Numerische Lineare Algebra

Oliver Ernst

Professur Numerische Mathematik

Wintersemester 2017/18

Mathematik !
TU Chemnitz

Inhalt I

Inhalt

Inhalt

Sei $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Bevor wir $f(\boldsymbol{A})$ für (beinahe) beliebige komplexwertige Funktionen f definieren, erinnern wir an

$$p(\boldsymbol{A}) := \alpha_0 \boldsymbol{I} + \alpha_1 \boldsymbol{A} + \cdots + \alpha_m \boldsymbol{A}^m,$$

wenn $p(\lambda) = \alpha_0 + \alpha_1\lambda + \cdots \alpha_m \lambda^m \in \mathscr{P}_m$ ein Polynom ist, und fassen einige der bekannten Eigenschaften von $p(\boldsymbol{A})$ zusammen.

Satz 4.1 (Eigenschaften von Polynomen in einer Matrix)

Für jedes Polynom $p \in \mathscr{P}_m$ gilt:

(1) Sind $\boldsymbol{A}$ und $\boldsymbol{B} \in \mathbb{C}^{n\times n}$ ähnliche Matrizen, d.h. existiert eine invertierbare Matrix $\boldsymbol{T} \in \mathbb{C}^{n\times n}$ mit $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}$, dann sind auch $p(\boldsymbol{A})$ und $p(\boldsymbol{B})$ ähnlich, genauer: $p(\boldsymbol{A}) = \boldsymbol{T} p(\boldsymbol{B}) \boldsymbol{T}^{-1}$.

(2) Ist $\boldsymbol{A} = \operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s)$ blockdiagonal mit quadratischen Diagonalblöcken $\boldsymbol{A}_j$, dann besitzt auch $p(\boldsymbol{A}) = \operatorname{diag}(p(\boldsymbol{A}_1), p(\boldsymbol{A}_2), \ldots, p(\boldsymbol{A}_s))$ Blockdiagonalstruktur, die mit der von $\boldsymbol{A}$ übereinstimmt.

Satz 4.1 (Fortsetzung)

(3) Ist

$$\boldsymbol{J}(\lambda) := \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \in \mathbb{C}^{n\times n} \tag{4.1}$$

ein Jordan-Block und bezeichnet $p^{(j)}$ die j-te Ableitung von p, so ist

$$p(\boldsymbol{J}(\lambda)) = \begin{bmatrix} \frac{p(\lambda)}{0!} & \frac{p'(\lambda)}{1!} & \cdots & \frac{p^{(n-2)}(\lambda)}{(n-2)!} & \frac{p^{(n-1)}(\lambda)}{(n-1)!} \\ & \frac{p(\lambda)}{0!} & \cdots & \frac{p^{(n-3)}(\lambda)}{(n-3)!} & \frac{p^{(n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \vdots & \vdots \\ & & & \frac{p(\lambda)}{0!} & \frac{p'(\lambda)}{1!} \\ & & & & \frac{p(\lambda)}{0!} \end{bmatrix}.$$

- Jede quadratische Matrix $\boldsymbol{A}$ ist zu einer Matrix in Jordan-Form ähnlich, also zu einer Blockdiagonalmatrix $\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_1, \boldsymbol{J}_2, \dots, \boldsymbol{J}_s)$, deren Diagonalblöcke $\boldsymbol{J}_\ell$ die Form (4.1) besitzen.
- Für $\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1}$ folgt aus Satz 4.1 dann

$$p(\boldsymbol{A}) = \boldsymbol{T} \operatorname{diag}(p(\boldsymbol{J}_1), p(\boldsymbol{J}_2), \dots, p(\boldsymbol{J}_s))\, \boldsymbol{T}^{-1}$$

und wir erkennen, dass $p(\boldsymbol{A})$ vollständig durch die Werte

$$p^{(\nu)}(\lambda_\mu) \qquad (\mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1)$$

festgelegt ist. Dabei sind $\lambda_1, \lambda_2, \dots, \lambda_k$ die (verschiedenen) Eigenwerte von $\boldsymbol{A}$ und n_μ die Dimension des größten Jordan-Blocks, der zu λ_μ gehört.
- Mit diesen Bezeichnungen ist das Minimalpolynom von $\boldsymbol{A}$ durch

$$m_{\boldsymbol{A}}(\lambda) = \prod_{\mu=1}^{k} (\lambda - \lambda_\mu)^{n_\mu} \in \mathscr{P}_K, \quad \lambda_i \neq \lambda_j \text{ für } i \neq j,\ K = \sum_{\mu=1}^{k} n_\mu, \tag{4.2}$$

gegeben. Eine unmittelbare Folgerung aus Satz 4.1 ist also

Korollar 4.2

Die Matrix $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ besitze das Minimalpolynom $m_{\boldsymbol{A}}$ aus (4.2). Für zwei Polynome p und q gilt genau dann $p(\boldsymbol{A}) = q(\boldsymbol{A})$, wenn die folgenden K Bedingungen erfüllt sind:

$$p^{(\nu)}(\lambda_\mu) = q^{(\nu)}(\lambda_\mu), \qquad \mu = 1, 2, \ldots, k, \quad \nu = 0, 1, \ldots, n_\mu - 1.$$

Anders formuliert: $p(\boldsymbol{A}) = q(\boldsymbol{A})$ gilt genau dann, wenn $m_{\boldsymbol{A}}$ die Differenz $p - q$ teilt. Insbesondere ist $p(\boldsymbol{A}) = \boldsymbol{O}$ genau dann, wenn p ein Vielfaches von $m_{\boldsymbol{A}}$ ist.

Jetzt definieren wir $f(\boldsymbol{A})$ für allgemeinere Funktionen f.

Definition 4.3

Die Matrix $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ besitze das Minimalpolynom $m_{\boldsymbol{A}}$ aus (4.2). Wir sagen die Funktion $f : \mathbb{C} \supseteq D \to \mathbb{C}$ sei in A definiert, kürzer gefasst $f(\boldsymbol{A})$ ist definiert, wenn f an allen Eigenwerten $\lambda = \lambda_\mu$ von $\boldsymbol{A}$ definiert ist und dort Ableitungen bis zur Ordnung $n_\mu - 1$ besitzt.
Mit anderen Worten: $f(\boldsymbol{A})$ ist definiert, wenn die $K = \sum_{\mu=1}^{k} n_\mu$ Ableitungen

$$f^{(\nu)}(\lambda_\mu), \qquad \mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1,$$

existieren. Ist dann $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_1, \boldsymbol{J}_2, \dots, \boldsymbol{J}_k)$ eine Jordansche Normalform von $\boldsymbol{A}$, so setzen wir

$$\boldsymbol{F}_j = f(\boldsymbol{J}_j) := \begin{bmatrix} \frac{f(\lambda_\mu)}{0!} & \frac{f'(\lambda_\mu)}{1!} & \cdots & \frac{f^{(\nu-2)}(\lambda_\mu)}{(\nu-2)!} & \frac{f^{(\nu-1)}(\lambda_\mu)}{(\nu-1)!} \\ & \frac{f(\lambda_\mu)}{0!} & \cdots & \frac{f^{(\nu-3)}(\lambda_\mu)}{(\nu-3)!} & \frac{f^{(\nu-2)}(\lambda_\mu)}{(\nu-2)!} \\ & & \ddots & \vdots & \vdots \\ & & & \frac{f(\lambda_\mu)}{0!} & \frac{f'(\lambda_\mu)}{1!} \\ & & & & \frac{f(\lambda_\mu)}{0!} \end{bmatrix} \tag{4.3}$$

für jeden Jordan-Block $\boldsymbol{J}_\ell = \boldsymbol{J}(\lambda_\mu) \in \mathbb{C}^{\nu\times\nu}$ $(\ell = 1, 2, \dots, s)$ (vgl. (4.1)) von $\boldsymbol{A}$ ($\lambda_\mu \in \{\lambda_1, \lambda_2, \dots, \lambda_k\}$, $0 \le \nu \le n_\mu$), und definieren schließlich

$$f(\boldsymbol{A}) := \boldsymbol{T} \operatorname{diag}(\boldsymbol{F}_1, \boldsymbol{F}_2, \dots, \boldsymbol{F}_k)\, \boldsymbol{T}^{-1}.$$

Definition 4.3 von $f(\boldsymbol{A})$ ist unabhangig von der gewählten Jordan-Form:
Denn diese ist durch $\boldsymbol{A}$ bis auf eine Permutation der Blöcke eindeutig bestimmt.
Ist also

$$\tilde{\boldsymbol{J}} = \tilde{\boldsymbol{T}}^{-1}\boldsymbol{A}\tilde{\boldsymbol{T}}$$

eine weitere Jordansche Normalform von $\boldsymbol{A}$, dann gibt es eine Permutation σ von $\{1, 2, \ldots, s\}$ mit

$$\tilde{\boldsymbol{J}} = \operatorname{diag}(\boldsymbol{J}_{\sigma(1)}, \boldsymbol{J}_{\sigma(2)}, \ldots, \boldsymbol{J}_{\sigma(s)}) = \boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{\mathsf{T}},$$

wobei $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ die zu σ gehörige Blockpermutationsmatrix bezeichnet. Wegen $\tilde{\boldsymbol{T}}\boldsymbol{P} = \boldsymbol{T}$ folgt

$$\begin{aligned}
\tilde{\boldsymbol{T}} \operatorname{diag}(\boldsymbol{F}_{\sigma(1)}, \boldsymbol{F}_{\sigma(2)}, \ldots, \boldsymbol{F}_{\sigma(s)})\, \tilde{\boldsymbol{T}}^{-1} &= \tilde{\boldsymbol{T}}\boldsymbol{P} \operatorname{diag}(\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_s)\, \boldsymbol{P}^{\mathsf{T}}\, \tilde{\boldsymbol{T}}^{-1} \\
&= \boldsymbol{T} \operatorname{diag}(\boldsymbol{F}_1, \boldsymbol{F}_2, \ldots, \boldsymbol{F}_s)\, \boldsymbol{T}^{-1}.
\end{aligned}$$

- Ist $f(\lambda) = 1/\lambda$, so ist $f(\boldsymbol{A})$ genau dann definiert, wenn $\boldsymbol{A}$ invertierbar ist und es gilt $f(\boldsymbol{A}) = \boldsymbol{A}^{-1}$. Dies folgt sofort aus

$$\begin{aligned}\boldsymbol{J}(\lambda)^{-1} &= (\lambda \boldsymbol{I} + \boldsymbol{S}_n)^{-1} = \tfrac{1}{\lambda}\left(\boldsymbol{I} + \tfrac{1}{\lambda}\boldsymbol{S}_n\right)^{-1} \\ &= \tfrac{1}{\lambda}\sum_{j=0}^{n-1}\left[\tfrac{-1}{\lambda}\right]^j \boldsymbol{S}_n^j = \sum_{j=0}^{n-1}\tfrac{f^{(j)}(\lambda)}{j!}\boldsymbol{S}_n^j \qquad (\text{ wenn } \lambda \neq 0).\end{aligned}$$

  Hier wurde $(\boldsymbol{I} + \frac{1}{\lambda}\boldsymbol{S}_n)\left(\sum_{j=0}^{n-1}\left[\frac{-1}{\lambda}\right]^j \boldsymbol{S}_n^j\right) = \boldsymbol{I}$ benutzt, was sich wiederum aus $\boldsymbol{S}_n^n = \boldsymbol{O}$ ergibt.

- Ist schließlich $f = p/q$ eine rationale Funktion mit Polynomen p und q, so ist $f(\boldsymbol{A})$ genau dann definiert, wenn kein Eigenwert von $\boldsymbol{A}$ Pol von f ist. Dann gilt

$$f(\boldsymbol{A}) = p(\boldsymbol{A})[q(\boldsymbol{A})]^{-1} = [q(\boldsymbol{A})]^{-1}p(\boldsymbol{A}).$$

Inhalt

Eine alternative Definition von $f(\boldsymbol{A})$ setzt Kenntnisse über polynomiale Interpolationsprozesse voraus, die hier vermittelt werden sollen:

Gegeben seien ein Knotenpolynom

$$w_K(\lambda) = \prod_{\mu=1}^{k} (\lambda - \xi_\mu)^{n_\mu} \in \mathscr{P}_K, \qquad \xi_i \neq \xi_j \text{ für } i \neq j, \quad K = \sum_{\mu=1}^{k} n_\mu, \tag{4.4}$$

sowie K Datenwerte

$$f_{\mu,\nu} \in \mathbb{C}, \qquad \mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1.$$

Wir suchen ein Polynom $p_{K-1} \in \mathscr{P}_{K-1}$ vom Grad $K-1$, welches die K Hermiteschen Interpolationsbedingungen

$$p_{K-1}^{(\nu)}(\xi_\mu) = f_{\mu,\nu}, \qquad \mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1 \tag{4.5}$$

erfüllt.

- In unserem Fall besitzt dieses Hermitesche Interpolationsproblem die Form

$$p_{K-1}^{(\nu)}(\xi_\mu) = f^{(\nu)}(\xi_\mu), \qquad \mu = 1, 2, \ldots, k, \quad \nu = 0, 1, \ldots, n_\mu - 1 \tag{4.6}$$

  mit einer gegebenen Funktion f, die in einer Umgebung der Menge $\{\xi_\mu\}_{\mu=1}^k$ genügend oft differenzierbar ist.
- Wir sagen, dass p_{K-1} die Funktion f an den Nullstellen von w_K im Hermiteschen Sinn interpoliert, wenn (4.6) erfüllt ist.

Satz 4.4 (Eindeutige Lösung der Hermiteschen Interpolationsaufgabe)

Das Hermitesche Interpolationsproblem (4.5) besitzt eine eindeutige Lösung $p_{K-1} \in \mathscr{P}_{K-1}$.

Sind die Knoten in (4.4) alle einfach, so ist

$$Z = \begin{bmatrix} 1 & \xi_1 & \xi_1^2 & \cdots & \xi_1^{K-1} \\ 1 & \xi_2 & \xi_2^2 & \cdots & \xi_2^{K-1} \\ 1 & \xi_3 & \xi_3^2 & \cdots & \xi_3^{K-1} \\ \vdots & & & & \vdots \\ 1 & \xi_K & \xi_K^2 & \cdots & \xi_K^{K-1} \end{bmatrix} \in \mathbb{C}^{K \times K}$$

eine (klassische) Vandermonde-Matrix. Sie ist nach Satz 4.4 invertierbar.

Das Interpolationspolynom p_{K-1} kann explizit angegeben werden:

Satz 4.5 (Explizite Form des Hermiteschen Interpolationspolynoms)

Die Lösung $p_{K-1} \in \mathscr{P}_{K-1}$ des Interpolationsproblems (4.6) kann in der Form

$$p_{K-1}(\lambda) = \sum_{\mu=1}^{k} \left(\sum_{\nu=0}^{n_\mu - 1} \frac{1}{\nu!} \left[\frac{f(\lambda)}{v_\mu(\lambda)} \right]^{(\nu)}_{\lambda=\xi_\mu} (\lambda - \xi_\mu)^\nu \right) v_\mu(\lambda), \qquad v_\mu(\lambda) := \frac{w_K(\lambda)}{(\lambda - \xi_\mu)^{n_\mu}} \in \mathscr{P}_{K-n_\mu} \tag{4.7}$$

geschrieben werden (vgl. (4.4)). Eine alternative Darstellung ist

$$p_{K-1}(\lambda) = \sum_{\mu=1}^{k} \left(\sum_{\nu=0}^{n_\mu - 1} f^{(\nu)}(\xi_\mu) \ell_{\mu,\nu}(\lambda) \right) \tag{4.8}$$

mit

$$\ell_{\mu,\nu}(\lambda) := \frac{1}{\nu!} \left(\sum_{j=0}^{n_\mu - 1 - \nu} \frac{1}{j!} \left[\frac{1}{v_\mu(\lambda)} \right]^{(j)}_{\lambda=\xi_\mu} (\lambda - \xi_\mu)^j \right) (\lambda - \xi_\mu)^\nu v_\mu(\lambda) \in \mathscr{P}_{K-1}. \tag{4.9}$$

Die Polynome $\ell_{\mu,\nu}$ hängen nicht von f ab. Sie sind durch das Knotenpolynom w_K eindeutig bestimmt und charakterisiert durch

$$\ell^{(\tau)}_{\mu,\nu}(\xi_\sigma) = \begin{cases} 1, & \text{falls } \mu = \sigma \text{ und } \nu = \tau, \\ 0, & \text{in allen anderen Fällen.} \end{cases} \tag{4.10}$$

Neben den Darstellungen des Hermiteschen Interpolationspolynoms aus Satz 4.5 ist auch die Newton-Form dieses Polynoms wichtig.

Definition 4.6 (Dividierte Differenzen für Polynome)

Es seien $\{\xi_j\}_{j\in\mathbb{N}} \subset \mathbb{C}$ eine Folge (nicht notwendig verschiedener) Zahlen sowie

$$w_0(\lambda) := 1 \quad \text{und} \quad w_j(\lambda) := \textstyle\prod_{\mu=1}^{j}(\lambda - \xi_\mu) \quad \text{für} \quad j = 1, 2, \ldots .$$

Weil $\{w_j\}_{j\geq 0}$ eine Basis des Polynomraums ist, kann jedes Polynom p eindeutig in der Form

$$p(\lambda) = \alpha_0 w_0(\lambda) + \alpha_1 w_1(\lambda) + \cdots + \alpha_{K-1} w_{K-1}(\lambda) + \cdots \tag{4.11}$$

geschrieben werden. (Natürlich ist $\alpha_j = 0$ für alle $j > \deg p$). Der j-te Koeffizient

$$\alpha_j =: \Delta(\xi_1, \ldots, \xi_{j+1})p = \Delta_{w_{j+1}} p$$

dieser Entwicklung heißt die dividierte Differenz von p bez. $\xi_1, \ldots, \xi_{j+1}$ oder die dividierte Differenz von p bez. der Nullstellen von w_{j+1}.
Die Kurzform dividierte Differenz der Ordnung j ist auch gebräuchlich.

Bemerkung 4.7

(a) Die Bezeichnungsweise $\Delta(\xi_1, \dots, \xi_{j+1})p$ geht auf W. Kahan zurück – ihre Vorteile gegenüber den üblichen Abkürzungen (wie etwa $p[\xi_1, \dots, \xi_{j+1}]$, die zum Beispiel von [Higham, 2008] verwendet wird) werden von [de Boor, 2005] dargelegt. Ein Grund besteht darin, dass $\Delta(\xi_1, \dots, \xi_{j+1})p$ die Rolle von

$$\Delta(\xi_1, \dots, \xi_{j+1}) : \; p \mapsto \Delta(\xi_1, \dots, \xi_{j+1})p$$

als lineares Funktional auf dem Raum aller Polynome betont. (Dass

$$\Delta(\xi_1, \dots, \xi_{j+1})(p+q) = \Delta(\xi_1, \dots, \xi_{j+1})p + \Delta(\xi_1, \dots, \xi_{j+1})q$$

und

$$\Delta(\xi_1, \dots, \xi_{j+1})(\alpha p) = \alpha \Delta(\xi_1, \dots, \xi_{j+1})p$$

gelten, ist leicht einzusehen.)

(b) Die Bezeichnung „dividierte Differenz“ für $\Delta(\xi_1, \dots, \xi_{j+1})p$ wird später, nach Herleitung einiger Eigenschaften der Abbildung Δ, gerechtfertigt.

Zunächst zeigen wir, dass $\Delta(\xi_1, \dots, \xi_{j+1})p$ tatsächlich eine Funktion von nur $\xi_1, \dots, \xi_{j+1}$, also unabhängig von der Wahl der ξ_m für $m > j+1$ ist.
Dazu definieren wir

$$p_j(\lambda) := \sum_{\mu=0}^{j} \Delta(\xi_1, \dots, \xi_{\mu+1})p \, w_\mu(\lambda). \tag{4.12}$$

Das Polynom p_j entsteht, wenn man die (endliche) Reihe (4.11) nach dem j-ten Summanden abbricht und es ist $p_j = p$ für alle $j \geq \deg p$. Da w_{j+1} natürlich w_μ ($\mu \geq j+1$) teilt, gilt

$$\begin{aligned} p(\lambda) &= \sum_{\mu=0}^{j} \Delta(\xi_1, \dots, \xi_{\mu+1})p \, w_\mu(\lambda) + w_{j+1}(\lambda) \sum_{\mu=j+1}^{\infty} \Delta(\xi_1, \dots, \xi_{\mu+1})p \, \tfrac{w_\mu}{w_{j+1}}(\lambda) \\ &= p_j(\lambda) + w_{j+1}(\lambda) r_j(\lambda) \end{aligned} \tag{4.13}$$

mit einem Polynom r_j (wir erinnern noch einmal daran, dass die Summe $\sum_{\mu=j+1}^{\infty}$ endlich ist). Aus dieser offensichtlichen Identität ziehen wir jetzt einige Schlüsse:

(i) Für jedes $j \geq 0$ ist p_j das Hermitesche Interpolationspolynom von p an den Nullstellen von w_{j+1}, also an $\xi_1, \xi_2, \ldots, \xi_{j+1}$.

(ii) Diese Beobachtung hat zwei Konsequenzen: Zum einen hängt $\Delta(\xi_1, \ldots, \xi_{j+1})p$ wie oben behauptet nicht von den restlichen Knoten $\xi_{j+2}, \xi_{j+3}, \ldots$ ab. Zum andern ist $\Delta(\xi_1, \ldots, \xi_{j+1})p$ der führende Koeffizient des Interpolationspolynoms p_j und, weil dieser Koeffizient unabhängig von der Nummerierung der Knoten ist, muss die dividierte Differenz $\Delta(\xi_1, \ldots, \xi_{j+1})p$ eine symmetrische, d.h. permutationsinvariante Funktion ihrer Argumente $\xi_1, \xi_2, \ldots, \xi_{j+1}$ sein.

(iii) Da der führende Koeffizient von p_j stetig von den Knoten $\xi_1, \xi_2, \ldots, \xi_{j+1}$ abhängt (wie etwa die explizite Darstellung des Interpolationspolynoms aus Satz 4.5 zeigt), ist $\Delta(\xi_1, \ldots, \xi_{j+1})p$ eine stetige Funktion von $\xi_1, \xi_2, \ldots, \xi_{j+1}$.

Als Nächstes bestimmen wir die Polynome r_j in (4.13).

- Wegen $r_j(\lambda) = 0$ für $j \geq \deg p$ setzen wir $j < \deg p$ voraus.
- Die Definition (4.12) impliziert

$$p_{j+1}(\lambda) = p_j(\lambda) + \Delta(\xi_1, \dots, \xi_{j+2})p \, w_{j+1}(\lambda)$$

und daher

$$p_{j+1}(\xi_{j+2}) = p_j(\xi_{j+2}) + \Delta(\xi_1, \dots, \xi_{j+2})p \, w_{j+1}(\xi_{j+2}).$$

- Weil $p_{j+1}(\xi_{j+2}) = p(\xi_{j+2})$ gilt, ergibt sich

$$p(\xi_{j+2}) = p_j(\xi_{j+2}) + \Delta(\xi_1, \dots, \xi_{j+2})p \, w_{j+1}(\xi_{j+2}).$$

- Andererseits folgt aus (4.13) noch

$$p(\xi_{j+2}) = p_j(\xi_{j+2}) + w_{j+1}(\xi_{j+2}) \, r_j(\xi_{j+2}),$$

was auf

$$\Delta(\xi_1, \dots, \xi_{j+2})p = r_j(\xi_{j+2}) \quad \text{für jedes } \xi_{j+2} \text{ mit } w_{j+1}(\xi_{j+2}) \neq 0$$

führt.

Betrachtet man jetzt $\Delta(\xi_1, \ldots, \xi_{j+1}, \lambda)p$ als Funktion ihres $(j+2)$-ten Arguments, so gilt

$$r_j(\lambda) = \Delta(\xi_1, \ldots, \xi_{j+1}, \lambda)p$$

für jedes λ, da beide Funktionen stetig sind. Nun schreiben wir (4.13) als

$$\begin{aligned} p(\lambda) &= p_j(\lambda) + w_{j+1}(\lambda)\,\Delta(\xi_1, \ldots, \xi_{j+1}, \lambda)p \quad \text{oder als} \\ \Delta(\xi_1, \ldots, \xi_{j+1}, \lambda)p &= \frac{p(\lambda) - p_j(\lambda)}{w_{j+1}(\lambda)} \in \mathscr{P}_{\deg p - j - 1}. \end{aligned} \tag{4.14}$$

Dann wählen wir $m > j$ und wenden das lineare Funktional $\Delta(\xi_{j+1},\dots,\xi_m)$ auf das Polynom $\Delta(\xi_1,\dots,\xi_j,\lambda)p \in \mathscr{P}_{\deg p - j}$ an, was wegen (4.14) und (4.13) schließlich

$$\begin{aligned}\Delta(\xi_1,\dots,\xi_j,\lambda)p = \tfrac{1}{w_j(\lambda)}\,(p(\lambda) - p_{j-1}(\lambda)) &= \sum_{\mu=j}^{\infty} \Delta(\xi_1,\dots,\xi_{\mu+1})p\;\tfrac{w_\mu(\lambda)}{w_j(\lambda)} \\ &= \Delta(\xi_1,\dots,\xi_{j+1})p + \Delta(\xi_1,\dots,\xi_{j+2})p\;(\lambda - \xi_{j+1}) + \\ &\quad \cdots + \Delta(\xi_1,\dots,\xi_m)p \textstyle\prod_{\mu=j+1}^{m-1}(\lambda - \xi_\mu) + \cdots\end{aligned}$$

liefert.

Damit ist eine der wichtigsten Beziehungen für dividierte Differenzen bewiesen, nämlich

$$\Delta(\xi_{j+1},\dots,\xi_m)\,(\Delta(\xi_1,\dots,\xi_j,\lambda)p) = \Delta(\xi_1,\dots,\xi_m)p. \tag{4.15}$$

Wir fassen zusammen:

Satz 4.8 (Eigenschaften dividierter Differenzen)

Seien komplexe Zahlen $\{\xi_j\}_{j\in\mathbb{N}}$ (nicht notwendig verschieden) und ein Polynom p gegeben. Die dividierte Differenz $\Delta(\xi_1,\ldots,\xi_j)p$ ist der führende Koeffizient des eindeutig bestimmten Polynoms $p_{j-1}\in\mathscr{P}_{j-1}$, welches p an den Nullstellen von $w_j(\lambda)=\prod_{\mu=1}^{j}(\lambda-\xi_\mu)$ im Hermiteschen Sinn interpoliert.
Daher stellt $\Delta(\xi_1,\ldots,\xi_j)p$ eine symmetrische und stetige Funktion iher Argumente $\xi_1,\ldots,\xi_j$ dar.
Ist das j-te Argument variabel, so ist $\Delta(\xi_1,\ldots,\xi_{j-1},\lambda)p$ ein Polynom in λ vom Grad höchstens $\deg p - j$ und es gilt

$$\Delta(\xi_{j+1},\ldots,\xi_m)\,(\Delta(\xi_1,\ldots,\xi_{j-1},\lambda)p)=\Delta(\xi_1,\ldots,\xi_m)p$$

für alle $m>j$. □

Wie wir wissen, gibt es ein eindeutig bestimmtes Polynom $p_{K-1} \in \mathscr{P}_{K-1}$, das eine gegebene Funktion f (im Hermiteschen Sinn) an den Nullstellen von (vgl. (4.4))

$$w_K(\lambda) = \prod_{\mu=1}^{k} (\lambda - \lambda_\mu)^{n_\mu}, \quad \lambda_i \neq \lambda_j \text{ für } i \neq j,$$

interpoliert. Für uns ist w_K gewöhnlich das Minimalpolynom einer Matrix und es ist bequemer, es in der Form

$$w_K(\lambda) = \prod_{\mu=1}^{K} (\lambda - \xi_\mu) \tag{4.16}$$

zu schreiben, wobei ξ_i nicht unbedingt verschieden von ξ_j ist für $i \neq j$. Die Reihenfolge der Knoten ist irrelevant, aber es vereinfacht die Notation, wenn wir gleiche Knoten direkt hintereinander nummerieren, also

$$\{\xi_1, \xi_2, \xi_3, \dots\} = \underbrace{\lambda_1, \lambda_1, \dots, \lambda_1}_{n_1}, \underbrace{\lambda_2, \lambda_2, \dots, \lambda_2}_{n_2}, \dots, \underbrace{\lambda_k, \lambda_k, \dots, \lambda_k}_{n_k}, \dots \tag{4.17}$$

Definition 4.9 (Dividierte Differenzen für allgemeinere Funktionen)

Sei f eine komplexwertige Funktion, die an den Stellen $\xi_1, \xi_2, \dots, \xi_K$ (nicht notwendig verschieden) definiert ist. Erscheint ξ_μ n_μ-mal in $\xi_1, \xi_2, \dots, \xi_K$, fordern wir, dass f an der Stelle ξ_μ $(n_\mu - 1)$-mal differenzierbar ist.
Mit $p_{j-1} \in \mathscr{P}_{j-1}$ bezeichnen wir das eindeutig bestimmte Polynom, das f an den Nullstellen von $w_j(\lambda) = \prod_{\mu=1}^{j}(\lambda - \xi_\mu)$ im Hermiteschen Sinn interpoliert $(j \in 1, 2, \dots, K)$.
Wir definieren

$$\Delta(\xi_1, \dots, \xi_j) f := \Delta(\xi_1, \dots, \xi_j) p_{j-1},$$

d.h. $\Delta(\xi_1, \dots, \xi_j) f$ ist der führende Koeffizient (also der Koeffizient bei λ^{j-1}) von p_{j-1}.

Die Newtonsche Darstellung des Interpolationspolynoms ergibt sich als unmittelbare Folgerung.

Satz 4.10 (Newton-Form des Interpolationspolynoms)

Mit den Bezeichnungen aus Definition 4.9 lässt sich das Interpolationspolynom p_{j-1} für f in der Form

$$p_{j-1}(\lambda) = \sum_{\mu=1}^{j} \Delta(\xi_1, \dots, \xi_\mu) f \; w_{\mu-1}(\lambda)$$

schreiben. Setzt man

$$\Delta(\xi_1, \dots, \xi_\mu, \lambda) f := \frac{f(\lambda) - p_{\mu-1}(\lambda)}{w_\mu(\lambda)},$$

dann gilt für $1 \leq \mu < j \leq K$

$$p_{j-1}(\lambda) = p_{\mu-1}(\lambda) + w_\mu(\lambda) q_{j-\mu-1}(\lambda), \tag{4.18}$$

wobei $q_{j-\mu-1} \in \mathscr{P}_{j-\mu-1}$ die Funktion $\Delta(\xi_1, \dots, \xi_\mu, \lambda) f$ an den Knoten $\xi_{\mu+1}, \xi_{\mu+2}, \dots, \xi_j$ interpoliert. □

- Die Gleichung (4.18) zeigt, wie man das Problem, f an $\xi_1, \dots, \xi_\mu, \dots, \xi_j$ zu interpolieren, in zwei Teilprobleme zerlegen kann:
- Man bestimmt zunächst das Interpolationspolynom für f an den Knoten $\xi_1, \dots, \xi_\mu$ und danach interpoliert man die Funktion $\Delta(\xi_1, \dots, \xi_\mu, \zeta) f$ an den restlichen Knoten $\xi_{\mu+1}, \dots, \xi_j$.

Schließlich soll noch die Bezeichnung 'dividierte Differenz' für Δ gerechtfertigt werden (gleichzeitig ergibt sich ein Algorithmus, mit dem diese Größen berechnet werden können).

Satz 4.11 (Rekursionsformel für dividierte Differenzen)

Angenommen, die Knoten $\xi_1, \dots, \xi_K$ sind so nummeriert, dass gleiche Knoten direkt aufeinander folgen (vgl. (4.17)). Dann gelten

$$\Delta(\xi_\mu) f = f(\xi_\mu) \qquad \text{für } \mu = 1, 2, \dots, K,$$

und für $1 \le \mu < j \le K$

$$\begin{aligned}
&\Delta(\xi_\mu, \xi_{\mu+1}, \dots, \xi_j) f \\
&\qquad = \begin{cases} \frac{1}{\xi_j - \xi_\mu} \left(\Delta(\xi_{\mu+1}, \dots, \xi_j) f - \Delta(\xi_\mu, \dots, \xi_{j-1}) f \right), & \xi_\mu \ne \xi_j, \\ \frac{1}{(j-\mu)!} f^{(j-\mu)}(\xi_\mu), & \xi_\mu = \xi_j. \end{cases}
\end{aligned}$$

Inhalt

- Definition 4.3 zeigt, dass $f(\boldsymbol{A})$ durch die Werte

$$f^{(\nu)}(\lambda_\mu) \qquad \mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1$$

vollständig bestimmt ist. Hier bezeichnet λ_μ eine n_μ-fache Nullstelle des Minimalpolynoms

$$m_{\boldsymbol{A}}(\lambda) = \prod_{\mu=1}^{k} (\lambda - \lambda_\mu)^{n_\mu} \in \mathscr{P}_K, \qquad \lambda_i \neq \lambda_j \text{ für } i \neq j,\ K = \sum_{\mu=1}^{k} n_\mu,$$

von $\boldsymbol{A}$ (vgl. (4.2)).

- Analog zu Korollar 4.2 gilt $f(\boldsymbol{A}) = g(\boldsymbol{A})$ genau dann, wenn

$$f^{(\nu)}(\lambda_\mu) = g^{(\nu)}(\lambda_\mu) \qquad \mu = 1, 2, \dots, k, \quad \nu = 0, 1, \dots, n_\mu - 1$$

erfüllt ist. Damit ist eine sehr nützliche Beschreibung von $f(\boldsymbol{A})$ gefunden, die man auch als Definition von $f(\boldsymbol{A})$ verwenden könnte.

Satz 4.12 (Darstellung von $f(\boldsymbol{A})$ mittels Interpolationspolynom)

Seien $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ und f eine komplexwertige Funktion, für die $f(\boldsymbol{A})$ definiert ist. Mit $p_{K-1} \in \mathscr{P}_{K-1}$ bezeichnen wir das eindeutig bestimmte Polynom vom Grad höchstens $K-1$, welches f an den Nullstellen des Minimalpolynoms $m_{\boldsymbol{A}}$ (4.2) von $\boldsymbol{A}$ (im Hermiteschen Sinn) interpoliert. Dann gilt

$$f(\boldsymbol{A}) = p_{K-1}(\boldsymbol{A}).$$

- Die Aussage von Satz 4.12 kann verallgemeinert werden (siehe Übungsaufgabe).
- Auf jeden Fall ist $f(\boldsymbol{A})$ ein Polynom in $\boldsymbol{A}$, wobei das Polynom allerdings von $\boldsymbol{A}$ abhängt.
- Viele Eigenschaften von Matrixfunktionen folgen unmittelbar aus dieser Beobachtung.

Satz 4.13 (Eigenschaften von Matrixfunktionen)

Es seien $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ mit dem Minimalpolynom $m_{\boldsymbol{A}}$ sowie f und g komplexwertige Funktionen, für die $f(\boldsymbol{A})$ bzw. $g(\boldsymbol{A})$ definiert sind.

(1) Ist $\boldsymbol{B} \in \mathbb{C}^{n\times n}$ ähnlich zu $\boldsymbol{A}$, d.h. gibt es eine invertierbare Matrix $\boldsymbol{T} \in \mathbb{C}^{n\times n}$ mit $\boldsymbol{B} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}$, dann ist $f(\boldsymbol{B})$ definiert und ähnlich zu $f(\boldsymbol{A})$. Es gilt $f(\boldsymbol{B}) = \boldsymbol{T} f(\boldsymbol{A}) \boldsymbol{T}^{-1}$.

(2) Ist $\boldsymbol{A} = \operatorname{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_s)$ blockdiagonal (mit quadratischen Diagonalblöcken $\boldsymbol{A}_j$), dann ist $f(\boldsymbol{A}_j)$ für jedes $j = 1, 2, \dots, s$ definiert und $f(\boldsymbol{A}) = \operatorname{diag}(f(\boldsymbol{A}_1), f(\boldsymbol{A}_2), \dots, f(\boldsymbol{A}_s))$ ist ebenfalls blockdiagonal.

(3) Die Matrizen $f(\boldsymbol{A})$ und $g(\boldsymbol{A})$ kommutieren, d.h. $f(\boldsymbol{A})g(\boldsymbol{A}) = g(\boldsymbol{A})f(\boldsymbol{A})$.

(4) Für die Funktion $h(\lambda) = f(\lambda) + g(\lambda)$ ist $h(\boldsymbol{A})$ definiert und es gilt $h(\boldsymbol{A}) = f(\boldsymbol{A}) + g(\boldsymbol{A})$.

(5) Für die Funktion $h(\lambda) = f(\lambda)g(\lambda)$ ist $h(\boldsymbol{A})$ definiert und es gilt $h(\boldsymbol{A}) = f(\boldsymbol{A})g(\boldsymbol{A}) = g(\boldsymbol{A})f(\boldsymbol{A})$.

Satz 4.13 (Fortsetzung)

(6) Aus $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ ($\lambda \in \mathbb{C}$, $\boldsymbol{v} \in \mathbb{C}^n \setminus \{\boldsymbol{0}\}$) folgt $f(\boldsymbol{A})\boldsymbol{v} = f(\lambda)\boldsymbol{v}$.
Ist $c_{\boldsymbol{A}}(\lambda) = \prod_{\mu=1}^k (\lambda - \lambda_\mu)^{m_\mu}$ ($\sum_{j=1}^k m_\mu = n$) das charakteristische Polynom von $\boldsymbol{A}$, so ist $c_{f(\boldsymbol{A})}(\lambda) = \prod_{\mu=1}^k (\lambda - f(\lambda_\mu))^{m_\mu}$.
Das Minimalpolynom $m_{f(\boldsymbol{A})}$ von $f(\boldsymbol{A})$ teilt $\prod_{\mu=1}^k (\lambda - f(\lambda_\mu))^{n_\mu}$.
Insbesondere ist $\deg m_{f(\boldsymbol{A})} \leq \deg m_{\boldsymbol{A}}$.

(7) Ist $\boldsymbol{A}$ is diagonalisierbar, so ist auch $f(\boldsymbol{A})$ diagonalisierbar.
Ist $\boldsymbol{A}$ normal, so ist auch $f(\boldsymbol{A})$ normal.

(8) Es gilt die Buchheimsche Formel [Buchheim, 1886]

$$f(\boldsymbol{A}) = \sum_{\mu=1}^{k} \left(\sum_{\nu=0}^{n_\mu - 1} \frac{1}{\nu!} \left[\frac{f(\lambda)}{v_\mu(\lambda)} \right]^{(\nu)}_{\lambda=\lambda_\mu} (\boldsymbol{A} - \lambda_\mu \boldsymbol{I})^\nu \right) v_\mu(\boldsymbol{A})$$

mit $v_\mu(\lambda) = m_{\boldsymbol{A}}(\lambda)/(\lambda - \lambda_\mu)^{n_\mu}$ (vgl. (4.7)).

Satz 4.13 (Fortsetzung)

(9) Eine andere Darstellung von $f(\boldsymbol{A})$ ist

$$f(\boldsymbol{A}) = \sum_{\mu=1}^{k} \left(\sum_{\nu=0}^{n_\mu - 1} f^{(\nu)}(\lambda_\mu) \boldsymbol{Z}_{\mu,\nu} \right),$$

wobei die Matrizen $\boldsymbol{Z}_{\mu,\nu} := \ell_{\mu,\nu}(\boldsymbol{A})$ (vgl. (4.8)) die Komponenten von $\boldsymbol{A}$ genannt werden.

Die Komponenten von $\boldsymbol{A}$ haben interessante Eigenschaften, die wir an dieser Stelle aber nicht formal herleiten werden. Wir betrachten stattdessen ein Beispiel.

Die Matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 20 & -7 & -8 & 9 & -6 \\ -3 & 1 & 2 & -2 & 1 \\ 3 & -1 & -1 & 2 & -1 \\ -14 & 6 & 6 & -6 & 5 \end{bmatrix} = \boldsymbol{T} \left[\begin{array}{ccc|c|c} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right] \boldsymbol{T}^{-1}$$

mit

$$\boldsymbol{T} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 2 \end{bmatrix} \quad \text{und} \quad \boldsymbol{T}^{-1} = \begin{bmatrix} 3 & -1 & -2 & 1 & -1 \\ 3 & -1 & -1 & 2 & -1 \\ 3 & -1 & -1 & 1 & -1 \\ -7 & 3 & 3 & -3 & 2 \\ -2 & 1 & 1 & -1 & 1 \end{bmatrix}.$$

besitzt die Eigenwerte $\lambda_1 = 1$, $\lambda_2 = -1$ und das Minimalpolynom

$$m_{\boldsymbol{A}}(\lambda) = (\lambda - 1)^3(\lambda + 1), \qquad \text{d.h. } n_1 = 3, n_2 = 1.$$

Die Polynome $\ell_{\mu,\nu}$ sind

$$
\begin{aligned}
\ell_{1,0}(\lambda) &= \left[\tfrac{1}{2} - \tfrac{1}{4}(\lambda-1) + \tfrac{1}{8}(\lambda-1)^2\right](\lambda+1) &&= \tfrac{1}{8}\left(\lambda^3 - 3\lambda^2 + 3\lambda + 7\right),\\
\ell_{1,1}(\lambda) &= \left[\tfrac{1}{2} - \tfrac{1}{4}(\lambda-1)\right](\lambda-1)(\lambda+1) &&= \tfrac{1}{4}\left(-\lambda^3 + 3\lambda^2 + \lambda - 3\right),\\
\ell_{1,2}(\lambda) &= \tfrac{1}{2}\left[\tfrac{1}{2}\right](\lambda-1)^2(\lambda+1) &&= \tfrac{1}{4}\left(\lambda^3 - \lambda^2 - \lambda + 1\right),\\
\ell_{2,0}(\lambda) &= \left[-\tfrac{1}{8}\right](\lambda-1)^3 &&= \tfrac{1}{8}\left(-\lambda^3 + 3\lambda^2 - 3\lambda + 1\right).
\end{aligned}
$$

Damit können wir die Komponenten von $\boldsymbol{A}$ berechnen, entweder über

$$\boldsymbol{Z}_{\mu,\nu} = \ell_{\mu,\nu}(\boldsymbol{A}) \quad \text{oder über} \quad \boldsymbol{Z}_{\mu,\nu} = \boldsymbol{T}\ell_{\mu,\nu}(\boldsymbol{J})\,\boldsymbol{T}^{-1}:$$

$$
\boldsymbol{Z}_{1,0} = \boldsymbol{T}\left[\begin{array}{ccc|c|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \end{array}\right]\boldsymbol{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 7 & -2 & -3 & 3 & -2 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -7 & 3 & 3 & -3 & 3 \end{bmatrix},
$$

$$
\boldsymbol{Z}_{1,1} = \boldsymbol{T}\left[\begin{array}{ccc|c|c} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right]\boldsymbol{T}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 6 & -2 & -2 & 3 & -2 \\ -3 & 1 & 1 & -2 & 1 \\ 3 & -1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

$$
\boldsymbol{Z}_{1,2} = \tfrac{1}{2}\,\boldsymbol{T}\left[\begin{array}{ccc|c|c} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right]\boldsymbol{T}^{-1} = \tfrac{1}{2}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 3 & -1 & -1 & 1 & -1 \\ -3 & 1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},
$$

$$
\boldsymbol{Z}_{2,0} = \boldsymbol{T}\left[\begin{array}{ccc|c|c} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right]\boldsymbol{T}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -7 & 3 & 3 & -3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & -3 & -3 & 3 & -2 \end{bmatrix}.
$$

Wir sehen, dass die Identitäten

$$\begin{aligned} \boldsymbol{Z}_{1,0} + \boldsymbol{Z}_{2,0} &= \boldsymbol{I}, \\ \boldsymbol{Z}_{1,0} + \boldsymbol{Z}_{1,1} - \boldsymbol{Z}_{2,0} &= \boldsymbol{A} \\ \text{und} \quad \boldsymbol{Z}_{1,0} + 2\boldsymbol{Z}_{1,1} + 2\boldsymbol{Z}_{1,2} + \boldsymbol{Z}_{2,0} &= \boldsymbol{A}^2 \end{aligned}$$

gelten. Diese Gleichungen sind alle Spezialfälle von Satz 4.13 (9), also von

$$f(\boldsymbol{A}) = f(1)\,\boldsymbol{Z}_{1,0} + f'(1)\,\boldsymbol{Z}_{1,1} + f''(1)\,\boldsymbol{Z}_{1,2} + f(-1)\,\boldsymbol{Z}_{2,0}.$$

Diese Beziehung gilt für alle Funktionen, die in $\{\pm 1\}$ definiert und an der Stelle $\lambda = 1$ zweimal differenzierbar sind (wähle $f(\lambda) = 1$, $f(\lambda) = \lambda$ bzw. $f(\lambda) = \lambda^2$).

Die Multipikationstafel der Komponenten von $\boldsymbol{A}$ ist gegeben durch

	$\boldsymbol{Z}_{1,0}$	$\boldsymbol{Z}_{1,1}$	$\boldsymbol{Z}_{1,2}$	$\boldsymbol{Z}_{2,0}$
$\boldsymbol{Z}_{1,0}$	$\boldsymbol{Z}_{1,0}$	$\boldsymbol{Z}_{1,1}$	$\boldsymbol{Z}_{1,2}$	$\boldsymbol{O}$
$\boldsymbol{Z}_{1,1}$	$\boldsymbol{Z}_{1,1}$	$2\boldsymbol{Z}_{1,2}$	$\boldsymbol{O}$	$\boldsymbol{O}$
$\boldsymbol{Z}_{1,2}$	$\boldsymbol{Z}_{1,2}$	$\boldsymbol{O}$	$\boldsymbol{O}$	$\boldsymbol{O}$
$\boldsymbol{Z}_{2,0}$	$\boldsymbol{O}$	$\boldsymbol{O}$	$\boldsymbol{O}$	$\boldsymbol{Z}_{2,0}$

Insbesondere sind $\boldsymbol{Z}_{1,0}$ und $\boldsymbol{Z}_{2,0}$ idempotent, d.h. $\boldsymbol{Z}_{\mu,0}^2 = \boldsymbol{Z}_{\mu,0}$ für $\mu = 1, 2$, während $\boldsymbol{Z}_{1,1}$ und $\boldsymbol{Z}_{1,2}$ nilpotent sind ($\boldsymbol{Z}_{1,1}^3 = 2\boldsymbol{Z}_{1,2}\boldsymbol{Z}_{1,1} = \boldsymbol{O}$, $\boldsymbol{Z}_{1,2}^2 = \boldsymbol{O}$). Außerdem gilt

$$\boldsymbol{A}\boldsymbol{Z}_{1,0} = (\boldsymbol{Z}_{1,0} + \boldsymbol{Z}_{1,1} - \boldsymbol{Z}_{2,0})\boldsymbol{Z}_{1,0} = \boldsymbol{Z}_{1,0} + \boldsymbol{Z}_{1,1}, \quad \text{d.h.} \quad (\boldsymbol{A} - \boldsymbol{I})\boldsymbol{Z}_{1,0} = \boldsymbol{Z}_{1,1}$$

und $\boldsymbol{A}^2\boldsymbol{Z}_{1,0} = (\boldsymbol{Z}_{1,0} + 2\boldsymbol{Z}_{1,1} + 2\boldsymbol{Z}_{1,2} + \boldsymbol{Z}_{2,0})\boldsymbol{Z}_{1,0} = \boldsymbol{Z}_{1,0} + 2\boldsymbol{Z}_{1,1} + 2\boldsymbol{Z}_{1,2}$, d.h.

$$(\boldsymbol{A} - \boldsymbol{I})^2\boldsymbol{Z}_{1,0} = (\boldsymbol{A}^2 - 2\boldsymbol{A} + \boldsymbol{I})\boldsymbol{Z}_{1,0} = 2\boldsymbol{Z}_{1,2}.$$

Satz 4.14 (Eigenschaften von Komponenten)

Gegeben sei $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ mit Minimalpolynom $m_{\boldsymbol{A}}(\lambda) = \prod_{\mu=1}^{k}(\lambda - \lambda_\mu)^{n_\mu}$ (vgl. (4.2)). Für die Komponenten $\boldsymbol{Z}_{\mu,\nu} = \ell_{\mu,\nu}(\boldsymbol{A})$ von $\boldsymbol{A}$ ($\mu = 1, 2, \ldots, k$, $\nu = 0, 1, \ldots, n_\mu - 1$) gelten folgende Aussagen:

(1) Zwei beliebige Komponenten von $\boldsymbol{A}$ kommutieren.
Jede Komponente von $\boldsymbol{A}$ kommutiert mit $\boldsymbol{A}$.

(2) Die $K = \sum_{\mu=1}^{k} n_\mu$ Matrizen $\boldsymbol{Z}_{\mu,\nu}$ sind linear unabhänging. Insbesondere ist keine gleich der Nullmatrix.

(3) $\boldsymbol{Z}_{\mu,\nu}\boldsymbol{Z}_{\sigma,\tau} = \boldsymbol{O}$ für $\mu \neq \sigma$ und

$$\boldsymbol{Z}_{\mu,\nu}\boldsymbol{Z}_{\mu,\tau} = \begin{cases} \binom{\nu+\tau}{\nu}\boldsymbol{Z}_{\mu,\nu+\tau} = \binom{\nu+\tau}{\tau}\boldsymbol{Z}_{\mu,\nu+\tau}, & \text{für } \nu+\tau = 0, 1, \ldots, n_\mu - 1, \\ \boldsymbol{O}, & \text{für } \nu+\tau = n_\mu, n_\mu + 1, \ldots \end{cases}$$

Insbesondere sind die Matrizen $\boldsymbol{Z}_{\mu,0}$ idempotent ($\boldsymbol{Z}_{\mu,0}^2 = \boldsymbol{Z}_{\mu,0}$), während $\boldsymbol{Z}_{\mu,\nu}$ nilpotent vom Index $n_j - \nu$ ist ($\boldsymbol{Z}_{\mu,\nu}^{n_j-\nu} = \boldsymbol{O} \neq \boldsymbol{Z}_{\mu,\nu}^{n_j-\nu-1}$) für $\nu = 1, 2, \ldots, n_j - 1$.

Satz 4.14 (Fortsetzung)

(4) Für $m = 0, 1, \dots$ ist

$$\boldsymbol{A}^m = \sum_{\mu=1}^{k} \left(\sum_{\nu=0}^{\min\{m, n_\mu - 1\}} \nu! \binom{m}{\nu} \lambda_\mu^{m-\nu} \boldsymbol{Z}_{\mu,\nu} \right).$$

Wir betonen zwei Spezialfälle dieser Formel: Für $m = 0$ ist $\boldsymbol{I} = \sum_{\mu=1}^{k} \boldsymbol{Z}_{\mu,0}$. Ist $\boldsymbol{A}$ diagonalisierbar (d.h. $n_\mu = 1$ für alle μ), so gilt $\boldsymbol{A} = \sum_{\mu=1}^{k} \lambda_\mu \boldsymbol{Z}_{\mu,0}$.

(5) Für $\mu = 1, 2, \dots, k$ ist

$$(\boldsymbol{A} - \lambda_\mu \boldsymbol{I})^\nu \boldsymbol{Z}_{\mu,0} = \nu! \boldsymbol{Z}_{\mu,\nu}, \quad \nu = 0, 1, \dots, n_\mu - 1.$$

Inhalt

- Grundlage dieses Abschnitts ist die Cauchysche Integralformel

$$f(\lambda) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{\zeta - \lambda} \, \mathrm{d}\zeta, \qquad \lambda \in \operatorname{int}\Gamma. \tag{4.19}$$

  Hier bezeichnet Γ eine einfach geschlossene, positiv orientierte Jordan-Kurve; die Funktion f sei analytisch in $\operatorname{int}\Gamma$, dem Innengebiet von Γ, und stetig in $\Gamma \cup \operatorname{int}\Gamma$.

- Die Formel (4.19) (und jede der folgenden Aussagen) gelten auch dann, wenn Γ Vereinigung endlich vieler (einfach geschlossener, positiv orientierter) Jordan-Kurven Γ_1, Γ_2, ..., Γ_k ist, sodass Γ_j im Außengebiet von allenΓ_ℓ, $\ell \neq j$, liegt $(j = 1, 2, \dots, k)$.

  In diesem Fall muss $f_j := f\big|_{\Gamma_j \cup \operatorname{int}\Gamma_j}$ stetig und analytisch in $\operatorname{int}\Gamma_j$ sein.

  Beachtenswert ist, dass die Funktionen f_j nicht notwendigerweise analytische Fortsetzungen voneinander sein müssen.

Der Residuensatz ist eine Verallgemeinerung von (4.19):

Ist f bis auf endlich viele Singularitäten $\lambda_1, \lambda_2, \dots, \lambda_m$ analytisch im Innengebiet Γ und stetig in $\Gamma \cup (\operatorname{int}\Gamma \setminus \{\lambda_1, \lambda_2, \dots, \lambda_m\})$, dann gilt

$$\frac{1}{2\pi i}\int_\Gamma f(\zeta)\,\mathrm{d}\zeta = \sum_{j=1}^m \operatorname{Res}_{\lambda_j} f, \tag{4.20}$$

wobei $\operatorname{Res}_{\lambda_j} f$ das Residuum von f an der Stelle λ_j bezeichnet. Einen wichtigen Spezialfall von (4.20) stellt

$$\frac{f^{(j)}(\lambda_0)}{j!} = \frac{1}{2\pi i}\int_\Gamma \frac{f(\zeta)}{(\zeta-\lambda_0)^{j+1}}\,\mathrm{d}\zeta \qquad (j = 0, 1, 2, \dots) \tag{4.21}$$

dar, wobei f analytisch in $\operatorname{int}\Gamma$, stetig in $\Gamma \cup \operatorname{int}\Gamma$ sowie $\lambda_0 \in \operatorname{int}\Gamma$ ist.

Wir wollen ein Analogon von (4.19) für Matrixfunktionen nachweisen, genauer

$$f(\boldsymbol{A}) = \frac{1}{2\pi i} \int_\Gamma f(\zeta)(\zeta \boldsymbol{I} - \boldsymbol{A})^{-1} \mathrm{d}\zeta, \tag{4.22}$$

wenn alle Eigenwerte von $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ im Innern von Γ liegen.

- Das Integral in (4.22) ist komponentenweise zu verstehen, d.h. für eine $(n \times n)$-Matrix $\boldsymbol{F}(\zeta) = [f_{i,j}(\zeta)]$ definieren wir

$$\int_\Gamma \boldsymbol{F}(\zeta)\,\mathrm{d}\zeta := \left[\int_\Gamma f_{i,j}(\zeta)\,\mathrm{d}\zeta\right]_{1 \le i,j \le n} \in \mathbb{C}^{n \times n}$$

  (vorausgesetzt, die n^2 Integrale existieren).
- Unmittelbare Folgerungen sind etwa

$$\int_\Gamma (\boldsymbol{T}\boldsymbol{F}(\zeta)\boldsymbol{S})\,\mathrm{d}\zeta = \boldsymbol{T}\left(\int_\Gamma \boldsymbol{F}(\zeta)\,\mathrm{d}\zeta\right)\boldsymbol{S} \text{ und}$$

$$\int_\Gamma \operatorname{diag}(\boldsymbol{F}_1(\zeta), \ldots, \boldsymbol{F}_s(\zeta))\,\mathrm{d}\zeta = \operatorname{diag}\left(\int_\Gamma \boldsymbol{F}_1(\zeta)\,\mathrm{d}\zeta, \ldots, \int_\Gamma \boldsymbol{F}_s(\zeta)\,\mathrm{d}\zeta\right).$$

Wir betrachten die Matrix $(\zeta \boldsymbol{I} - \boldsymbol{J}(\lambda))^{-1}$ mit dem Jordan-Block $\boldsymbol{J}(\lambda) \in \mathbb{C}^{n\times n}$ (vgl. (4.1)) und $\zeta \neq \lambda$.

Die Identität

$$(\zeta \boldsymbol{I} - \boldsymbol{J}(\lambda))^{-1} = \begin{bmatrix} \frac{1}{\zeta-\lambda} & \frac{1}{(\zeta-\lambda)^2} & \cdots & \frac{1}{(\zeta-\lambda)^{n-1}} & \frac{1}{(\zeta-\lambda)^{n}} \\ & \frac{1}{\zeta-\lambda} & \cdots & \frac{1}{(\zeta-\lambda)^{n-2}} & \frac{1}{(\zeta-\lambda)^{n-1}} \\ & & \ddots & \vdots & \vdots \\ & & & \frac{1}{\zeta-\lambda} & \frac{1}{(\zeta-\lambda)^2} \\ & & & & \frac{1}{\zeta-\lambda} \end{bmatrix} \tag{4.23}$$

ist leicht nachzuweisen.

Zusammen mit (4.21) und (4.3) liefert sie

$$\frac{1}{2\pi i}\int_\Gamma f(\zeta)\left(\zeta \boldsymbol{I} - \boldsymbol{J}(\lambda)\right)^{-1}\,\mathrm{d}\zeta = \begin{bmatrix} \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & \frac{f(\lambda)}{0!} & \cdots & \frac{f^{(n-3)}(\lambda)}{(n-3)!} & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \vdots & \vdots \\ & & & \frac{f(\lambda)}{0!} & \frac{f'(\lambda)}{1!} \\ & & & & \frac{f(\lambda)}{0!} \end{bmatrix}$$
$$= f\left(\boldsymbol{J}(\lambda)\right),$$

wenn Γ eine Jordan-Kurve ist, die λ in ihrem Innengebiet enthält, und wenn f analytisch in $\operatorname{int}\Gamma$ sowie stetig in $\Gamma \cup \operatorname{int}\Gamma$ ist.

Für $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ beliebig mit Jordanscher Normalform $\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_1, \dots, \boldsymbol{J}_s) = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ folgt

$$
\begin{aligned}
f(\boldsymbol{A}) &= \boldsymbol{T} \operatorname{diag}\left(f(\boldsymbol{J}_1), \dots, f(\boldsymbol{J}_s)\right) \boldsymbol{T}^{-1} \\
&= \frac{1}{2\pi i} \boldsymbol{T} \operatorname{diag}\left(\int_\Gamma f(\zeta)(\zeta \boldsymbol{I} - \boldsymbol{J}_1)^{-1} \,\mathrm{d}\zeta, \dots, \int_\Gamma f(\zeta)(\zeta \boldsymbol{I} - \boldsymbol{J}_s)^{-1} \,\mathrm{d}\zeta\right) \boldsymbol{T}^{-1} \\
&= \frac{1}{2\pi i} \int_\Gamma f(\zeta)\, \boldsymbol{T} \operatorname{diag}\left((\zeta I - \boldsymbol{J}_1)^{-1}, \dots, (\zeta \boldsymbol{I} - \boldsymbol{J}_s)^{-1}\right) \boldsymbol{T}^{-1} \,\mathrm{d}\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma f(\zeta)\, \boldsymbol{T} \left[\operatorname{diag}\left(\zeta \boldsymbol{I} - \boldsymbol{J}_1, \dots, \zeta \boldsymbol{I} - \boldsymbol{J}_s\right)\right]^{-1} \boldsymbol{T}^{-1} \,\mathrm{d}\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma f(\zeta) \left[\boldsymbol{T} \operatorname{diag}\left(\zeta \boldsymbol{I} - \boldsymbol{J}_1, \dots, \zeta \boldsymbol{I} - \boldsymbol{J}_s\right) \boldsymbol{T}^{-1}\right]^{-1} \mathrm{d}\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma f(\zeta) \left[\zeta \boldsymbol{I} - \boldsymbol{T} \operatorname{diag}\left(\boldsymbol{J}_1, \dots, \boldsymbol{J}_s\right) \boldsymbol{T}^{-1}\right]^{-1} \mathrm{d}\zeta \\
&= \frac{1}{2\pi i} \int_\Gamma f(\zeta)\, (\zeta \boldsymbol{I} - \boldsymbol{A})^{-1} \,\mathrm{d}\zeta,
\end{aligned}
$$

wenn Γ ein Jordan-Kurve ist, deren Innengebiet alle Eigenwerte von $\boldsymbol{A}$ enthält.

Damit ist (4.22) bewiesen.

Satz 4.15 (Darstellung von $f(A)$ als Resolventenintegral)

Es seien $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, Γ eine Jordan-Kurve sowie f eine Funktion, die analytisch im Innengebiet $\operatorname{int}\Gamma$ von Γ und stetig in $\Gamma \cup \operatorname{int}\Gamma$ ist. Sind alle Eigenwerte von $\boldsymbol{A}$ in $\operatorname{int}\Gamma$ enthalten, so ist $f(\boldsymbol{A})$ definiert und es gilt

$$f(\boldsymbol{A}) = \frac{1}{2\pi i}\int_\Gamma f(\zeta)(\zeta\boldsymbol{I} - \boldsymbol{A})^{-1}\mathrm{d}\zeta.$$

- Wie in den vorigen Abschnitten bezeichnen wir mit $p_{K-1} \in \mathscr{P}_{K-1}$ das eindeutig bestimmte Polynom vom Grad (höchstens) $K-1$, welches eine gegebene Funktion f an den Nullstellen ξ_j von w_K interpoliert (siehe (4.5), (4.6)).
- Ist f analytisch in einer Umgebung der Knoten $\xi_1, \xi_2, \ldots, \xi_k$, dann besitzt auch p_{K-1} eine Darstellung als Cauchy-Integral, die wir als nächstes herleiten:

Satz 4.16 (Integraldarstellung des Interpolationspolynoms)

Seien Γ eine positiv orientierte Jordan-Kurve (oder eine endliche Vereinigung von Jordan-Kurven, wie sie oben beschrieben wurde) und f eine in $\operatorname{int}\Gamma$ analytische Funktion, die stetig in $\Gamma \cup \operatorname{int}\Gamma$ ist. Sind die Nullstellen des Knotenpolynoms w_K alle in $\operatorname{int}\Gamma$ enthalten, dann kann man das zugehörige Interpolationspolynom p_{K-1} schreiben in der Form

$$p_{K-1}(\lambda) = \frac{1}{2\pi i} \int_\Gamma \frac{w_K(\zeta) - w_K(\lambda)}{\zeta - \lambda} \, \frac{f(\zeta)}{w_K(\zeta)} \, \mathrm{d}\zeta \qquad (\lambda \in \operatorname{int}\Gamma).$$

Der Interpolationsfehler gegeben ist durch

$$f(\lambda) - p_{K-1}(\lambda) = \frac{1}{2\pi i} \int_\Gamma \frac{w_K(\lambda)}{w_K(\zeta)} \, \frac{f(\zeta)}{\zeta - \lambda} \, \mathrm{d}\zeta \qquad (\lambda \in \operatorname{int}\Gamma).$$

Integraldarstellung von dividierten Differenzen analytischer Funktionen:
Aus Satz 4.10 und dem Beweis von Satz 4.16 kann man folgende Identität ableiten:

$$\begin{aligned}\Delta(\xi_1, \dots, \xi_\mu, \lambda) f &= \frac{f(\lambda) - p_{k-1}(\lambda)}{w_K(\lambda)} \\ &= \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{w_K(\zeta)(\zeta - \lambda)} \, \mathrm{d}\zeta \qquad (\lambda \in \operatorname{int} \Gamma).\end{aligned}$$

Inhalt

Für Matrizen $\boldsymbol{A}$ mit spezieller Struktur ist $f(\boldsymbol{A})$ manchmal (mehr oder weniger) explizit bekannt, dies gilt etwa für Diagonalmatrizen oder Jordan-Blöcke.

Hier ist eine Verallgemeinerung (cf. [Opitz, 1964]): wir betrachten zunächst spezielle obere Bidiagonalmatrizen der Form

$$\boldsymbol{A} = \begin{bmatrix} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & 1 \\ & & & & \lambda_n \end{bmatrix} \in \mathbb{C}^{n\times n}. \tag{4.24}$$

Satz 4.17 (Funktionen bidiagonaler Matrizen)

Ist $\boldsymbol{A}$ eine obere Bidiagonalmatrix der Form (4.24) (und ist $f(\boldsymbol{A})$ definiert), so gilt

$$f(\boldsymbol{A}) = \begin{bmatrix} f(\lambda_1) & \Delta(\lambda_1,\lambda_2)f & \cdots & \Delta(\lambda_1,\dots,\lambda_{n-1})f & \Delta(\lambda_1,\dots,\lambda_n)f \\ & f(\lambda_2) & \cdots & \Delta(\lambda_2,\dots,\lambda_{n-1})f & \Delta(\lambda_2,\dots,\lambda_n)f \\ & & \ddots & \vdots & \vdots \\ & & & f(\lambda_{n-1}) & \Delta(\lambda_{n-1},\lambda_n)f \\ & & & & f(\lambda_n) \end{bmatrix}.$$

Beweis siehe [Opitz, 1964].

Der Fall allgemeiner Bidiagonalmatrizen

$$\boldsymbol{B} = \begin{bmatrix} \lambda_1 & \beta_2 & & & \\ & \lambda_2 & \beta_3 & & \\ & & \ddots & \ddots & \\ & & & \lambda_{n-1} & \beta_n \\ & & & & \lambda_n \end{bmatrix} \in \mathbb{C}^{n\times n}$$

lässt sich einfach auf Satz 4.17 zurückführen.

- Wir können annehmen, dass die Eintrage in der oberen Nebendiagonalen alle von Null verschieden sind, denn andernfalls ist $\boldsymbol{B}$ blockdiagonal und $f(\boldsymbol{B})$ kann mithilfe von Satz 4.13 (ii) ausgewertet werden.
- Die Diagonalmatrix $\boldsymbol{D} := \operatorname{diag}\left(1, \beta_2, \prod_{j=2}^{3} \beta_j, \dots, \prod_{j=2}^{n} \beta_j\right)$ ist dann invertierbar.
- Sie transformiert $\boldsymbol{B}$ nach $\boldsymbol{A}$ (aus Satz 4.17): $\boldsymbol{D}\boldsymbol{B}\boldsymbol{D}^{-1} = \boldsymbol{A}$. Folglich ist

$$f(\boldsymbol{B}) = \boldsymbol{D}^{-1} f(\boldsymbol{A}) \boldsymbol{D}.$$