Bachelorarbeit

Der Satz von Perron-Frobenius

Borbala Mercedes Gerhat
0625423
Betreuer: Ao.Univ.Prof. Dipl.-Ing. Dr.Techn. Michael Kaltenbäck

15.01.2012

Inhaltsverzeichnis

Diese Arbeit beschäftigt sich mit reellen, positiven bzw. nichtnegativen Matrizen und gewissen Aussagen über deren Spektrum. Das Hauptresultat ist einerseits der Satz von Perron, der eine schöne Charakterisierung der Eigenwerte und Eigenvektoren positiver Matrizen liefert, andererseits der Satz von Perron-Frobenius, der die Resultate unter gewissen Voraussetzungen für nichtnegative Matrizen verallgemeinert. Diese Arbeit stammt rein aus der linearen Algebra, man kann jedoch den Begriff der positiven Matrix auf den unendlichdimensionalen Fall ausdehnen, nämlich durch den des positiven Operators zwischen Banachräumen. Eine Verallgemeinerung des Satzes von Perron-Frobenius findet sich auf dem Gebiet der Funktionalanalysis im Satz von Krein-Rutman wieder.

Die Arbeit ist in drei Teile gegliedert. Im ertsen Teil finden sich beweistechnisch relevante Aussagen und Definitionen aus der linearen Algebra, im zweiten Teil der Satz von Perron, der den Fall einer strikt positiven Matrix behandelt und im dritten Teil der Satz von Perron-Frobenius, der die Aussagen bestmöglich auf den Fall einer nichtnegativen Matrix überträgt.

Kapitel 1

Einige Resultate der Linearen Algebra

Der Vollständigkeit halber folgen zunächst einige Definitionen und Aussagen aus der Spektraltheorie komplexer Matrizen. Sie werden zum Teil ohne Beweis zitiert, sind für spätere Beweise von Bedeutung und sollen gleichzeitig als Einführung in diese Arbeit dienen.

Definition 1.1. Seien $A, B \in \mathbb{C}^{n\times m}$ mit $n, m \in \mathbb{N}$.

- A heißt *positiv* (in Zeichen $A > 0$), falls

  $$a_{ij} \in \mathbb{R},\, a_{ij} > 0 \quad \text{für alle } i = 1, \ldots, n,\, j = 1, \ldots, m.$$

  A heißt *nichtnegativ* (i.Z. $A \geq 0$), falls

  $$a_{ij} \in \mathbb{R},\, a_{ij} \geq 0 \quad \text{für alle } i = 1, \ldots, n,\, j = 1, \ldots, m.$$

  Entsprechend definiert man $A < B$ bzw. $A \leq B$, falls $B - A > 0$ bzw. $B - A \geq 0$.

- $|A|$ sei die Matrix der Absolutbeträge,

  $$|A| \;:= (|a_{ij}|)_{\substack{i=1,\ldots,n\\ j=1,\ldots,m}}\,.$$

Für $m = 1$ umfasst diese Definition auch Vektoren. //

Definition 1.2. Sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$ und $A \in \mathbb{C}^{n\times n}$. Man definiert die von $\|\cdot\|$ *induzierte Matrixnorm* $\|\cdot\|_M$ durch

$$\|A\|_M := \sup_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|}$$

//

Bemerkung 1.3. Definiere für ein beliebiges $x \in \mathbb{C}^n\setminus\{0\}$ den Vektor $\tilde{x} := \frac{x}{\|x\|}$. Dann gilt $\|\tilde{x}\| = 1$ und

$$\frac{\|A\tilde{x}\|}{\|\tilde{x}\|} = \frac{\left\|A\frac{x}{\|x\|}\right\|}{\left\|\frac{x}{\|x\|}\right\|} = \frac{\frac{1}{\|x\|}\,\|Ax\|}{\frac{1}{\|x\|}\,\|x\|} = \frac{\|Ax\|}{\|x\|}$$

und damit

$$\sup_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} = \sup_{\substack{x\in\mathbb{C}^n\\ \|x\|=1}} \frac{\|Ax\|}{\|x\|}.$$

Da die Menge $\{ x \in \mathbb{C}^n \mid \|x\| = 1 \}$ in $\mathbb{C}^n$ kompakt ist, nimmt die Abbildung

$$x \mapsto \frac{\|Ax\|}{\|x\|}$$

auf ihr ein Maximum an. Es gilt also

$$\sup_{\substack{x\in\mathbb{C}^n\\ \|x\|=1}} \frac{\|Ax\|}{\|x\|} = \max_{\substack{x\in\mathbb{C}^n\\ \|x\|=1}} \frac{\|Ax\|}{\|x\|}$$

und daher auch

$$\|A\|_M = \sup_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} = \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|}.$$

//

Satz 1.4. *Sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$. Dann ist die in Definition 1.2 definierte Abbildung $\|\cdot\|_M : \mathbb{C}^{n\times n} \mapsto \mathbb{R}$ tatsächlich eine Norm auf $\mathbb{C}^{n\times n}$ und für $x \in \mathbb{C}^n$ und $A, B \in \mathbb{C}^{n\times n}$ gilt*

$$\|Ax\| \leq \|A\|_M \|x\|$$

und

$$\|AB\|_M \leq \|A\|_M \|B\|_M .$$

Beweis. Dass $\|\cdot\|_M$ eine Norm ist, folgt leicht aus der Tatasache, dass $\|\cdot\|$ eine ist. Es gilt klarerweise $\|A\|_M \geq 0$ für alle $A \in \mathbb{C}^{n\times n}$ und für $A = 0$ ist $\|A\|_M = 0$. Sei umgekehrt $\|A\|_M = 0$ für ein $A \in \mathbb{C}^{n\times n}$. Dann muss wegen

$$0 = \|A\|_M = \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|}$$

für alle $x \in \mathbb{C}^n$ schon $\|Ax\| = 0$ und damit $Ax = 0$ gelten. Also haben wir $\|A\|_M = 0$ genau dann, wenn $A = 0$. Für $\lambda \in \mathbb{C}$ gilt

$$\|\lambda A\|_M = \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|\lambda Ax\|}{\|x\|} = \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{|\lambda|\,\|Ax\|}{\|x\|} = |\lambda| \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} = |\lambda|\,\|A\|_M .$$

Um die Dreiecksungleichung einzusehen, seien $A, B \in \mathbb{C}^{n\times n}$. Dann gilt wegen

$$\max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|(A+B)x\|}{\|x\|} \leq \max_{x\in\mathbb{C}^n\setminus\{0\}} \left(\frac{\|Ax\|}{\|x\|} + \frac{\|Bx\|}{\|x\|} \right) \leq \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|}{\|x\|} + \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Bx\|}{\|x\|}$$

auch $\|A+B\|_M = \|A\|_M + \|B\|_M$ und damit ist $\|\cdot\|_M$ eine Norm. Für $x \in \mathbb{C}^n$ und $A \in \mathbb{C}^{n\times n}$ haben wir

$$\|A\|_M = \max_{y\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ay\|}{\|y\|} \geq \frac{\|Ax\|}{\|x\|}$$

und damit $\|Ax\| \leq \|A\|_M \|x\|$. Daraus erhält man für $A, B \in \mathbb{C}^{n\times n}$

$$\max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|ABx\|}{\|x\|} \leq \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|A\|_M \|Bx\|}{\|x\|} \leq \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|A\|_M \|B\|_M \|x\|}{\|x\|},$$

also $\|AB\|_M \leq \|A\|_M \|B\|_M$, womit alle Aussagen bewiesen sind. □

Beispiel 1.5. Wählt man auf $\mathbb{C}^n$ speziell die Maximumsnorm,

$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|, \quad x \in \mathbb{C}^n,$$

so gilt wegen

$$\max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\max_{i=1,\dots,n}|\sum_{j=1}^n a_{ij}x_j|}{\|x\|_\infty} \le \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\max_{i=1,\dots,n}\sum_{j=1}^n |a_{ij}||x_j|}{\|x\|_\infty} \le \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$$

für die induzierte Matrixnorm die Ungleichung

$$\|A\|_{\infty,M} = \max_{x\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ax\|_\infty}{\|x\|_\infty} \le \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|.$$

Sei $i_0 \in \{1,\dots,n\}$ jener Index, sodass $\max_{i=1,\dots,n}\sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{i_0 j}|$ und setze

$$x_j = \begin{cases} \frac{|a_{i_0 j}|}{a_{i_0 j}} & \text{wenn } a_{i_0 j} \neq 0 \\ 0 & \text{wenn } a_{i_0 j} = 0 \end{cases}$$

Dann gilt $\|x\|_\infty = 1$ und damit

$$\|A\|_{\infty,M} = \max_{y\in\mathbb{C}^n\setminus\{0\}} \frac{\|Ay\|_\infty}{\|y\|_\infty} \ge \frac{\|Ax\|_\infty}{\|x\|_\infty} = \|Ax\|_\infty = \sum_{j=1}^n |a_{i_o j}| = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|.$$

Insgesamt erhalten wir also

$$\|A\|_{\infty,M} = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|.$$

Man nennt $\|\cdot\|_{\infty,M}$ auch *Zeilensummennorm* und bezeichnet sie der Einfachheit halber auch mit $\|\cdot\|_\infty$, wenn aus dem Zusammenhang klar ist, dass es sich um die Matrixnorm handelt. //

Definition 1.6. Sei $A \in \mathbb{C}^{n\times n}$.

- Jedes $\lambda \in \mathbb{C}$ mit
$$\ker(A - \lambda I) \neq \{0\}$$
heißt *Eigenwert* von A. Jedes $x \in \mathbb{C}^n$ mit
$$x \in \ker(A - \lambda I)\setminus\{0\}$$
heißt *Eigenvektor* von A zum Eigenwert λ. Das Paar (λ, x) heißt *Eigenpaar* von A, der lineare Unterraum $\ker(A - \lambda I)$ heißt *Eigenraum* von A zum Eigenwert λ.

- Das *Spektrum* $\sigma(A)$ bezeichnet die Menge aller Eigenwerte von A,
$$\sigma(A) = \{\ \lambda \in \mathbb{C} \mid \ker(A - \lambda I) \neq \{0\}\ \}.$$

- Das Polynom $p(\lambda) = \det(A - \lambda I)$ heißt *charakteristisches Polynom* von A.

- Man definiert den *Spektralradius* von A als
$$r(A) := \max\{\ |\lambda| \mid \lambda \in \sigma(A)\ \}.$$

//

Bemerkung 1.7. $\lambda \in \mathbb{C}$ ist genau dann ein Eigenwert von A, wenn die Matrix $(A - \lambda I)$ nichttrivialen Kern hat, also wenn sie singulär ist. Das ist genau dann der Fall, wenn

$$p(\lambda) = \det(A - \lambda I) = 0.$$

Die Eigenwerte von A sind also genau die Nullstellen des charakteristischen Polynoms. //

Definition 1.8. Sei $A \in \mathbb{C}^{n\times n}$. Die Vielfachheit eines Eigenwertes λ von A als Nullstelle des charakteristischen Polynoms wird als *algebraische Vielfachheit* von λ, die Dimension des zu λ gehörigen Eigenraumes als *geometrische Vielfachheit* bezeichnet. //

Bemerkung 1.9. Sei $A \in \mathbb{C}^{n\times n}$ und $\lambda \in \mathbb{C}$ ein Eigenwert von A. Zur Verallgemeinerung von Eigenvektor und Eigenraum führt man die Begriffe Hauptvektor und Hauptraum ein. Ein Vektor $x \in \mathbb{C}^n \backslash \{0\}$ heißt *Hauptvektor* von A zum Eigenwert λ, wenn es ein $l \in \mathbb{N}$ gibt mit

$$x \in \ker(A - \lambda I)^l \backslash \{0\}.$$

Die Zahl

$$m := \min \left\{ \, l \in \mathbb{N} \, \middle| \, x \in \ker(A - \lambda I)^l \, \right\} \in \mathbb{N}$$

heißt *Stufe* des Hauptvektors x. Hierbei sei bemerkt, dass Hauptvektoren der Stufe 1 genau den Eigenvektoren entsprechen. Das Konzept des Hauptvektors stellt also wirklich eine Verallgemeinerung des Eigenvektors dar. Klarerweise gilt für $k, l \in \mathbb{N}$ mit $k \leq l$

$$\ker(A - \lambda I)^k \subseteq \ker(A - \lambda I)^l.$$

Also gibt es zu jedem Eigenwert λ eine kleinste Zahl $d \in \mathbb{N}$ mit

$$\ker(A - \lambda I)^d = \ker(A - \lambda I)^{d+r} \quad \text{für alle } r \in \mathbb{N}.$$

Der lineare Unterraum $\ker(A - \lambda I)^d$ heißt *Hauptraum* von A zum Eigenwert λ. An dieser Stelle sei ohne Beweis bemerkt, dass die Dimension des Hauptraumes genau mit der algebraischen Vielfachheit von λ übereinstimmt.

Betrachte die Matrix A als lineare Abbildung

$$A : \left\{ \begin{array}{ccc} \mathbb{C}^n & \to & \mathbb{C}^n \\ x & \mapsto & Ax \end{array} \right.$$

und bezeichne mit $H_\lambda \subseteq \mathbb{C}^n$ den Hauptraum von A zum Eigenwert λ. Wieder bemerken wir ohne Beweis, dass man durch Einschränkung von A auf den Unterraum H_λ eine lineare Selbstabbildung

$$A|_{H_\lambda} : \left\{ \begin{array}{ccc} H_\lambda & \to & H_\lambda \\ x & \mapsto & Ax \end{array} \right.$$

erhält, dessen einziger Eigenwert λ ist. A bildet also den linearen Unterraum H_λ in sich selbst ab und es gilt

$$\sigma(A|_{H_\lambda}) = \{\lambda\}.$$

Der Hauptraum von A zu einem Eigenwert λ entspricht also genau jenem unter A invarianten Unterraum, auf dem A nur den einzigen Eigenwert λ besitzt. Zudem ist die direkte Summe aller H_λ, wobei $\lambda \in \sigma(A)$ läuft, genau $\mathbb{C}^n$. //

Mit Hilfe der in Bemerkung 1.9 gebrachten Tatsachen zeigt man folgendes wichtiges Ergebnis.

Satz 1.10 (Jordan'sche Normalform). *Sei $A \in \mathbb{C}^{n\times n}$ und sei*

$$p(\lambda) = \alpha(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_r)^{n_r}, \qquad \alpha \in \mathbb{C}$$

das charakteristische Polynom von A. Die Matrix A habe also genau $r > 0$ verschiedene Eigenwerte $\lambda_1, \lambda_2, \dots \lambda_r \in \mathbb{C}$. Dann ist A ähnlich zu einer Matrix $J \in \mathbb{C}^{n\times n}$, d.h. es gibt eine reguläre Matrix $P \in \mathbb{C}^{n\times n}$ mit $A = PJP^{-1}$, die die Gestalt

$$J = \begin{pmatrix} C_{\lambda_1} & 0 & \cdots & 0 \\ 0 & C_{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & C_{\lambda_r} \end{pmatrix}$$

hat, wobei jedes C_{λ_j} eine $n_j \times n_j$ Matrix von folgender Gestalt ist:

$$C_{\lambda_j} = \operatorname{diag}\left(\underbrace{J_{m_1(\lambda_j)}(\lambda_j), \dots, J_{m_1(\lambda_j)}(\lambda_j)}_{k_1(\lambda_j)\geq 1}, \underbrace{J_{m_2(\lambda_j)}(\lambda_j), \dots, J_{m_2(\lambda_j)}(\lambda_j)}_{k_2(\lambda_j)\geq 1}, \dots, \underbrace{J_{m_p(\lambda_j)}(\lambda_j), \dots, J_{m_p(\lambda_j)}(\lambda_j)}_{k_p(\lambda_j)\geq 1} \right)$$

mit $m_1(\lambda_j) > m_2(\lambda_j) > \cdots > m_p(\lambda_j) \geq 1$ und $\dim\ker(A - \lambda_j I) = \sum_{i=1}^{p(\lambda_j)} k_i(\lambda_j)$, wobei

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{m\times m}.$$

Die Matrix J heißt Jordan'sche Normalform von A.

Definition 1.11. Sei $A \in \mathbb{C}^{n\times n}$, $\lambda \in \mathbb{C}$ ein Eigenwert von A. Man definiert den *Index* von λ als die Dimension des größten zu λ gehörigen Jordan-Blocks in der Jordan-Zerlegung von A, $\operatorname{index}(\lambda) = m_1(\lambda)$ (vgl. Satz 1.10). //

Definition 1.12. Sei $p \in \mathbb{C}[X]$ ein Polynom, $p(x) = \sum_{i=1}^m p_i x^i$. Für $A \in \mathbb{C}^{n\times n}$ setzt man

$$p(A) := \sum_{i=1}^m p_i A^i \in \mathbb{C}^{n\times n}.$$

//

Satz 1.13. *Sei $p \in \mathbb{C}[X]$ ein Polynom, $p(x) = \sum_{i=1}^m p_i x^i$, und $A \in \mathbb{C}^{n\times n}$. Dann gilt*

$$\sigma(p(A)) = p(\sigma(A)),$$

wobei

$$p(\sigma(A)) := \{\ p(\lambda) \mid \lambda \in \sigma(A)\ \}.$$

Beweis. Wir zeigen zuerst $\sigma(p(A)) \supseteq p(\sigma(A))$. Sei dazu $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ ein Eigenpaar von A, $x \neq 0$. Zu zeigen ist, dass $p(\lambda) \in \sigma(p(A))$. Nun gilt aber

$$(p(A) - p(\lambda)I)x = p(A)x - p(\lambda)x = \sum_{i=1}^{m} p_i A^i x - \sum_{i=1}^{m} p_i \lambda^i x = \sum_{i=1}^{m} p_i \lambda^i x - \sum_{i=1}^{m} p_i \lambda^i x = 0.$$

Also ist $p(\lambda)$ ein Eigenwert von $p(A)$ zum Eigenvektor x und das liefert $p(\lambda) \in \sigma(p(A))$. Für die andere Inklusion sei $\lambda \in \sigma(p(A))$. Betrachte das Polynom

$$q(x) := p(x) - \lambda \in \mathbb{C}[X].$$

Über $\mathbb{C}$ zerfällt bekanntlich jedes Polynom in Linearfaktoren, d.h. es gibt $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ und $a \in \mathbb{C}$, sodass

$$q(x) = a(x - \lambda_1) \cdots (x - \lambda_m)$$

und damit auch

$$q(A) = p(A) - \lambda I = a(A - \lambda_1) \cdots (A - \lambda_m).$$

Wegen $\lambda \in \sigma(p(A))$ ist $q(A)$ singulär, also muss es ein λ_i geben, $i \in \{1, \dots, m\}$, sodass $(A - \lambda_i I)$ singulär ist. Das bedeutet aber genau $\lambda_i \in \sigma(A)$. Für dieses λ_i gilt

$$p(\lambda_i) - \lambda = q(\lambda_i) = 0,$$

also $p(\lambda_i) = \lambda$ und damit $\lambda \in p(\sigma(A))$. $\square$

Bemerkung 1.14.

(i). Im ersten Beweisteil von Satz 1.13 haben wir gesehen, dass für jedes Eigenpaar

$$(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n, \quad x \neq 0$$

von A gilt, dass $(p(\lambda), x)$ ein Eigenpaar von $p(A)$ ist. Insbesondere ist daher die geometrische Vielfachheit von $p(\lambda)$ als Eigenwert von $p(A)$ größer oder gleich der geometrischen Vielfachheit von λ als Eigenwert von A.

(ii). Sei $\lambda \in \mathbb{C}$ ein Eigenwert von A. Dann ist laut Bemerkung 1.9 der Hauptraum H_λ von A zu λ unter A invariant und es gilt

$$\sigma(A|_{H_\lambda}) = \{\lambda\}.$$

Daraus folgt aber, dass der lineare Unterraum H_λ auch unter $p(A)$ invariant ist und mit Satz 1.13 gilt

$$\sigma(p(A)|_{H_\lambda}) = \sigma(p(A|_{H_\lambda})) = \{p(\lambda)\}.$$

Also muss H_λ ein Teilraum des Hauptraumes von $p(A)$ zum Eigenwert $p(\lambda)$ sein, insbesondere muss die algebraische Vielfachheit von $p(\lambda)$ größer oder gleich der algebraischen Vielfachheit von λ sein.

//

Satz 1.15. *Sei $A \in \mathbb{C}^{n \times n}$. Dann gilt $\lim_{k\to\infty} A^k = 0$ genau dann, wenn $r(A) < 1$.*

Beweis. Sei $J \in \mathbb{C}^{n\times n}$ die Jordan'sche Normalform von A, d.h. $A = PJP^{-1}$ für ein reguläres $P \in \mathbb{C}^{n\times n}$. Dann gilt klarerweise wegen $A^k = PJ^kP^{-1}$, dass $\lim_{k\to\infty} A^k = 0$ genau dann, wenn $\lim_{k\to\infty} J^k = 0$. Wir erinnern uns daran, dass in der Diagonale von J Jordan-Blöcke der Gestalt

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{m\times m}$$

mit $\lambda \in \sigma(A)$ und $m \in \mathbb{N}$ stehen. Damit gilt, dass $\lim_{k\to\infty} J^k = 0$ genau dann, wenn für alle diese Blöcke $J_m(\lambda)$ gilt, dass $\lim_{k\to\infty} J_m(\lambda)^k = 0$. Zu zeigen ist also, dass für jedes $\lambda \in \sigma(A)$ und $m \in \mathbb{N}$, $\lim_{k\to\infty} J_m(\lambda)^k = 0$ genau dann gilt, wenn $|\lambda| < 1$. Für $\lambda = 0$ gilt $J_m(\lambda)^m = 0$ (siehe (1.2)), daher ist die Aussage in diesem Fall trivial, und wir können uns im Folgenden auf $\lambda \neq 0$ beschränken. Wegen dem Binomialsatz

$$(a+b)^k = \sum_{j=0}^{k} \binom{k}{j} a^j b^{k-j}$$

gilt

$$x^k = ((x-\lambda)+\lambda)^k = \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j}(x-\lambda)^j.$$

Aus dieser Darstellung erhalten wir mit $N := (J_m(\lambda) - \lambda I) \in \mathbb{C}^{m\times m}$

$$J_m(\lambda)^k = \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j}(J_m(\lambda)-\lambda I)^j = \sum_{j=0}^{k} \binom{k}{j} \lambda^{k-j} N^j. \tag{1.1}$$

Für die Matrix N gilt

$$N^0 = I = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}, \quad N^1 = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix},$$

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}, \quad \ldots, \quad N^{m-1} = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & 1 \\ \vdots & \ddots & & & & 0 \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} \tag{1.2}$$

und $N^j = 0$ für alle $j \geq m$. Setzt man $\binom{k}{j} = 0$ für $k < j$, so erhält man aus (1.1)

$$J_m(\lambda)^k = \begin{pmatrix} \lambda^k & \binom{k}{1}\lambda^1 & \binom{k}{2}\lambda^2 & \cdots & \binom{k}{m-1}\lambda^{k-m+1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \binom{k}{2}\lambda^2 \\ \vdots & & \ddots & \ddots & \binom{k}{1}\lambda^1 \\ 0 & \cdots & \cdots & 0 & \lambda^k \end{pmatrix}. \tag{1.3}$$

Gilt nun $\lim_{k\to\infty} J_m(\lambda)^k = 0$, so muss aufgrund der obigen Darstellung von $J_m(\lambda)^k$ schon $\lim_{k\to\infty} \lambda^k = 0$ gelten. Das aber impliziert wiederum $|\lambda| < 1$. Sei umgekehrt $|\lambda| < 1$. Für jedes feste $j \in \{1, \dots, m-1\}$ gilt

$$\binom{k}{j} = \frac{k(k-1)\cdots(k-j+1)}{j!} \leq \frac{k^j}{j!}.$$

Damit erhalten wir

$$|\binom{k}{j}\lambda^{k-j}| \leq \frac{k^j}{j!}|\lambda|^{k-j}.$$

Definiere nun die Funktionen

$$f(k) := k^j$$

$$g(k) := |\lambda|^{-k},$$

dann gilt

$$f^{(j)}(k) = j!$$

$$g^{(j)} = (-\ln|\lambda|)^j |\lambda|^{-k}.$$

Wegen $|\lambda| < 1$ gilt

$$\lim_{k\to\infty} \frac{j!}{(-\ln|\lambda|)^j|\lambda|^{-k}} = \lim_{k\to\infty} \frac{j!}{(-\ln|\lambda|)^j}|\lambda|^k = 0.$$

Daraus und aus j-maliger Anwendung der Regel von de L'Hospital folgt

$$0 = \lim_{k\to\infty} \frac{j!}{(-\ln|\lambda|)^j|\lambda|^{-k}} = \lim_{k\to\infty} \frac{f^{(j)}(k)}{g^{(j)}(k)} = \cdots = \lim_{k\to\infty} \frac{f(k)}{g(k)} = \lim_{k\to\infty} \frac{k^j}{|\lambda|^{-k}}$$

und damit

$$\lim_{k\to\infty} |\binom{k}{j}\lambda^{(k-j)}| \leq \lim_{k\to\infty} \frac{k^j}{j!}|\lambda|^{k-j} = \lim_{k\to\infty} \frac{|\lambda|^{-j}}{j!}\frac{k^j}{|\lambda|^{-k}} = 0.$$

Also konvergiert jeder Eintrag von $J_m(\lambda)^k$ für $k \to \infty$ gegen 0, was $\lim_{k\to\infty} J_m(\lambda)^k = 0$ zur Folge hat. □

Satz 1.16. *Sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$ und $\|\cdot\|_M$ die davon induzierte Matrixnorm. Seien $A, B \in \mathbb{C}^{n\times n}$, dann gilt für den Spektralradius*

(i). $r(A) \leq \|A\|_M$

(ii). $r(A) = \lim_{k\to\infty} \left\|A^k\right\|_M^{\frac{1}{k}}$

(iii). $0 \leq A \leq B \;\Rightarrow\; r(A) \leq r(B)$

Beweis. ad (i): Sei $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, $x \neq 0$ ein beliebiges Eigenpaar von A. Dann gilt

$$|\lambda| = \frac{|\lambda|\,\|x\|}{\|x\|} = \frac{\|\lambda x\|}{\|x\|} = \frac{\|Ax\|}{\|x\|} \leq \max_{x \in \mathbb{C}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \|A\|_M\,,$$

also $|\lambda| \leq \|A\|_M$ für alle $\lambda \in \sigma(A)$. Daraus folgt aber insbesondere $r(A) \leq \|A\|_M$.
ad (ii): Aus Satz 1.13 und (i) folgt

$$r(A)^k = \left(\max_{\lambda \in \sigma(A)} |\lambda| \right)^k = \max_{\lambda \in \sigma(A)} |\lambda|^k = \max_{\lambda \in \sigma(A)} |\lambda^k| = \max_{\lambda \in \sigma(A^k)} |\lambda| = r(A^k) \leq \left\| A^k \right\|_M$$

und damit $r(A) \leq \left\| A^k \right\|^{\frac{1}{k}}$. Sei nun $\epsilon > 0$ beliebig und betrachte die Matrix $B := \frac{A}{r(A)+\epsilon}$. Dann gilt wieder wegen Satz 1.13

$$r(B) = \max_{\lambda \in \sigma(B)} |\lambda| = \max_{\lambda \in \sigma(A)} \frac{|\lambda|}{r(A)+\epsilon} = \frac{r(A)}{r(A)+\epsilon} < 1.$$

Mit Satz 1.15 erhält man daraus, dass $\lim_{k\to\infty} B^k = 0$ und damit

$$\lim_{k\to\infty} \frac{\|A^k\|_M}{(r(A)+\epsilon)^k} = \lim_{k\to\infty} \left\| B^k \right\|_M = 0.$$

Das bedeutet aber, dass es zu $\epsilon > 0$ einen Index $K_\epsilon \in \mathbb{N}$ gibt, sodass

$$\frac{\|A^k\|_M}{(r(A)+\epsilon)^k} < 1 \quad \text{für alle } k \geq K_\epsilon,$$

bzw. gleichbedeutend

$$\left\| A^k \right\|_M^{\frac{1}{k}} < r(A) + \epsilon \quad \text{für alle } k \geq K_\epsilon.$$

Insgesamt gilt also

$$r(A) \leq \left\| A^k \right\|^{\frac{1}{k}} < r(A) + \epsilon \quad \text{für alle } k \geq K_\epsilon$$

und da $\epsilon > 0$ beliebig war, folgt daraus $r(A) = \lim_{k\to\infty} \left\| A^k \right\|^{\frac{1}{k}}$.
ad (iii): Die Aussage folgt nun unmittelbar aus der Darstellung des Spektralradius in (ii) über die Zeilensummennorm (vgl. Beispiel 1.5), denn für $0 \leq A \leq B$ gilt sicher $A^k \leq B^k$ und damit

$$\left\| A^k \right\|_\infty \leq \left\| B^k \right\|_\infty,$$

$$\left\| A^k \right\|_\infty^{\frac{1}{k}} \leq \left\| B^k \right\|_\infty^{\frac{1}{k}}.$$

Daraus folgt

$$r(A) = \lim_{k\to\infty} \left\| A^k \right\|_\infty^{\frac{1}{k}} \leq \lim_{k\to\infty} \left\| B^k \right\|_\infty^{\frac{1}{k}} = r(B).$$

□

Kapitel 2

Positive Matrizen, Satz von Perron

In diesem Abschnitt behandeln wir ausschließlich positive Matrizen. Wie sich herausstellen wird, liefert der Satz von Perron eine sehr gute Beschreibung ihres Spektrums. Wir werden sehen, dass der Spektralradius $r(A)$ ein Eigenwert mit algebraischer Vielfachheit 1 ist und dass es einen positiven Eigenvektor zu diesem Eigenwert gibt. Außerdem besagt der Satz von Perron, dass es keine weiteren positiven Eigenvektoren geben kann und dass jeder andere Eigenwert betragsmäßig kleiner als $r(A)$ ist. Die folgende grundlegende Bemerkung wird die Situation um einiges vereinfachen.

Bemerkung 2.1. Für $A \in \mathbb{R}^{n\times n}$, $A > 0$ gilt stets $r(A) > 0$. Um das einzusehen, betrachte die Jordan'sche Normalform von A, d.h. $A = PJP^{-1}$ für ein reguläres $P \in \mathbb{C}^{n\times n}$. Nehmen wir an, dass $\sigma(A) = \{0\}$. Dann wäre J nilpotent, d.h. es wäre $J^n = 0$ für hinreichend großes $n \in \mathbb{N}$. Wegen

$$A^n = PJ^nP^{-1} \quad \text{für alle } n \in \mathbb{N}$$

wäre aber auch $A^n = 0$ für n hinreichend groß. Da aber die Einträge von A^n endliche Summen von endlichen Produkten der Einträge von A sind, ist das ein Widerspruch zu $a_{ij} > 0$ für alle $i, j = 1, \dots, n$. Also muss $r(A) > 0$ gelten.

Betrachte nun die Matrix $\tilde{A} := \frac{1}{r(A)}A$. Laut Satz 1.13 gilt dann $r(\tilde{A}) = 1$ und alle Resultate über Eigenwerte und Eigenvektoren von $\tilde{A}$, die wir in den folgenden Abschnitten erhalten werden, können direkt auf A übertragen werden. Daher können wir im Weiteren o.B.d.A. annehmen, dass $r(A) = 1$. //

Bemerkung 2.2. Seien $A \in \mathbb{R}^{n\times n}$, $x, y \in \mathbb{R}^n$. Es gelten folgende Implikationen:

$$A > 0,\ x \geq 0,\ x \neq 0 \quad \Longrightarrow \quad Ax > 0 \tag{2.1}$$
$$A \geq 0,\ x \geq y \geq 0 \quad \Longrightarrow \quad Ax \geq Ay \tag{2.2}$$
$$A \geq 0,\ x > 0,\ Ax = 0 \quad \Longrightarrow \quad A = 0 \tag{2.3}$$
$$A > 0,\ x > y > 0 \quad \Longrightarrow \quad Ax > Ay \tag{2.4}$$

Um (2.1) einzusehen, seien $A > 0$ und $x \geq 0$, $x \neq 0$. Dann gibt es ein $j_0 \in \{1, \dots, n\}$ mit $x_{j_0} > 0$. Damit gilt für beliebiges $i = 1, \dots, n$

$$(Ax)_i = \sum_{j=1}^{n} a_{ij}x_j \geq a_{ij_0}x_{j_0} > 0,$$

also $Ax > 0$. Seien nun $A \geq 0$ und $x \geq y \geq 0$. Dann ist wegen $x_j \geq y_j$ für $j \in \{1, \ldots, n\}$

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j \geq \sum_{j=1}^{n} a_{ij} y_j = (Ay)_i$$

für alle $i = 1, \ldots, n$ und damit $Ax \leq Ay$ und (2.2) gezeigt. Für die dritte Implikation seien $A \geq 0$, $x > 0$ und $Ax = 0$. Das liefert für alle $i = 1, \ldots, n$

$$0 = (Ax)_i = \sum_{j=1}^{n} a_{ij} x_j$$

und da jeder Summand nichtnegativ ist, muss

$$a_{ij} x_j = 0 \qquad \text{für alle } j = 1, \ldots, n$$

gelten. Wegen $x > 0$ impliziert das aber $A = 0$. Schließlich seien $A > 0$ und $x > y > 0$. Dann gilt für jedes $i = 1, \ldots, n$

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j > \sum_{j=1}^{n} a_{ij} y_j = (Ay)_i,$$

da $a_{ij} > 0$ und $x_j > y_j$ für $i, j = 1, \ldots, n$. Das bedeutet aber genau $Ax > Ay$.

Mit der Dreiecksungleichung gilt für $A \in \mathbb{C}^{n \times n}$ und $z \in \mathbb{C}^n$, dass

$$(|Az|)_i = |\sum_{j=1}^{n} a_{ij} x_j| \leq \sum_{j=1}^{n} |a_{ij}||x_j| = (|A||z|)_i$$

für alle $i = 1, \ldots, n$ und damit

$$|Az| \leq |A||z|. \tag{2.5}$$

//

Der Beweis des Satzes von Perron ist eher länglich. Deswegen beweisen wir die Aussagen in mehreren Lemmata, die dann am Ende dieses Abschnittes wiederholend in einem Satz zusammengefasst sind.

Das nächste Lemma liefert uns die erste wesentliche Aussage über das Spektrum einer positiven Matrix, nämlich dass der Spektralradius ein Eigenwert ist und es einen positiven Eigenvekor dazu gibt.

Lemma 2.3. *Sei $A \in \mathbb{R}^{n \times n}$, $A > 0$. Dann gelten folgende Aussagen:*

(i). $r(A) \in \sigma(A)$

(ii). $x \in \mathbb{C}^n \backslash \{0\}$, $\lambda \in \mathbb{C}$ *mit* $|\lambda| = r(A)$ *und* $Ax = \lambda x \quad \Rightarrow \quad (\ A|x| = r(A)|x| \wedge |x| > 0\ )$

A hat also ein Eigenpaar $(r(A), y) \in \mathbb{R} \times \mathbb{R}^n$ mit $y > 0$.

Beweis. Laut Bemerkung 2.1 können wir o.B.d.A. annehmen, dass $r(A) = 1$. Aussage (i) folgt unmittelbar aus (ii). Sei also $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ ein beliebiges Eigenpaar von A mit $|\lambda| = r(A) = 1$ und $x \neq 0$. Dann gilt wegen (2.5) und wegen $A > 0$

$$|x| = |\lambda||x| = |\lambda x| = |Ax| \leq |A||x| = A|x|,$$

also
$$|x| \leq A|x|. \tag{2.6}$$
Ziel ist es, die Gleichheit zu zeigen, denn dann hätten wir einen Eigenvektor von A, nämlich
$$|x| = A|x| > 0,$$
zum Eigenwert $r(A) = 1$ gefunden und somit (ii) gezeigt. Setze $z := A|x|$ und $y := z - |x|$. Dann bedeutet Gleichung (2.6) genau, dass $y \geq 0$. Zu zeigen ist, dass $y = 0$. Nehmen wir also an, dass $y \neq 0$. Daraus folgt aber mit (2.1), dass $Ay > 0$ und, weil $|x| \neq 0$, ebenso $z = A|x| > 0$. Wegen $Ay > 0$ existiert ein $\epsilon > 0$, sodass
$$Ay = Az - A|x| = Az - z > \epsilon z$$
und somit
$$\frac{A}{1+\epsilon} z > z > 0.$$
Setzt man in dieser Ungleichung $B := \frac{A}{1+\epsilon}$, dann schreibt sie sich als $Bz > z$, und man erhält durch wiederholte Multiplikation mit B von links die Ungleichungen
$$B^2 z > Bz > z, \quad B^3 z > Bz > z, \quad \ldots$$
also
$$B^n z > z \quad \text{für alle } n \in \mathbb{N}.$$
Laut Satz 1.16 (ii) und Satz 1.13 gilt aber
$$\lim_{n\to\infty} \|B^n\|^{\frac{1}{n}} = r(B) = r(\frac{A}{1+\epsilon}) = \frac{1}{1+\epsilon} r(A) = \frac{1}{1+\epsilon} < 1$$
und wegen Satz 1.15 folgt daraus
$$\lim_{n\to\infty} B^n = 0.$$
Also erhalten wir aus der obigen Ungleichung für $n \to \infty$, dass $0 \geq z$. Wir haben aber oben schon gesehen, dass $z > 0$ gilt, und das führt hier zu einem Widerspruch zu unserer Annahme $y \neq 0$. Es gilt also
$$0 = y = z - |x| = A|x| - |x|$$
und daher $A|x| = |x| > 0$. □

Als nächstes zeigen wir, dass $r(A)$ der einzige Eigenwert auf dem Kreis
$$\{\ z \in \mathbb{C} \mid |z| = r(A)\ \}$$
ist und dass $\text{index}(r(A)) = 1$ gilt. In der Jordan-Zerlegung treten also zum Eigenwert $r(A)$ nur Blöcke der Größe 1 auf, was gleichbedeutend ist damit, dass die algebraische und geometrische Vielfachheit von $r(A)$ übereinstimmen.

Lemma 2.4. *Sei $A \in \mathbb{R}^{n\times n}$, $A > 0$. Dann gilt:*

(i). *$r(A)$ ist der einzige Eigenwert von A, dessen Betrag gleich $r(A)$ ist, d.h. aus $\lambda \in \sigma(A)$, $|\lambda| = r(A)$ folgt $\lambda = r(A)$.*

(ii). $\text{index}(r(A)) = 1$*; insbesondere stimmen algebraische und geometrische Vielfachheit des Eigenwertes $r(A)$ überein.*

Beweis. Es sei wieder o.B.d.A. $r(A) = 1$ (siehe Bemerkung 2.1).
ad (i): Aus Lemma 2.3 (ii) wissen wir, dass für jedes Eigenpaar $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$ von A mit $|\lambda| = 1$ und $x \neq 0$ gilt, dass $0 < |x| = A|x|$. Also haben wir für ein beliebiges, aber festes $k \in \{1, \ldots, n\}$

$$0 < |x_k| = (A|x|)_k = \sum_{j=1}^{n} a_{kj}|x_j|.$$

Es gilt aber auch

$$|x_k| = |\lambda||x_k| = |(\lambda x)_k| = |(Ax)_k| = \left|\sum_{j=1}^{n} a_{kj}x_j\right|,$$

womit

$$\left|\sum_{j=1}^{n} a_{kj}x_j\right| = \sum_{j=1}^{n} a_{kj}|x_j| = \sum_{j=1}^{n}|a_{kj}x_j|. \tag{2.7}$$

Bekannterweise gilt für Vektoren $z_1, \ldots, z_n \in \mathbb{C}^m \backslash \{0\}$ die Gleichheit in der Dreiecksungleichung, also

$$\left\|\sum_{j=1}^{n} z_j\right\|_2 = \sum_{j=1}^{n} \|z_j\|_2$$

genau dann, wenn es $\alpha_2, \ldots, \alpha_n \in (0, +\infty)$ gibt, sodass $z_j = \alpha_j z_1$ für alle $j = 2, \ldots, n$. Mit $m = 1$ gilt diese Aussage speziell für Skalare aus $\mathbb{C}$. Damit erhalten wir aus Gleichung (2.7), dass es zu festem $k \in \{1, \ldots, n\}$ Zahlen $\alpha_j(k) > 0, j = 2, \ldots, n$ gibt, sodass

$$a_{kj}x_j = \alpha_j(k)(a_{k1}x_1),$$

bzw. äquivalent dazu,

$$x_j = \pi_j(k)x_1 \quad \text{mit } \pi_j(k) = \frac{\alpha_j(k)a_{k1}}{a_{kj}} > 0.$$

Für jedes Eigenpaar (λ, x) mit $|\lambda| = 1$ und $x \neq 0$ gilt also $x = x_1 p(k)$ mit

$$p(k) = (1, \pi_2(k), \ldots, \pi_n(k))^T > 0.$$

Wegen $\lambda x = Ax$ folgt

$$\lambda p(k) = Ap(k) = |Ap(k)| = |\lambda p(k)| = |\lambda|p(k) = p(k).$$

Also gilt für jeden Eigenwert λ mit $|\lambda| = 1$ schon $\lambda = 1$ und somit ist (i) gezeigt.
ad (ii): Sei nun angenommen, dass $\operatorname{index}(r(A)) = m > 1$. Betrachte die Jordan'sche Normalform von A, $A = PJP^{-1}$ und $P \in \mathbb{C}^{n\times n}$ regulär. Dann bedeutet $\operatorname{index}(r(A)) = m > 1$, dass J einen $m \times m$-Block

$$\widetilde{J} := J_m(r(A))$$

zu $r(A) = 1$ enthält, vgl. Satz 1.10. Wegen (1.3) gilt

$$\widetilde{J}^k = \begin{pmatrix} 1 & \binom{k}{1} & \binom{k}{2} & \cdots & \binom{k}{m-1} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \binom{k}{2} \\ \vdots & & \ddots & \ddots & \binom{k}{1} \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{C}^{m\times m},$$

wobei wieder $\binom{k}{j} = 0$ für $j > k$ gesetzt wird. An dieser Darstellung von $\widetilde{J}^k$ erkennt man, dass zumindest

$$\left\| \widetilde{J}^k \right\|_\infty \geq \binom{k}{1} = k$$

gelten muss, was

$$\lim_{k\to\infty} \left\| \widetilde{J}^k \right\|_\infty = \infty$$

und damit natürlich auch

$$\lim_{k\to\infty} \left\| J^k \right\|_\infty = \infty$$

zur Folge hat. Wegen

$$\left\| J^k \right\|_\infty = \left\| P^{-1} A^k P \right\|_\infty \leq \left\| P^{-1} \right\|_\infty \left\| A^k \right\|_\infty \|P\|_\infty$$

gilt

$$\left\| A^k \right\|_\infty \geq \frac{\left\| J^k \right\|_\infty}{\left\| P^{-1} \right\|_\infty \|P\|_\infty}$$

und somit auch $\lim_{k\to\infty} \left\| A^k \right\|_\infty = \infty$. Für $k \in \mathbb{N}$ schreibe nun $A^k = (a_{ij}^k)_{i,j=1,\ldots,n}$ und sei $i_k \in \{1, \ldots, n\}$ jener Zeilenindex, sodass

$$\left\| A^k \right\|_\infty = \sum_{j=1}^{n} a_{i_k j}^k.$$

Aus Lemma 2.3 wissen wir, dass es ein $p \in \mathbb{R}^n$ gibt, $p > 0$, mit $p = Ap$ und daher auch $p = A^k p$ für alle $k \in \mathbb{N}$. Für so einen Eigenvektor p gilt nun

$$\|p\|_\infty \geq p_{i_k} = (A^k p)_{i_k} = \sum_{j=1}^{n} a_{i_k j}^k p_j \geq \left(\sum_{j=1}^{n} a_{i_k j}^k \right) (\min_{i=1,\ldots,n} p_i) = \left\| A^k \right\|_\infty (\min_{i=1,\ldots,n} p_i),$$

wobei die rechte Seite für $k \to \infty$ gegen ∞ konvergiert Wir erhalten also im Grenzübergang einen Widerspruch, da p ein konstanter Vektor in $\mathbb{R}^n$ ist. Also ist $\operatorname{index}(A) = 1$, womit (ii) gezeigt ist. $\square$

Nun wissen wir bereits, dass algebraische und geometrische Vielfachheit von $r(A)$ als Eigenwert von A übereinstimmen. Das nächste Resultat sagt aus, dass diese Vielfachheiten beide gleich 1 sind, womit der zu $r(A)$ gehörige Eigenraum eindimensional ist.

Lemma 2.5. *Sei $A \in \mathbb{R}^{n\times n}$, $A > 0$. Dann ist die algebraische Vielfachheit, und somit auch die geometrische Vielfachheit des Eigenwertes $r(A)$ gleich 1. $r(A)$ ist also ein einfacher Eigenwert von A und es gilt*

$$\dim(\ker(A - r(A)I)) = 1.$$

Beweis. Sei wieder o.B.d.A. $r(A) = 1$ (vgl.Bemerkung 2.1). Angenommen, $r(A)$ hat algebraische Vielfachheit $m > 1$. Wir wissen bereits aus Lemma 2.4, dass die geometrische Vielfachheit von $r(A)$ mit der algebraischen übereinstimmt. Also gibt es $m > 1$ linear unabhängige Eigenvektoren zum Eigenwert $r(A) = 1$. Sind x und y zwei linear unabhängige Eigenvektoren zum Eigenwert 1, so gilt $x \neq \alpha y$ für alle $\alpha \in \mathbb{C}$. Sei nun $i \in \{1, \ldots, n\}$ so, dass $y_i \neq 0$, und setze

$$z := x - \frac{x_i}{y_i} y.$$

Dann gilt

$$Az = Ax - \frac{x_i}{y_i} Ay = z,$$

und aus Lemma 2.3 folgt, dass auch $A|z| = |z| > 0$. Das ist aber ein Widerspruch zu

$$z_i = x_i - \frac{x_i}{y_i} y_i = 0.$$

Damit haben wir gezeigt, dass $m = 1$ gelten muss. □

Bemerkung 2.6. Laut Satz 2.5 ist der Eigenraum zum Eigenwert $r(A)$ eindimensional, d.h. es gibt einen eindeutigen Eigenvektor

$$p \in \ker(A - r(A)I), \quad p > 0 \quad \text{und} \quad \sum_{j=1}^{n} p_j = 1.$$

Es gilt außerdem $A > 0$ genau dann, wenn $A^T > 0$ und auch $r(A) = r(A^T)$. Die bisherigen Resultate kann man also analog auf A^T anwenden und erhält, dass es einen eindeutigen Vektor $q \in \mathbb{R}^n$ gibt mit

$$A^T q = r(A) q \quad (\text{ bzw. } q^T A = r(A) q^T \text{ }), \quad q > 0 \quad \text{und} \quad \sum_{j=1}^{n} q_j = 1.$$

//

Definition 2.7. Sei $A \in \mathbb{R}^{n \times n}$, $A > 0$. Dann heißt $r(A)$ die *Perron-Wurzel* von A. $p \in \mathbb{R}^n$ heißt *rechter Perron-Vektor* zu A, falls

$$Ap = r(A)p, \quad p > 0 \quad \text{und} \quad \sum_{j=1}^{n} p_j = 1.$$

$q^T \in \mathbb{R}^n$ heißt *linker Perron-Vektor* zu A, falls

$$q^T A = r(A) q^T, \quad q > 0 \quad \text{und} \quad \sum_{j=1}^{n} q_j = 1.$$

//

Laut Bemerkung 2.6 sind diese Vektoren eindeutig bestimmt und die Definition damit sinnvoll.

Das nächste Lemma liefert das erstaunliche Resultat, dass A außer dem Perron-Vektor und seinen positiven Vielfachen keine weiteren nichtnegativen Eigenvektoren mehr haben kann.

Lemma 2.8. *Sei $A \in \mathbb{R}^{n \times n}$, $A > 0$. Dann hat A außer dem rechten Perron-Vektor p und seinen positiven Vielfachen keine weiteren nichtnegativen Eigenvektoren, egal zu welchem Eigenwert.*

Beweis. Sei $(\lambda, y) \in \mathbb{C} \times \mathbb{R}^n$ ein beliebiges Eigenpaar von A mit $y \geq 0, y \neq 0$. Sei q der linke Perron-Vektor von A. Aus (2.1) folgt $q^T y > 0$ und wegen $r(A) q^T = q^T A$ auch

$$r(A) q^T y = q^T A y = \lambda q^T y.$$

Also muss $\lambda = r(A)$ gelten und damit $y > 0$ ein positives Vielfaches des rechten Perron-Vektors p sein. □

Das folgende Lemma liefert einen hilfreichen Zusammenhang zwischen dem Spektralradius und den Einträgen der Matrix A.

Lemma 2.9 (Collatz-Wielandt Formel). *Sei $A \in \mathbb{R}^{n\times n}$, $A > 0$. Dann ist die Perron-Wurzel gegeben durch $r(A) = \max_{x\in\mathcal{N}} f(x)$, wobei*

$$f(x) = \min_{\substack{1\leq i\leq n\\ x_i\neq 0}} \frac{(Ax)_i}{x_i}$$

$$\mathcal{N} = \{\ x \in \mathbb{R}^n \mid x \geq 0 \textit{ und } x \neq 0\ \}$$

Beweis. Wir zeigen zuerst, dass $f(x) \leq r(A)$ für alle $x \in \mathcal{N}$. Dazu sei $x \in \mathcal{N}$ beliebig und setze $\xi := f(x) > 0$. Dann gilt

$$0 \leq \xi x \leq Ax,$$

wobei die erste Ungleichung wegen $x \in \mathcal{N}$ und die zweite aus der Definition von f folgt. Seien p und q der jeweils rechte und linke Perron-Vektor von A. Wegen (2.1) gilt dann $q^T x > 0$. Damit und mit (2.2) folgt aus $\xi x \leq Ax$

$$0 < \xi q^T x \leq q^T Ax = r(A) q^T x.$$

Also gilt $\xi \leq r(A)$ und wir erhalten $f(x) \leq r(A)$ für alle $x \in \mathcal{N}$.
Nachdem aber $f(p) = r(A)$ und $p \in \mathcal{N}$ gilt, haben wir $r(A) = \max_{x\in\mathcal{N}} f(x)$. □

Nun haben wir alle Aussagen des Satzes von Perron bewiesen. Zusammenfassend sind sie in dem nächsten Satz noch einmal aufgelistet.

Satz 2.10 (Perron). *Sei $A \in \mathbb{R}^{n\times n}$ mit $A > 0$. Für $r := r(A)$ gelten die folgenden Aussagen:*

(i). $r > 0$

(ii). $r \in \sigma(A)$ *(r heißt die Perron-Wurzel von A)*

(iii). *r ist der einzige Eigenwert von A, dessen Betrag gleich $r(A)$ ist.*

(iv). *r ist ein einfacher Eigenwert von A, d.h. die algebraische Vielfachheit von r ist 1.*

(v). *Es gibt einen Eigenvektor $x \in \mathbb{R}^n$ von A zum Eigenwert r mit $x > 0$.*

(vi). *Es gibt einen eindeutigen Eigenvektor $p \in \mathbb{R}^n$ von A mit*

$$Ap = rp, \quad p > 0, \quad \|p\|_1 = \sum_{j=1}^{n} p_j = 1$$

(p heißt rechter Perron-Vektor zu A).
Außer allen positiven Vielfachen von p hat A keine weiteren nichtnegativen Eigenvektoren.

(vii). *(Collatz-Wielandt Formel) Es ist*

$$r = \max_{x\in\mathcal{N}} f(x)$$

mit

$$f(x) = \min_{\substack{1\leq i\leq n\\ x_i\neq 0}} \frac{(Ax)_i}{x_i}$$

$$\mathcal{N} = \{\ x \in \mathbb{R}^n \mid x \geq 0 \textit{ und } x \neq 0\ \}$$

Kapitel 3

Nichtnegative Matrizen, Satz von Perron-Frobenius

In diesem Abschnitt wollen wir die Voraussetzungen abschwächen und auch Matrizen $A \in \mathbb{R}^{n\times n}$, die nur $A \geq 0$ erfüllen, zulassen. Ziel dieses Teils der Arbeit ist es, möglichst viele Aussagen des vorigen Abschnittes auch unter diesen schwächeren Voraussetzungen zu beweisen. Das wird uns allerdings nicht vollständig gelingen, die Eigenschaft (iii) aus Satz 2.10 wird gänzlich verloren gehen und um die Eigenschaften (i), (iv), (v) und (vi) zu zeigen, werden wir die Voraussetzungen noch verschärfen müssen, indem wir die Irreduzibilität der Matrix A fordern.

Das nächste Lemma zeigt, dass die Eigenschaften (ii) und (vii) aus Satz 2.10 erhalten bleiben, wenn man zu nichtnegativen Matrizen übergeht und dass man zumindest die Existenz eines nichtnegativen Eigenvektors zum Eigenwert $r(A)$ nachweisen kann.

Lemma 3.1. *Sei $A \in \mathbb{R}^{n\times n}$, $A \geq 0$. Dann gilt:*

(i). $r(A) \in \sigma(A)$ *($r(A) = 0$ nicht ausgeschlossen)*

(ii). $Az = r(A)z$ *für ein* $z \in \mathcal{N} = \{\ x \in \mathbb{R}^n \mid x \geq 0$ *und* $x \neq 0\ \}$

(iii). $r(A) = \max_{x\in\mathcal{N}} f(x)$, *wobei*

$$f(x) = \min_{\substack{1\leq i\leq n \\ x_i\neq 0}} \frac{(Ax)_i}{x_i}$$

Beweis. ad (i): Sei $E = (e_{ij})$ mit $e_{ij} = 1$ für alle $i, j = 1, \ldots, n$, und betrachte die Folge $A_k = A + \frac{1}{k}E$, $k \in \mathbb{N}$. Dann gilt

$$A_k \in \mathbb{R}^{n\times n}, \quad A_k > 0 \quad \text{für alle } k \in \mathbb{N}.$$

Sind nun r_k und p_k jeweils die Perron-Wurzel und der rechte Perron-Vektor zu A_k, dann gilt

$$r_k > 0, \quad p_k > 0 \quad \text{und} \quad \|p_k\|_1 = 1 \quad \text{für alle } k \in \mathbb{N}.$$

Die Folge $(p_k)_{k\in\mathbb{N}}$ ist also in der Einheitssphäre des $\mathbb{R}^n$ enthalten, die nach dem Satz von Bolzano-Weierstrass kompakt ist. Also gibt es eine konvergente Teilfolge $(p_{k_i})_{i\in\mathbb{N}}$ von $(p_k)_{k\in\mathbb{N}}$ mit $\lim_{i\to\infty} p_{k_i} = z$ für ein $z \in \mathbb{R}^n$. Wegen $\lim_{i\to\infty}(p_{k_i})_j = z_j$ für alle $j = 1, \ldots, n$ folgt $z_j \geq 0$ für $j = 1, \ldots, n$, also $z \geq 0$, und wegen $\|p_k\|_1 = 1$ für alle $k \in \mathbb{N}$ gilt auch $\|z\|_1 = 1$; insbesondere $z \neq 0$. Wegen

$$A_1 > A_2 > A_3 > \cdots > A \geq 0$$

und Satz 1.16 (iii) folgt

$$r_1 \geq r_2 \geq r_3 \geq \cdots \geq r(A)$$

Also ist die Folge $(r_k)_{k\in\mathbb{N}}$ eine monoton fallende und durch $r(A)$ nach unten beschränkte Folge in $\mathbb{R}$. Also ist sie gegen ein $r^* \geq r(A)$ konvergent. Insbesondere konvergiert auch die Teilfolge $(r_{k_i})_{i\in\mathbb{N}}$ gegen r^*. Es gilt klarerweise $\lim_{k\to\infty} A_k = A$ und daher auch $A_{k_i} \to A$ für $i \to \infty$. Also konvergiert auch $A_{k_i}p_{k_i}$, und zwar gegen Az und es gilt

$$Az = \lim_{i\to\infty} A_{k_i}p_{k_i} = \lim_{i\to\infty} r_{k_i}p_{k_i} = r^*z,$$

womit $r^* \in \sigma(A)$ und daher $r^* \leq r(A)$. Insgesamt erhalten wir $r^* = r(A)$, $Az = r(A)z$ und $z \geq 0$, $z \neq 0$. Damit sind (i) und gleichzeitig auch (ii) bewiesen.
ad (iii): Betrachte wieder die Folge $(A_k)_{k\in\mathbb{N}}$ und die dazugehörige Folge der Perron-Wurzeln $(r_k)_{k\in\mathbb{N}}$. Betrachte außerdem die Folge der linken Perron-Vektoren $(q_k^T)_{k\in\mathbb{N}}$, d.h.

$$q_k^T A_k = r(A_k)q_k^T, \quad q_k > 0 \quad \text{für alle } k \in \mathbb{N}.$$

Damit gilt für jedes $x \in \mathcal{N}$ wegen (2.2), dass

$$0 \leq f(x)x \leq Ax \leq A_kx$$

und daher

$$0 \leq f(x)q_k^T x \leq q_k^T A_k x = r_k q_k^T x,$$

also $f(x) \leq r_k$ für alle $k \in \mathbb{N}$. Daraus folgt wegen $r_k \to r^* = r(A)$ für $k \to \infty$, dass $f(x) \leq r(A)$ für alle $x \in \mathcal{N}$. Wegen $z \in \mathcal{N}$ und $f(z) = r(A)$ (siehe Beweisteil (i)) gilt daher

$$\max_{x\in\mathcal{N}} f(x) = r(A).$$

□

In Lemma 3.1 haben wir gesehen, welche Eigenschaften sich ohne Verschärfung der Voraussetzungen auf den Fall der nichtnegativen Matrix übertragen lassen. Weiter kann man ohne zusätzliche Voraussetzungen tatsächlich nicht gehen, wie das nächste Beispiel zeigt.

Beispiel 3.2. Betrachte die zwei nichtnegativen Matrizen

$$A_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

A_1 hat einen Eigenwert $\lambda_1 = 0$ mit algebraischer Vielfachheit 2, es ist also $\sigma(A_1) = \{0\}$ und damit auch $r(A_1) = 0$. Die geometrische Vielfachheit von λ_1 ist 1, der Eigenraum ist also eindimensional und ein Eigenvektor zum Eigenwert $\lambda_1 = 0$ wäre zum Beispiel

$$x_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

A_2 hat die Eigenwerte $\lambda_{2,1} = 1$ und $\lambda_{2,2} = -1$, klarerweise beide mit algebraischer und geometrischer Vielfachheit 1. Damit ist $\sigma(A_2) = \{-1, 1\}$ und $r(A_2) = 1$. Eigenvektoren zu $\lambda_{2,1}$ und $\lambda_{2,2}$ wären zum Beispiel

$$x_{2,1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad x_{2,2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Anhand von A_1 sieht man, dass die Eigenschaften (i), (iv) und (v) aus Satz 2.10 auf beliebige nichtnegative Matrizen sicher nicht übertragbar sind. Die Matrix A_2 zeigt, dass man (iii) auch nicht retten kann. //

Beispiel 3.3. Betrachte die nichtnegativen Matrizen

$$A_3 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

A_3 hat den Eigenwert $\lambda_3 = 1$ mit algebraischer Vielfachheit 2. Daher ist $\sigma(A_3) = \{1\}$ und $r(A_3) = 1$. Der Eigenraum zu λ_3 ist eindimensional, λ_3 hat also geometrische Vielfachheit 1 und ein Eigenvektor ist zum Beispiel

$$x_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A_4 hat die Eigenwerte $\lambda_{4,1} = \frac{1+\sqrt{5}}{2}$ und $\lambda_{4,2} = \frac{1-\sqrt{5}}{2}$, wobei geometrische und algebraische Vielfachheit jeweils 1 betragen. Es gilt $\sigma(A_4) = \{\frac{1-\sqrt{5}}{2}, \frac{1+\sqrt{5}}{2}\}$ und $r(A_4) = \frac{1+\sqrt{5}}{2}$. Eigenvektoren zu $\lambda_{4,1}$ und $\lambda_{4,2}$ wären zum Beispiel

$$x_{4,1} = \begin{pmatrix} 2 \\ \sqrt{5} - 1 \end{pmatrix} \quad \text{und} \quad x_{4,2} = \begin{pmatrix} 1 - \sqrt{5} \\ 2 \end{pmatrix}.$$

Für A_4 sind (iv) und (v) aus Satz 2.10 erfüllt, für A_3 allerdings nicht. Das lässt vermuten, dass vielleicht die Positionen der Nulleinträge eine wichtige Rolle spielen und dass wir unter geeigneten Voraussetzungen noch weitere Eigenschaften aus Satz 2.10 zeigen können. //

Es stellt sich heraus, dass tatsächlich nicht die Nulleinträge an sich das Problem darstellen, sondern die Positionen der Nullen. Um diese Vermutung zu konkretisieren und eine sinnvolle Verschärfung der Voraussetzungen möglich zu machen, brauchen wir zunächst einige Definitionen und Hilfsresultate.

Definition 3.4. Sei $A \in \mathbb{R}^{n \times n}$.

- $P \in \mathbb{R}^{n \times n}$, $P = (p_{ij})_{i,j=1}^n$, heißt *Permutationsmatrix*, falls

  $$p_{ij} \in \{0, 1\} \quad \text{für alle } i, j = 1, \ldots, n$$

  und in jeder Zeile und jeder Spalte von P genau ein Einser steht.

- A heißt *reduzibel*, falls es eine Permutationsmatrix P gibt, sodass

  $$P^T A P = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix},$$

  wobei $X \in \mathbb{R}^{m \times m}$, $Y \in \mathbb{R}^{k \times k}$ mit $m + k = n$.

- A heißt *irreduzibel*, wenn A nicht reduzibel ist.

//

Bemerkung 3.5. Sei $A \in \mathbb{R}^{n\times n}$, $P \in \mathbb{R}^{n\times n}$ eine Permutationsmatrix und setze

$$B := P^T AP.$$

Dann gibt es eine zu P gehörige Permutation $\sigma \in S_n$, sodass $B \in \mathbb{R}^{n\times n}$ genau jene Matrix ist, die durch Permutation der Zeilen und Spalten gemäß der Permutaion σ aus A hervorgeht. Genauer gilt

$$b_{ij} = a_{\sigma^{-1}(i)\sigma^{-1}(j)} \quad \text{für alle } i, j = 1, \ldots, n.$$

//

Das nächste Lemma zeigt, dass die Irreduzibilität einer Matrix eng mit der Position ihrer Nulleinträge zusammenhängt.

Lemma 3.6. *Sei $A \in \mathbb{R}^{n\times n}$. Dann sind folgende Aussagen äquivalent:*

(i). *A ist irreduzibel.*

(ii). *Für alle Indexpaare $(i, j) \in \{1, \ldots, n\}^2$ gibt es ein $l \in \mathbb{N}$ und eine Abfolge von Indizes*

$$k_2, \ldots, k_{l-1} \in \{1, \ldots, n\},$$

sodass mit $k_1 = i$, $k_l = j$ immer

$$a_{k_s k_{s+1}} \neq 0, \quad s = 1, \ldots, l-1$$

gilt.

Beweis. (i) $\Rightarrow$ (ii): Angenommen, es gibt ein Indexpaar $(i_0, j_0) \in \{1, \ldots, n\}^2$, für das es keine solche Indexabfolge gibt. Wir zeigen, dass A dann reduzibel sein muss. Sei

$$I_- := \{h_1, \ldots, h_r\}, \quad r \in \mathbb{N}$$

die Menge aller Indizes h, sodass es keine solche Indexabfolge von i_0 nach $h \in I_-$ gibt. Dabei seien die h_s, $s = 1, \ldots, r$, so gewählt, dass $h_1 = j_0$. Weiters sei

$$I_+ := \{h_{r+1}, \ldots, h_{n-1}\}$$

die Menge aller Indizes $h \neq i_0$, sodass es eine solche Indexabfolge von i_0 nach $h \in I_+$ gibt. Dann kann es keine Indexabfolge von $i \in I_+$ nach $j \in I_-$ geben. Denn angenommen, es gibt eine solche Abfolge

$$i = k_1, \ldots, k_l = j, \quad l \in \mathbb{N}$$

von i nach j und sei

$$i_0 = l_1, \ldots, l_m = i, \quad m \in \mathbb{N}$$

die wegen $i \in I_+$ existente Abfolge von i_0 nach i. Dann ist klarerweise

$$i_0 = l_1, \ldots, l_m = i = k_1, \ldots, k_l = j$$

eine Indexabfolge von i_0 nach j, was wegen $j \in I_-$ nicht möglich ist. Betrachte nun $B := P^T AP$, wobei $P \in \mathbb{R}^{n\times n}$ jene Permutationsmatrix sei, sodass

$$b_{ij} = a_{\sigma^{-1}(i)\sigma^{-1}(j)} \quad \text{für alle } i, j = 1, \ldots, n \tag{3.1}$$

mit

$$\sigma(\overbrace{h_1}^{=j_0}) = 1, \quad \ldots, \quad \sigma(h_r) = r,$$

$$\sigma(h_{r+1}) = r + 1, \quad \ldots, \quad \sigma(h_{n-1}) = n - 1, \quad \sigma(i_0) = n.$$

Dann gibt es für die Matrix B keine Abfolge von Indizes von $i \in \{r+1, \ldots, n\}$ nach $j \in \{1, \ldots, r\}$. Denn angenommen, es gäbe eine Indexabfolge von $i \in \{r + 1, \ldots, n\}$ nach $j \in \{1, \ldots, r\}$, d.h. es gäbe Indizes

$$i = k_1, \ldots, k_l = j, \quad l \in \mathbb{N}$$

mit

$$b_{k_s k_{s+1}} \neq 0, \quad s = 1, \ldots, l - 1.$$

Wegen (3.1) ist das aber gleichbedeutend mit

$$a_{\sigma^{-1}(k_s)\sigma^{-1}(k_{s+1})} \neq 0, \quad s = 1, \ldots, l - 1,$$

was bedeuten würde, dass es für A eine Abfolge von Indizes von $\sigma^{-1}(i)$ nach $\sigma^{-1}(j)$ geben würde. Das ist aber wegen $\sigma^{-1}(i) \in I_+ \cup \{i_0\}$ und $\sigma^{-1}(j) \in I_-$ nicht möglich. Es gibt also für B keine Abfolge von solchen Indizes von $i \in \{r + 1, \ldots, n\}$ nach $j \in \{1, \ldots, r\}$, was klarerweise $b_{ij} = 0$ für $i \in \{r + 1, \ldots, n\}$ und $j \in \{1, \ldots, r\}$ impliziert. B hat also die Gestalt

$$B = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$$

mit $X \in \mathbb{R}^{r\times r}$, $Z \in \mathbb{R}^{(n-r)\times(n-r)}$ und $Y \in \mathbb{R}^{r\times(n-r)}$. Das bedeutet aber, dass A reduzibel ist.
(ii) ⇒ (i): Sei $P \in \mathbb{R}^{n\times n}$ eine Permutationsmatrix mit

$$P^T AP = B = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}, \quad X \in \mathbb{R}^{r\times r}, \quad Z \in \mathbb{R}^{(n-r)\times(n-r)}, \quad Y \in \mathbb{R}^{r\times(n-r)}$$

und

$$b_{ij} = a_{\sigma^{-1}(i)\sigma^{-1}(j)} \quad \text{für alle } i, j = 1, \ldots, n$$

für eine zugehörige Permutation $\sigma \in S_n$. Angenommen, es gibt für $i > r$ und $j \leq r$ Indizes

$$i = k_1, \ldots, k_l = j, \quad l \in \mathbb{N}$$

mit

$$b_{k_s k_{s+1}} \neq 0, \quad s = 1, \ldots, l - 1.$$

Da $i = k_1 > r$ ist, gilt $b_{ik} = 0$ für alle $k \leq r$. Da aber $b_{ik_2} \neq 0$ gilt, muss $k_2 > r$ sein. Verfährt man so weiter, erhält man

$$k_m > r \quad \text{für alle } m = 1, \ldots, l.$$

Das führt aber wegen $k_l = j \leq r$ zu einem Widerspruch. B erfüllt also nicht die Eigenschaft (ii), d.h. es gibt ein Indexpaar (i_0, j_0), sodass es keine Abfolge von i_0 nach j_0 gibt. Dann gibt es aber für die Matrix A ein Indexpaar, nämlich $(\sigma^{-1}(i_0), \sigma^{-1}(j_0))$, das keine Indexabfolge verbindet. Denn angenommen, es gäbe Indizes

$$\sigma^{-1}(i_0) = k_1, \ldots, k_l = \sigma^{-1}(j_0), \quad l \in \mathbb{N},$$

sodass

$$a_{k_s k_{s+1}} \neq 0, \quad s = 1, \ldots, l - 1.$$

Damit wäre aber auch

$$b_{\sigma(k_s)\sigma(k_{s+1})} \neq 0, \quad s = 1, \dots, l-1,$$

was aber ein Widerspruch zu unserer Annahme ist. A erfüllt also nicht die Eigenschaft (ii). □

Das folgende Lemma ist der Schlüssel zum Beweis der weiteren Aussagen.

Lemma 3.7. *Sei $A \in \mathbb{R}^{n\times n}$, $A \geq 0$ irreduzibel. Dann ist $(I+A)^{n-1} > 0$.*

Beweis. Da A irreduzibel ist, gibt es laut Lemma 3.6 zu jedem Paar $(i,j) \in \{1,\dots,n\}^2$ eine Abfolge von Indizes

$$i = k_1, \dots, k_l = j, \quad l \in \mathbb{N}$$

mit

$$a_{k_s k_{s+1}} \neq 0, \quad s = 1, \dots, l-1.$$

Nun gilt für die kleinste solche Zahl l sicher $l \leq n$. Denn falls $l > n$ gilt, gibt es einen Index i_0, der in der Abfolge doppelt vorkommt,

$$i = k_1, \dots, k_{m-1}, k_m = i_0, k_{m+1}, \dots, k_{p-1}, k_p = i_0, k_{p+1}, \dots, k_l = j$$

mit $m, p \in \{1, \dots, l\}$ und die Abfolge

$$i = k_1, \dots, k_{m-1}, i_0, k_{p+1}, \dots, k_l = j$$

wäre kürzer. Man kann also zu einem Indexpaar (i,j) eine Indexabfolge der Länge $l \leq n$ finden. Sei nun $A^k = (a_{ij}^k)$ die k-te Potenz von A. Man erkennt aus

$$(A^2)_{ij} = \sum_{k=1}^{n} a_{ik} a_{kj},$$

$$(A^3)_{ij} = \sum_{l=1}^{n} \left(\sum_{k=1}^{n} a_{ik} a_{kl} \right) a_{lj} = \sum_{l=1}^{n} \sum_{k=1}^{n} a_{ik} a_{kl} a_{lj},$$

$$\vdots$$

dass der (i,j)-Eintrag von A^k gegeben ist durch

$$a_{ij}^k = \sum_{h_1,\dots,h_{k-1}=1}^{n} a_{ih_1} a_{h_1 h_2} \cdots a_{h_{k-1} j}. \tag{3.2}$$

Wegen $A \geq 0$ gilt sicher $a_{ij}^k \geq 0$, und aus Gleichung (3.2) erkennt man, dass $a_{ij}^k > 0$ genau dann gilt, wenn es zu (i,j) Indizes $i = h_1, \dots, h_{k+1} = j$ gibt, sodass

$$a_{ih_2} > 0, \quad a_{h_2 h_3} > 0, \quad \dots, \quad a_{h_k j} > 0.$$

Wegen dem anfangs Gezeigten kann man zu jedem Paar (i,j) ein k finden, für das $a_{ij}^k > 0$ gilt, wobei dieses k die Ungleichung $k+1 \leq n$ bzw. $k \leq n-1$ erfüllt. Daraus folgt für jedes Paar (i,j), dass

$$\left((I+A)^{n-1}\right)_{ij} = \left(\sum_{k=0}^{n-1} \binom{n-1}{k} A^k \right)_{ij} = \sum_{k=0}^{n-1} \binom{n-1}{k} a_{ij}^k > 0,$$

womit wir $(I+A)^{n-1} > 0$ gezeigt haben. □

Mit diesem Hilfslemma können wir nun den Satz von Perron-Frobenius beweisen. Die Aussage des Satzes gleicht beinahe der des Satzes von Perron. Nur die Eigenschaft (iii) geht gänzlich verloren. Die entscheidende Voraussetzung, die man neben der Nichtnegativität fordern muss, ist die Irreduzibilität der Matrix A.

Satz 3.8 (Perron-Frobenius). *Sei $A \in \mathbb{R}^{n\times n}$, $A \geq 0$ irreduzibel. Für $r := r(A)$ gelten folgende Aussagen:*

(i). $r > 0$

(ii). $r \in \sigma(A)$ *(r heißt die Perron-Wurzel von A)*

(iii). *r ist ein einfacher Eigenwert von A, d.h. die algebraische Vielfachheit von r ist* 1

(iv). *Es gibt einen Eigenvektor $x \in \mathbb{R}^n$ von A zum Eigenwert r mit $x > 0$*

(v). *Es gibt einen eindeutigen Eigenvektor $p \in \mathbb{R}^n$ von A mit*

$$Ap = rp, \quad p > 0, \quad \|p\|_1 = \sum_{j=1}^{n} p_j = 1$$

(p heißt rechter Perron-Vektor zu A)
Außer allen positiven Vielfachen von p hat A keine weiteren nichtnegativen Eigenvektoren

(vi). *(Collatz-Wielandt Formel) Es ist*

$$r = \max_{x\in\mathcal{N}} f(x)$$

mit

$$f(x) = \min_{\substack{1\leq i\leq n\\ x_i\neq 0}} \frac{(Ax)_i}{x_i}$$

$$\mathcal{N} = \{\ x \in \mathbb{R}^n \mid x \geq 0 \text{ und } x \neq 0\ \}$$

Beweis. ad (ii): Siehe Lemma 3.1, (i).
ad (iii): Sei $B := (I + A)^{n-1} > 0$ wie in Lemma 3.7. Dann gilt laut Satz 1.13 für $r(B)$

$$r(B) = \max_{\mu\in\sigma(B)} |\mu| = \max_{\lambda\in\sigma(A)} |(1+\lambda)^{n-1}| = \max_{\lambda\in\sigma(A)} |1+\lambda|^{n-1} = \left(\max_{\lambda\in\sigma(A)} |1+\lambda|\right)^{n-1}.$$

Nun ist aber $\sigma(A)$ in der Kreisscheibe mit Radius r enthalten und es gilt $r \in \sigma(A)$. Es ist also $1 + \sigma(A)$ in der Kreisscheibe mit Radius $1 + r$ enthalten und $1 + r \in (1 + \sigma(A))$. Daher ist

$$\max_{\lambda\in\sigma(A)} |1+\lambda| = 1 + r$$

und somit

$$r(B) = (1 + r)^{n-1}.$$

Weil $B > 0$ gilt, hat laut Satz 2.10 $r(B) = (1 + r)^{n-1} \in \sigma(B)$ algebraische Vielfachheit 1. Laut Bemerkung 1.14 (ii) hat auch $r \in \sigma(A)$ algebraische Vielfachheit 1 und (iii) ist gezeigt.
ad (iv): Wegen Bemerkung 1.14 (i) gilt für jedes Eigenpaar $(\lambda, x) \in \mathbb{C} \times \mathbb{C}^n$, $x \neq 0$ von A, dass $((1 + \lambda)^{n-1}, x)$ ein Eigenpaar von B ist. Aus Lemma 3.1 (ii) wissen wir bereits, dass es ein $x \in \mathbb{R}^n$, $x \geq 0$ gibt, sodass x ein Eigenvektor von A zum Eigenwert r ist. Dann ist x aber auch ein Eigenvektor von B zum Eigenwert

$$(1 + r)^{n-1} = r(B).$$

Deswegen und wegen $B > 0$ sowie Satz 2.10 (v) muss x ein positives Vielfaches des rechten Perron-Vektors zu B sein. Insbesondere gilt $x > 0$.
ad (i): Angenommen, es gilt $r = 0$. Laut (iv) gibt es ein $x > 0$ mit

$$Ax = rx = 0x = 0.$$

Aus (2.3) würde aber dann $A = 0$ folgen und somit führt $r = 0$ auf einen Widerspruch.
ad (v): Siehe Beweis von Lemma 2.8.
ad (vi): Siehe Lemma 3.1. $\square$

Bemerkung 3.9. Der Schlüssel zum Erhalt von (i), (iv), (v) und (vi) aus Satz 2.10 ist also, dass A irreduzibel ist. In der Tat ist die Matrix A_4 aus Beispiel 3.3 irreduzibel, die Matrix A_3 hingegen reduzibel. //

Literaturverzeichnis

[HA] Hans Havlicek: *Lineare Algebra für Technische Mathematiker*, Heldermann Verlag, 2006.

[ME] Carl D: Meyer: *Matrix Analysis and Applied Linear Algebra*, Society for Industrial Mathematics, 2000.

[KA] Harald Woracek, Michael Kaltenbäck, Martin Blümlinger: *Funktionalanalysis*, Skriptum zur Vorlesung SS2009.