots & \vdots & & \vdots & & & \\
\vdots & & \vdots & \vdots & & \vdots & \vdots & \ddots & \\
N^{p-1}v_1 & \dots & N^{p-1}v_\ell & N^{p-2}v_{\ell+1} & \dots & N^{p-2}v_{\ell_1} & N^{p-3}v_{\ell_1+1} & \dots & v_{\ell_{p-1}}
\end{array}
$$

die eine Basis von $\mathbb{C}^m$ bilden. Geeignet angeordnet hat N in dieser Basis die gewünschte Matrixdarstellung. Jede Spalte gibt einen Jordanblock. $\square$

Der Bewies des Satzes über die Jordansche Normalform ergibt sich durch Zusammenfügen der Schritte 1,2,3. $\square$

Korollar. *Sei* $A \in \mathbb{C}^{n\times n}$. *Dann ist* $\dim \operatorname{Eig}(A, \lambda)$ *die Anzahl der Jordanblöcke* $J(\lambda, r_j)$ *von* A *zum Eigenwert* λ. $\square$

Ende Teil 4