Plan

- Wir werden uns weiter mit Jordanschen Normalform beschäftigen: wir besprechen noch Anwendungen und auch wie man die Ergebnisse auf reelle Matrizen übertragen kann (weil nicht alle Polynome mit reellen Koeffizienten in Linearfaktoren zerfallen).
- Allerdigs mussen wir als Vorbereitung Ihnen den aus Ana I bekannten Begriff "Ableitung" "algebraisch" definieren.

Exkurs: Ableitung von Polynomen über $\mathbb{K}$

- Aus Ana I oder sogar aus der Schule wissen wir, was die Ableitung einer reell- oder komplexwertigen Funktion f ist.
- Wir wissen auch, dass
  - ein $\mathbb{R}$- oder $\mathbb{C}$-Polynom im Wesentlichen dasselbe wie eine Funktion ist (Ein Polynom $\sum_i a_i x^i$ definiert immer eine Funktion $\tilde{f}(z) = \sum a_i z^i$. Wir haben letztes Mal gesehen, dass es für einige $\mathbb{K}$ mehrere Polynome gibt, die einer Funktion entsprechen.)
  - und dass die Ableitung eines solchen Polynoms auch ein Polynom ist.
- Was tun wir aber wenn unseres Polynom ein $\mathbb{K}$-Polynom ist für einen Körper $\mathbb{K} \notin \{\mathbb{R}, \mathbb{C}\}$?

Def. Sei $P = \sum_{i \geq 0} a_i x^i \in \mathbb{K}[x]$. Wir definieren P' durch $P' = \sum_{i \geq 1} i a_i x^{i-1} = \sum_{i \geq 0} (i+1) a_{i+1} x^i$ (wobei x^0 wie immer gleich 1 gesetzt wird).

Bsp. Sei $\mathbb{K} = \mathbb{R}$ und $P = x^3 + 2x + 19$. Dann ist $P' = 3x^2 + 2$.

Was ist übrigens i? In $\mathbb{R}$ bzw. $\mathbb{C}$ es ist klar, was i ist. (Beachte: die Bezeichnung sind suboptimal gewähl: i in diesem Kontext ist nicht $\sqrt{-1}$, sondern die ganze Zahl i). In einem beliebigen Körper $\mathbb{K}$ definieren wir $i := \underbrace{1 + ... + 1}_{i \text{ Stück}}$. Z.B. in $\mathbb{Z}_3$ ist $5 = 1 + 1 + 1 + 1 + 1 = 2$. Eigentlich ist eine formal bessere Formel in der Definition des Polynoms $P' = \sum_{i \geq 0} \left(\sum_{\ell=1}^{i+1} a_{i+1} x^i \right)$.

Bemerkung. Für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ fällt die Definition oben mit der Ableitung im Sinne Analysis I bzw. der Schule überein.

Rechnen wir zusammen in $\mathbb{Z}_2$

Leiten Sie bitte $f = x^3 + x^2 - x + 2$ ab.

Antwort. $f = (1+1+1)x^2 + (1+1)x - 1 \cdot 1 = x^2 + 1.$

Jetzt nehmen wir an, dass $\mathbb{K} = \mathbb{Z}_3$ ist. Leiten Sie das gleiche Polynom $f = x^3 + x^2 - x + 2$ ab.

Antwort. $-x - 1$ (oder $2x + 2$)

Ableitung ist eine lineare Abbildung.

$$\boxed{P' = \sum_{i>1} i a_i x^{i-1} = \sum_{i>0} (i+1) a_{i+1} x^i}$$

Wir wissen, dass die Polynome $\mathbb{K}$ einen Vektorraum (bzgl. $+$ und Multiplizieren mit Konstanten) bilden.

Lemma 5. Die Abbildung von $\mathbb{K}[x]$ nach $\mathbb{K}[x]$, die einem Polynom f dessen Ableitung zuordnet, ist linear.

Beweis ist eine einfache Übung, bitte selber machen.

Leibniz-Regel für die Ableitung

Lemma 6. Für zwei beliebige Polynome $P, Q \in \mathbb{K}$ gilt:
$(PQ)' = P'Q + PQ'$.

Beweisidee. Im Prinzip kann man die Definition der Ableitung

$$\boxed{P' = \sum_{i>0}(i+1)a_{i+1}x^i}$$

und des Produkts von Polynomen benutzen, um direkt die Koeffizienten der Polynome $(PQ)'$, $P'Q + PQ'$ auszurechnen und zu beweisen, dass sie gleich sind.

Das ist möglich, aber technisch nicht trivial – Sie können das zu Hause ausprobieren.
In unserem Beweis benutzen wir zweimal Linearität der Ableitung (Lemma 5) und reduzieren damit den Beweis zur Aussage, dass für die *Monome*, d.h. für die Polynome der Form $P = x^k$ und $Q = x^m$, das Lemma gilt. Für die Monomen ist das Lemma rechnerisch einfach verifizierbar.

Beweis ist auch als Demonstration einer Stärke der Linearen Algebra gedacht: Wie wir es öfter gemacht haben, reduziert man den Beweis zu überprüfen der Aussage für Basis-Vektoren.

Lemma 6. Für zwei beliebige Polynome $P, Q \in \mathbb{K}$ gilt:
$(PQ)' = P'Q + PQ'$.

Beweis. Wir betrachten zwei Abbildungen $A : \mathbb{K}[x] \to \mathbb{K}[x]$, $B : \mathbb{K}[x] \to \mathbb{K}[x]$ definiert durch

$$A(P) = (PQ)', \quad B(P) = P'Q + PQ'.$$

Ich betone: Q ist ein festes Polynom.

Unseres Ziel ist zu zeigen, dass diese zwei Abbildungen zusammenfallen: d.h., $\forall P \in \mathbb{K}[x]$ gilt $A(P) = B(P)$. Wenn wir das gezeigt haben, haben wir auch das Lemma 6 bewiesen, weil das Polynom Q eigentlich auch beliebig ist.

Die Abbildungen A und B sind linear: z.B. wenn wir $A(P_1 + P_2)$ nehmen, bekommen wir

$$A(P_1+P_2) = ((P_1+P_2)Q)' \overset{\text{Rechenregeln}}{=} (P_1Q+P_2Q)' \overset{\text{Weil Ableitung linear ist}}{=} (P_1Q)'+(P_2Q)' = A(P_1)+A(P_2).$$

(Beweis der Linearitätseigenschaften für $A(\lambda \cdot P)$, $B(P_1 + P_2)$, $B(\lambda \cdot P)$ sind analog.)

Da die Abbildungen A, B linear sind, genügt es, wenn wir die Gleichung $A = B$ nur für Basiselemente, also für die Monome $P = x^k$ beweisen. Jetzt beweisen wir also dass für jedes Monom x^k gilt: $A(x^k) = B(x^k)$. Zu zeigen ist, dass (für jedes Q) $(x^kQ)' = (x^k)'Q + x^kQ'$.

Jetzt beweisen wir also dass für jedes Monom x^k gilt: $A(x^k) = B(x^k)$. Zu zeigen ist dass (für jedes Q) $(x^k Q)' = (x^k)'Q + x^k Q'$.

Dazu wiederholen wir die Argumentation: Wir betrachten die Abbildungen $\tilde{A} : \mathbb{K}[x] \to \mathbb{K}[x]$ und $\tilde{B} : \mathbb{K}[x] \to \mathbb{K}[x]$, definiert durch

$$\tilde{A}(Q) = (x^k Q)', \quad \tilde{B}(Q) = (x^k)'Q + x^k Q'.$$

Die Abbildungen $\tilde{A}$, $\tilde{B}$ sind linear. Deswegen genügt es, wenn wir die Gleichung $\tilde{A} = \tilde{B}$ nur für Basiselemente, also für die Monome $Q = x^m$ beweisen.

Also: Um das Lemma 6 zu beweisen, reicht es, wenn wir es nur für Monome $P = x^k$ und $Q = x^m$ beweisen.

Beweis für $P = x^k$ und $Q = x^m$: Für $k > 0, m > 0$ gilt:
$(x^k x^m)' = (x^{k+m})' = (k+m)x^{k+m-1}$ ist gleich
$(x^k)'x^m + x^k(x^m)' = kx^{k-1}x^m + mx^k x^{m-1} = (k+m)x^{k+m-1}$ wie wir es wollen.

Für $k = 0$ (Der Fall $m = 0$ ist analog.) gilt: $(x^0 x^m)' = (x^m)'$ und
$(x^0)'x^m + x^0(x^m)' = (x^m)'$. $\square$

Wiederholung

Satz 8 (Jordansche Normalform) Sei $\mathbb{K}$ ein Körper, sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum und sei $\phi : V \to V$ ein Endomorphismus, dessen Minimalpolynom in Linearfaktoren zerfällt. Dann existiert eine Basis von V, so dass die Matrix von ϕ eine Jordan-Matrix ist. Diese Jordansche Normalform ist bis auf Reihenfolge der Jordanblöcke eindeutig.

Wiederholung: *Jordan-Matrizen* sind die Matrizen der Form

$$\begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$$

, wobei J_λ^k die folgende $k \times k$-Matrix ist

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

Wiederholung

- **Beobachtung 1** Sei $P \in \mathbb{K}[x]$ ein Polynom, $A \in Mat(n,n,\mathbb{K})$, $B \in GL(n,\mathbb{K})$. Dann gilt: $P(B^{-1}AB) = B^{-1}P(A)B$.
- **Beobachtung 2** Für eine Block-diagonale Matrix

$$M = \begin{pmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_m} \end{pmatrix}$$

, wobei A_i eine $k_i \times k_i$-Matrix ist, ist

$$P(M) = \begin{pmatrix} \boxed{P(A_1)} & & \\ & \ddots & \\ & & \boxed{P(A_m)} \end{pmatrix}.$$

Beobachtung 3 ist Hausaufgabe

- **Beobachtung 3** Sei $J = J_\lambda^k$ ein $k \times k$ Jordan Block über $\mathbb{K}$ mit Eigenwert λ. Dann gilt:
$$P(J) = \begin{pmatrix} P(\lambda) & \frac{1}{1!}P'(\lambda) & \frac{1}{2!}P^{(2)}(\lambda) & \cdots & \frac{1}{n!}P^{(n)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \frac{1}{2!}P^{(2)}(\lambda) \\ & & & P(\lambda) & \frac{1}{1!}P'(\lambda) \\ & & & & P(\lambda) \end{pmatrix},$$
wobei $P^{(j)}$ die $j-$te Ableitung des Polynoms P ist.
(Z.B. für das Polynom $P = z^3$ ist $P' = 3z^2$, $P^{(2)} = 6z$, $P^{(3)} = 6$, $P^{(4)} = P^{(5)} = ... = 0$.
Allgemein: Für ein Polynom $P = \sum_{k=0}^m a_k x^k \in \mathbb{C}[x]$ ist
$P' = \sum_{k=1}^m k \cdot a_k x^{k-1} \in \mathbb{K}[x]$ $(*)$.
- Das wird Ihre Hausaufgabe.Ein Hinweis: da die Ableitung eine lineare Abbildung von $\mathbb{K}[x]$ auf $\mathbb{K}[x]$ ist , genügt es, die Formel für Monome (d.h. für die Polynome der Form x^k) zu prüfen, und das müssen Sie tun (und soll nicht besonders schwierig sein).

Die Beobachtungen 2,3 geben uns eine Formel für die Polynomen der Matrix in Jordannormaldorm: Ich mache es wegen Platzmangel nur für zwei Blöcke; Sie werden die Antwort sofort für den Fall von k Blöcke verallgemeinern können. Wir betrachten die Matrix $A = \begin{pmatrix} J_\lambda^k & \\ 0 & J_\mu^m \end{pmatrix}$ und ein Polynom P. Dann gilt:

$$P(A) = \begin{pmatrix} P(J_\lambda^k) & \\ 0 & P(J_\mu^m) \end{pmatrix} =$$

$$\begin{pmatrix} \boxed{\begin{matrix} P(\lambda) & \frac{1}{1!}P'(\lambda) & \frac{1}{2!}P^{(2)}(\lambda) & \cdots & \frac{1}{n!}P^{(n)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \frac{1}{2!}P^{(2)}(\lambda) \\ & & & P(\lambda) & \frac{1}{1!}P'(\lambda) \\ & & & & P(\lambda) \end{matrix}} & \\ & \boxed{\begin{matrix} P(\mu) & \frac{1}{1!}P'(\mu) & \frac{1}{2!}P^{(2)}(\mu) & \cdots & \frac{1}{n!}P^{(n)}(\mu) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \frac{1}{2!}P^{(2)}(\mu) \\ & & & P(\mu) & \frac{1}{1!}P'(\mu) \\ & & & & P(\mu) \end{matrix}} \end{pmatrix}$$

Die Beobachtung 1 gibt uns dann die Methode, das Polynom von beliebigen Matrix $A \in Mat(\mathbb{C}, n \times n)$ auszurechnen.

Analytische Funktionen von Matrizen

Sei $a_0, a_1, ..., a_k, ... \in \mathbb{C}$ eine Folge, $f(z) = a_0 + a_1 z + a_2 z^2 + ... + a_k z^k + ... = \sum_{k=0}^{\infty} a_k z^k$ die entsprechende analytische Funktion und A eine Matrix. Wir betrachten die Folge von Matrizen(partielle Summen): $a_0 \cdot Id$, $a_0 \cdot Id + a_1 A$, ..., $a_0 \cdot Id + a_1 \cdot A + a_2 \cdot A^2 + ... + a_m A^m = \sum_{k=0}^{m} a_k \cdot A^k := P_m(A)$,...
Wir sagen, dass die Folge konvergiert, falls sie komponentenweise konvergiert, d.h., für jede $i, j \in n$ konvergieren die $(i,j)-$ Einträge von Matrizen $a_0 \cdot Id$, $a_0 \cdot Id + a_1 A$, ..., $P_m(A)$.

Def. Falls die Folge konvergiert, setzen wir

$$f(A) = \lim_{m \to \infty} \sum_{k=0}^{m} a_k \cdot A^k.$$

(Falls für irgendwelche i, j die Folge divergiert, ist $f(A)$ nicht definiert.)
Bemerkung: Für Polynomen $\sum_{k=0}^{m} a_k \cdot A^k$ gilt:
$\sum_{k=0}^{m} a_k \cdot (B^{-1}AB)^k = B^{-1} \left(\sum_{k=0}^{m} a_k \cdot A^k \right) B$. Wir werden bald sehen: $f(B^{-1}AB) = B^{-1} f(A) B$.
Dies bedeutet, dass die Definition nicht von der Wahl der Basis abhängt (wir also in Wirklichkeit analytische Funktionen von Endomorphismen definiert haben).

Natürliche Fragen

Unten ist $f = \sum_{k=0}^{\infty} a_k z^k$ stets eine analytische Funktion und A eine $n \times n$-Matrix.

1. Für welche A existiert $f(A)$? (d.h. für welche A konvergiert die Reihe $a_0 \cdot Id$, $a_0 \cdot Id + a_1 A$, ..., komponentenweise)
2. Wie kann man $f(A)$ ausrechnen?
   Bemerkung. Es ist fast hoffnungslos, bereits für 3×3-Matrizen einfache analytische Funktionen von Matrizen, z.B. e^A, nach Definition auszurechnen.
3. Wozu sollte man analytische Funktionen von Matrizen ausrechnen?
   Bemerkung. In dem Kurs gew. Differentialgleichungen werden analytische Funktionen von linearen Abbildungen auch behandelt(sogar für Banach-Räume), teilweise parallel zu linearen Algebra. Das bedeutet, dass sie für Differentialgleichungenund deswegen auch alle Naturwissenschaften wichtig sind.

Satz 9 Angenommen, alle Eigenwerte von $A \in Mat(n, n, \mathbb{C})$ liegen im Konvergenzkreis von f. Dann gilt: $f(A)$ ist definiert.

Satz 9 beantwortet die erste Frage: wann $f(A)$ definiert ist.

Satz 10 (Spektralsatz) Seien $\lambda_1, ..., \lambda_m$ verschiedene Eigenwerte von $A \in Mat(n, n, \mathbb{C})$, $k_1, ..., k_m$ ihre algebraische Vielfachheiten $(k_1 + ... + k_m = n)$. Angenommen, der Konvergenzradius von f ist größer als $max|\lambda_i|$.
Dann gilt: Es gibt n Matrizen
$Z_{\lambda_1,1}, ..., Z_{\lambda_1,k_1}, Z_{\lambda_2,1}, ..., Z_{\lambda_2,k_2}, ..., Z_{\lambda_m,1}, ..., Z_{\lambda_m,k_m} \in Mat(n, n, \mathbb{C})$, die nur von A abhängen, so dass für jede analytische Funktion f gilt: ist $f(A)$ definiert, so ist
$$f(A) = \sum_{i=1}^{m} \left(f(\lambda_i) Z_{\lambda_i,1} + ... + f^{(k_i-1)}(\lambda_i) Z_{\lambda_i,k_i} \right) \qquad (*)$$
In Worten. Für jede Matrix A gibt es n Matrizen $Z_{\lambda_i,j}$, sodass jede analytische Funktion $f(A)$ mit genügend großem Konvergenzradius die Linearkombination $(*)$ ist.

Bemerkung. Die Matrizen $Z_{\lambda_i,j}$ hängen nicht von f ab, nur von A. Satz 10 beantwortet die zweite Frage und gibt uns eine effektive Methode, die Matrix $f(A)$ auszurechnen.

Beweis der Sätze 9/10.

Beweisstrategie: Wir zeigen zuerst, dass es genügt, die Sätze nur für Jordan-Blöcke zu beweisen und dann die Formeln für die Blöcke zu prüfen.

Wir benutzen die folgende Aussage aus Ana I:
Konvergieren die Folgen $\alpha_0, ..., \alpha_k, ...$; $\beta_0, ..., \beta_k, ...$ gegen α bzw. β, so konwergiert die Folge $C_1 \cdot \alpha_i + C_2 \cdot \beta_i$ gegen $C_1 \cdot \alpha + C_2 \cdot \beta$.

Ist $f(B^{-1}AB)$ definiert, dann ist $f(A)$ auch definiert. In der Tat, für jedes Polynom $P_k = a_0 + ... + a_k z^k$ ist $P_k(A) = BP_k(B^{-1}AB)B^{-1}$. Also sind die Einträge von $P_k(A)$ lineare Ausdrücke in den Einträgen von $P_k(B^{-1}AB)$ mit konstanten Koeffizienten, die von B kommen. Falls alle Einträge von $P_k(B^{-1}AB)$ für $k \to \infty$ konvergieren, konvergieren deswegen auch die Einträge von $P_k(A) = BP_k(B^{-1}AB)B^{-1}$, und zwar gegen $Bf(B^{-1}AB)B^{-1}$. Deswegen genügt es, die Sätze 9, 10 für Jordan-Matrizen zu beweisen.

Da für eine Block-diagonale Matrix $M = \begin{pmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_m} \end{pmatrix}$, wobei A_i eine $k_i \times k_i$-Matrix ist, $P_n(M) = \begin{pmatrix} \boxed{P_n(A_1)} & & \\ & \ddots & \\ & & \boxed{P_n(A_m)} \end{pmatrix}$ gilt, genügt es, die Sätze 9, 10 für die einzelnen Jordan-Blöcke zu beweisen.

Beweis für Jordan-Blöcke: Ausrechnen

Wie wir es oben erklärt haben, ist

$$P_n(J_\lambda^k) = \begin{pmatrix} P_n(\lambda) & \frac{1}{1!}P_n'(\lambda) & \frac{1}{2!}P_n^{(2)}(\lambda) & \cdots & \frac{1}{k!}P_n^{(k)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \vdots \\ & & \ddots & & \frac{1}{2!}P_n^{(2)}(\lambda) \\ & & & P_n(\lambda) & \frac{1}{1!}P_n'(\lambda) \\ & & & & P_n(\lambda) \end{pmatrix}$$

Falls λ im Konvergenzkreis von f liegt, konvergiert $P_n(\lambda)$ gegen $f(\lambda)$. Folglich konvergieren die Diagonalelemente gegen $f(\lambda)$. Falls λ im Konvergenzkreis von f liegt, konvergiert $P_n'(\lambda)$ gegen $f'(\lambda)$, folglich konvergieren die $(i, i+1)$ Einträge gegen $\frac{1}{1!}f'(\lambda)$, u.s.w. Satz 9 ist bewiesen.

Wir sehen, dass $f(J)$ eine Linearkombination der Matrizen

$$D_i := \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \\ & \ddots & & \ddots & \vdots \\ & & 0 & & 1 \\ & & & \ddots & \vdots \\ & & & & 0 \end{pmatrix}$$ (die nicht von f abhängen)mit den Koeffizienten $\frac{1}{i!}f^{(i)}(\lambda)$ ist. Satz 10 ist bewiesen.

Folgerung (Rechnenregeln für analytische Funktionen von Matrizen) Es gilt:

1. $Af(A) = f(A)A$ (falls $f(A)$ definiert ist).
2. Sei $h = f + g$. Dann gilt $h(A) = f(A) + g(A)$ (falls $h(A), f(A)$ definiert sind.)
3. Sei $h = fg$. Dann gilt $h(A) = f(A)g(A)$ (falls definiert)
4. Sei $h(z) = f(g(z))$. Dann gilt $h(A) = f(g(A))$ (falls definiert)
5. Sei $f(\lambda) \neq 0$ für jeden Eigenwert λ von A, $h = 1/f$. Dann $h(A) = (f(A))^{-1}$.

Beweis. Benutzen Sie Satz 10 oder die Definition. Z.B. in (1) sei $g(z) = zf(z) = f(z)z$. Dann ist $g(A) = Af(A) = f(A)A$.

Anwendung: die Exponentialabbildung von Matrizen

Die gewöhnliche Exponentialfunktion e^x in der Analysis ist durch jede der beiden folgenden Eigenschaften charakterisiert:
1. Als Lösung der Differentialgleichung $y' = y$ mit der Anfangsbedingung $y(0) = 1$,
2. Als (überall konvergente) Potenzreihe $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.
Wir betrachten jetzt e^A, wobei A eine Matrix ist, und zeigen, dass Sie die beiden Eigenschaften hat.
Nach Def. von $f(A)$ ist $e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$.

e^J, wobei $J = J_\lambda^k$ ein Jordan-Block ist

Im Beweis der Sätze 9, 10 haben wir gezeigt, dass

$$f(J) = \begin{pmatrix} f(\lambda) & \frac{1}{1!}f'(\lambda) & \frac{1}{2!}f^{(2)}(\lambda) & \cdots & \frac{1}{n!}f^{(n)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \vdots \\ & & \ddots & & \frac{1}{2!}f^{(2)}(\lambda) \\ & & & f(\lambda) & \frac{1}{1!}f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}$$

Da für $f(z) = e^z$ auch $f^{(j)}(z) = e^z$ gilt, ist $f^{(j)}(\lambda) = e^\lambda$und deswegen

$$e^J = e^\lambda \begin{pmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{n!} \\ & \ddots & \ddots \ \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 1 & \frac{1}{1!} \\ & & & & 1 \end{pmatrix}$$

Für eine Jordan-Matrix $M = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_m} \end{pmatrix}$, wobei $J_i := J_{\lambda_i}^{k_i}$ ein $k_i \times k_i$-Jordanblock ist, ist $e^M = \begin{pmatrix} \boxed{e^{J_1}} & & \\ & \ddots & \\ & & \boxed{e^{J_m}} \end{pmatrix}$.

Insbesondere gilt: $det(e^M) = e^{trace(M)}$ (beide Seiten sind gleich $e^{\sum_i k_i \lambda_i}$).

Satz 11 Die Exponentialabbildung hat die folgenden Eigenschaften:
(Unten sind A, B stets $n \times n$-Matrizen über $\mathbb{C}$.)
(1) e^A konvergiert für jede Matrix A.
(2) Die Zuordnung $A \mapsto e^A$ ist eine C^∞ – Abbildung (sogar reell analytisch) (als Abbildung von $\underbrace{\mathbb{C}^{n^2}}_{=Mat(n,n,\mathbb{C})}$ auf sich selbst.)
(3) Es gilt $e^{A+B} = e^A \cdot e^B$, falls $AB = BA$ (dies ist das Exponentialgesetz)
(4) e^A ist stets invertierbar, und es gilt $(e^A)^{-1} = e^{-A}$ sowie $det(e^A) = e^{trace(A)}$
(5) Die Abbildung $t \mapsto e^{tA}$ ist bzgl. t differenzierbar mit $e^{0 \cdot A} = Id$ und $\frac{d}{dt} e^{t \cdot A} = Ae^{tA}$.
Beweis (1): Da der Konvergenzradius der e^z-Funktion unendlich ist, ist e^A für alle $A \in Mat(n, n, \mathbb{C})$ nach Satz 9 definiert.
Beweis (2): Weil die Einträge von e^A konvergente Reihen von Einträgen von A sind.

Beweis (3): Die Gleichung $AB = BA$ impliziert die binomische Formel $(A+B)^2 = A^2 + AB + BA + B^2 = A^2 + 2AB + B^2$, allgemein $(A+B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$, wobei $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
Es folgt
$e^{A+B} = \sum_{n=0}^{\infty} \frac{1}{n!}(A+B)^n$
$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}$
$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} A^k B^{n-k}$
$= \sum_{n=0}^{\infty} \sum_{k+l=n} \frac{1}{k!l!} A^k B^l = \sum_{k,l=0}^{\infty} \frac{1}{k!l!} A^k B^l \overset{(*)}{=}$
$\left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k\right) \left(\sum_{l=0}^{\infty} \frac{1}{l!} B^l\right) = e^A e^B$.
Bemerkung. In $(*)$ habe ich benutzt, dass unsere Folgen absolut konvergent sind; deswegen es ist egal in welcher Reinordnung wir sie aufsummieren.

Beweis (4):

Setze $A := -B$. Offensichtlich, $AB = BA (= -B^2)$. Dann folgt nach Teil (3): $Id = e^{\mathbf{0}} = e^{A-A} = e^A \cdot e^{-A}$, also $e^{-A} = \left(e^A\right)^{-1}$.
Die Gleichung $det(e^A) = e^{trace(A)}$ bekommt man sofort, wenn A eine Jordan-Matrix ist, wie wir vor dem Satz ausgerechnet haben. In diesem Fall stehen auf der Diagonale von e^A die Werte e^{λ_i}, wobei λ_i die Eigenwerte sind. $e^{trace(A)}$ ist ebenfalls Produkt von Exponentialfunktionen von Eigenwerten (mit Vielfachheiten). Da *det* und *trace* nicht von Wahl der Basis abhängen, gilt die Gleichung $det(e^A) = e^{trace(A)}$ für alle A.
Beweis (5): Hier rechnet man mit gliedweiser Differentiation folgendes aus:

$$\frac{d}{dt}(e^{t\cdot A}) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \cdot A^k = \sum_{k=0}^{\infty} \frac{d}{dt} \frac{1}{k!} t^k \cdot A^k$$
$$= \sum_{k=1}^{\infty} \frac{d}{dt} \frac{1}{k!} \cdot k \cdot t^{k-1} \cdot A^k$$
$$\overset{k-1=m}{=} A \sum_{m=0}^{\infty} \frac{d}{dt} \frac{1}{m!} \cdot t^m \cdot A^m.$$

Anwendung: Lösen von linearen Differentialgleichungssystemen mit konstanten Koeffizienten

Für eine gegebene $n \times n$-Matrix A wird ein Vektor $u = \begin{pmatrix} u_1(x) \\ \vdots \\ u_n(x) \end{pmatrix}$ gesucht,

dessen Einträge u_i Funktionen von x sind, s.d.

$(*) \quad u'(x) = Au$

d.h., $\begin{pmatrix} u_1'(x) \\ \vdots \\ u_n'(x) \end{pmatrix} = A \begin{pmatrix} u_1(x) \\ \vdots \\ u_n(x) \end{pmatrix}$

Bsp. Für $n = 1$ ist $(*)$ die bekannte Gleichung $u'(x) = au(x)$, und deren Lösung ist $u(x) = C \cdot e^{ax}$ wobei C eine Konstante ist.

$(*)$ ist eine äußerst wichtige Gleichung in Physik und kommt selbstverständlich im Kurs über gewöhnliche Differentialgleichungen vor.

Nach Satz 11 löst

$$u(x) := e^{Ax} C \qquad (**)$$

diese Gleichung , wobei C ein (konstanter) Vektor ist. Nach dem Satz von Picard-Lindelöf (kommt noch im Kurs gewöhnliche Differentialgleichungen) haben alle Lösungen von $(*)$ die Form $(**)$.

Algorithmus zur Berechnung der Jordan-Normalform für eine Matrix A: erste Schritte

Wir haben gesehen, wie effektiv die Jordan-Normalform ist. Wie man die Jordan-Normalform bestimmen kann, haben wir ebenfalls besprochen. Wir werden es wiederholen. Wir besprechen jetzt, wie man eine Basis findet, sodass die darstellende Matrix die Jordan-Normalform hat.

Die gleiche Frage in der algebraischen Sprache: Gegeben sei eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$. Gesucht ist eine Matrix B, sodass

$$B^{-1}AB$$

die Jordan-Normalform hat.

Algorithmus zur Berechnung der Jordan-Normalform für eine Matrix A: erste Schritte

1. Eigenwerte berechnen
2. Algebraische und geometrische Vielfachheiten berechnen
3. Verallgemeinerte Eigenräume bestimmen
4. Jordan-Normalform bestimmen
5. Jordan-Ketten aufbauen
6. Matrix B konstruieren, sodass $B^{-1}AB$ die Jordan-Normalform hat.

1. Eigenwerte berechnen

Berechne das charakteristische Polynom:

$$\chi_A(t) = \det(A - t\mathrm{Id})$$

und finde die Eigenwerte $\lambda_1, \ldots, \lambda_k$ als Nullstellen dieses Polynoms.

Wir werden den Algorithmus an zwei Beispielen ausführen: eine Matrix steht bereits in JNF, die zweite nicht:

$$A_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -2 & 15 & -1 \\ -4 & 13 & 0 \\ -3 & 7 & 4 \end{pmatrix}$$

Wir berechnen:

$$\chi_{A_1}(t) = \det(A_1 - t\mathrm{Id}) = (2-t)^3 \Rightarrow \text{Eigenwert } \lambda = 2$$

(algebraische Vielfachheit 3)

$$\chi_{A_2}(t) = \det(A_2 - t\mathrm{Id}) = (5-t)^3$$

Der einzige Eigenwert ist $\lambda = 5$, mit algebraischer Vielfachheit 3.

2. Algebraische und geometrische Vielfachheiten

Für jeden Eigenwert λ finde:

- **Algebraische Vielfachheit** $alg_A(\lambda)$: Vielfachheit von λ als Nullstelle von $\chi_A(t)$, d.h. die Anzahl der Faktoren $(\lambda - t)$ in $\chi_A(t)$.
- **Geometrische Vielfachheit** $geo_A(\lambda)$: Dimension des Eigenraums:
$$U_\lambda = \dim(\mathrm{Kern}(A - \lambda I)).$$

Für unsere Beispielmatrizen A_1 und A_2 haben wir bereits die algebraische Vielfachheit berechnet: $alg_{A_1}(2) = alg_{A_2}(5) = 3$. Wir berechnen jetzt geo_{A_1} und geo_{A_2}:

Bsp.

$$A_1 - 2\mathrm{Id} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \mathrm{Kern}(A_1 - 2\mathrm{Id}) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Geometrische Vielfachheit: 1.

Bsp.

$$\mathrm{Kern}(A_2 - 5\mathrm{Id}) = \mathrm{Kern}\begin{pmatrix} -7 & 15 & -1 \\ -4 & 8 & 0 \\ -3 & 7 & -1 \end{pmatrix}$$

Der Rang ist 2, daher ist der Kern 1-dimensional, also $geo_{A_2}(5) = 1$.

3. Verallgemeinerte Eigenräume bestimmen

Berechne für jedes λ den verallgemeinerten Eigenraum:

$$U_\lambda = \mathrm{Kern}(A - \lambda\,\mathrm{Id})^{\mathrm{alg}_\lambda(A)}$$

("Berechne" bedeutet: Finde eine Basis von U_λ.)

Da wir in beiden obigen Beispielen nur einen Eigenwert haben, ist der verallgemeinerte Eigenraum U_λ gleich dem gesamten Vektorraum $\mathbb{C}^3$. Daher entfällt dieser Schritt.

5. Jordan-Normalform bestimmen

Wie im Beweis von Satz 7 ersichtlich ist, können wir die Jordan-Form bestimmen, ohne zuvor die Transformationsmatrix B zu berechnen. Wir benötigen lediglich die Ränge von $(A - \lambda_i \operatorname{Id})^k$ für alle Eigenwerte λ_i und für alle $k \leq \operatorname{alg}_{\lambda_i}(A)$.

Beispiel:

$$A_1 - 2\operatorname{Id} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow (A_1 - 2\operatorname{Id})^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (A_1 - 2\operatorname{Id})^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Wir haben: $\lambda = 2$, $\operatorname{Rang}(A - \lambda \operatorname{Id}) = 2$, $\operatorname{Rang}((A - \lambda \operatorname{Id})^2) = 1$, $\operatorname{Rang}((A - \lambda \operatorname{Id})^3) = 0$.

Beispiel:

$$A_2 - 5\,\mathrm{Id} = \begin{pmatrix} -7 & 15 & -1 \\ -4 & 8 & 0 \\ -3 & 7 & -1 \end{pmatrix}, \quad (A_2 - 5\,\mathrm{Id})^2 = \begin{pmatrix} -8 & 8 & 8 \\ -4 & 4 & 4 \\ -4 & 4 & 4 \end{pmatrix}, \quad (A_2 - 5\,\mathrm{Id})^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Es gilt: $\lambda = 5$, $\mathrm{Rang}(A - \lambda\,\mathrm{Id}) = 2$, $\mathrm{Rang}((A - \lambda\,\mathrm{Id})^2) = 1$, $\mathrm{Rang}((A - \lambda\,\mathrm{Id})^3) = 0$.
Jordan-Normalform:

$$J = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

5. Jordan-Ketten aufbauen

Def. Jordan-Kette (zu Eigenwert λ) ist das Tupel $((A-\lambda\mathrm{Id})^{k-1}v, (A-\lambda\mathrm{Id})^{k-2}v, \ldots, v)$ sodass $(A-\lambda\mathrm{Id})^{k}v = \vec{0}$, aber $(A-\lambda\mathrm{Id})^{k-1}v \neq \vec{0}$. Die Vektoren in einer Jordan-Kette sind automatisch linear unabhängig. Wenn ein Basis-Tupel aus nacheinander-stehenden Jordan-Ketten besteht, steht die darstellende Matrix in Jordan-Normalform.

Wir nehmen ein Eigenwert λ und bauen nacheinander die Jordan-Ketten auf. Sei k_1 die Dimensional des größten Jordan-Block mit EW λ, und sei m_1 die Anzahl von solchen Blöcken. Dann nehmen wir m_1 linear unabhängigen Lösungen Vektoren, welche nicht in $\mathrm{Kern}(A-\lambda\mathrm{Id})^{k_1-1}$ liegen.

Für jede von diesen m_1 linear unabhängigen Vektoren $v_{k_1,1}, ..., v_{k_1,m_1}$ bauen wir die Jordan Kette wie folgt aus:

$$((A-\lambda\mathrm{Id})^{k_1-1}v_{k_1,i}(A-\lambda\mathrm{Id})^{k_1-2}v_{k_1,i}, \ldots, v_{k_1,i},), \quad i = 1, \ldots, m_1.$$

Wir nehmen die Vektoren aus Ketten als ersten $k_1 \times m_1$ Basisvektoren, und bezeichnen

$$\begin{aligned} U_{k_1} \quad = \quad & \mathrm{span}(\{(A-\lambda \mathrm{Id})^{k_1-1} v_{k_1,1}, (A-\lambda \mathrm{Id})^{k_1-2} v_{k_1,1}, \ldots, v_{k_1,1}, \\ & (A-\lambda \mathrm{Id})^{k_1-1} v_{k_1,2}, (A-\lambda \mathrm{Id})^{k_1-2} v_{k_1,2}, \ldots, v_{k_1,2}, \ldots, \\ & (A-\lambda \mathrm{Id})^{k_1-1} v_{k_1,m_1} v_{k_1,i}, (A-\lambda \mathrm{Id})^{k_1-2} v_{k_1,m_1}, \ldots, v_{k_1,m_1}\}) \end{aligned} \quad .$$

Danach nehmen wir die zweit-größte Dimension k_2 von Jordan-Blöcke, sei m_2 die Anzahl von solchen λ-Blöcke. Dann nehmen wir m_2 linear unabhängigen Vektoren, welche die Bedingungen $(A-\lambda \mathrm{Id})^{k_2} x = \vec{0}$, $(A-\lambda \mathrm{Id})^{k_2-1} x \neq \vec{0}$, $x \notin U_{k_1}$ erfüllen. und konstruieren die entsprechenden Ketten, u.s.w.

Jordan-Kette für Matrix A_1

Wir haben einen eindimensionalen Eigenraum $\operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$. Also gibt es nur eine Jordan-Kette. Wir erhalten sie, indem wir ein v finden, sodass $(A_1 - 2\mathrm{Id})^2 v \neq \vec{0}$.

Wenn wir $v = \begin{pmatrix}0\\0\\1\end{pmatrix}$ wählen, ergibt sich die Kette:

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Selbstverständlich könnten wir auch ein anderes v wählen, etwa $v_1 = \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}$. Diese Wahl ändert die Basis, in der der Endomorphismus in Jordan-Normalform dargestellt wird, und somit auch die Matrix B; sie ändert jedoch nicht die Jordan-Normalform selbst.

Jordan-Kette für Matrix A_2

Die Bedingung $(A_2 - 5\mathrm{Id})^2 v \neq \vec{0}$ ist für fast alle Vektoren erfüllt.

Wenn wir etwa mit dem Vektor $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ beginnen, ergibt sich die

Kette:

$$\left(\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

6. Die Matrix B konstruieren

Die Matrix B ist die Transformationsmatrix von der Standardbasis zu der Basis, die aus den oben konstruierten Vektoren besteht. Die Spalten von B sind also genau diese Vektoren.
Die Vektoren sind nach Eigenwerten geordnet: zuerst kommen $\mathrm{alg}_{\lambda_1}(A)$ Vektoren aus dem verallgemeinerten Eigenraum zu λ_1, in Jordan-Ketten organisiert, danach die zu λ_2, usw.

Die Matrix B für A_1

Für die Matrix

$$A_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

haben wir nur einen Eigenwert. In dem zugehörigen Eigenraum haben wir eine Jordan-Kette der Länge 3 gefunden:

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$$

Für diese Basis ist $B = \mathrm{Id}$, und somit gilt:

$$B^{-1} A_1 B = A_1$$

Was nicht verwundert, da A_1 bereits in Jordan-Normalform ist.
Wie oben besprochen, könnten wir aber eine andere Kette wählen.
Für eine solche wäre $B \neq \mathrm{Id}$, trotzdem gilt:

$$B^{-1} A_1 B = A_1$$

Die Matrix B für A_2

Für die Matrix

$$A_2 = \begin{pmatrix} -2 & 15 & -1 \\ -4 & 13 & 0 \\ -3 & 7 & 4 \end{pmatrix}$$

haben wir nur einen verallgemeinerten Eigenraum. Darin haben wir eine Jordan-Kette der Länge 3 gefunden:

$$\left(\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

Sie bildet ein Basistupel. Für diese Basis ist:

$$B = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Nach etwas Rechnen ergibt sich:

$$B^{-1} A_2 B = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

Beispiel mit mehreren Jordan-Ketten (aus Wikipedia)

Als Beispiel betrachte man den Endomorphismus mit der Matrix:

$$A := \begin{pmatrix} 25 & -16 & 30 & -44 & -12 \\ 13 & -7 & 18 & -26 & -6 \\ -18 & 12 & -21 & 36 & 12 \\ -9 & 6 & -12 & 21 & 6 \\ 11 & -8 & 15 & -22 & -3 \end{pmatrix}$$

Es gilt: $\chi_A(t) = (3-t)^5$, $\mathrm{Rank}(A - 3\mathrm{Id}) = 2$, und $(A - 3\mathrm{Id})^2 = \mathbf{0}$.
Die Jordan-Normalform lautet daher:

$$J = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Wir suchen nun die Transformationsmatrix B.

Man wähle $u_1 \notin \mathrm{Kern}(A - 3\mathrm{Id})$ beliebig, z.B.

$$u_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Dann ergibt sich:

$$v_1 := (A - 3\mathrm{Id})u_1 = \begin{pmatrix} 22 \\ 13 \\ -18 \\ -9 \\ 11 \end{pmatrix}$$

Daraus erhalten wir bereits die ersten beiden Spalten der Transformationsmatrix:

$$B := \begin{pmatrix} 22 & 1 & * & * & * \\ 13 & 0 & * & * & * \\ -18 & 0 & * & * & * \\ -9 & 0 & * & * & * \\ 11 & 0 & * & * & * \end{pmatrix}$$

Nun suchen wir die zweite Kette der Länge 2.
Wir wählen

$$u_2 \in \operatorname{Kern}(A - 3\mathrm{Id})^2 \setminus (\operatorname{Kern}(A - 3\mathrm{Id}) \cup \operatorname{Span}\{u_1\})$$

beispielsweise:

$$u_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Dann ist

$$v_2 := (A - 3\mathrm{Id})u_2 = \begin{pmatrix} -16 \\ -10 \\ 12 \\ 6 \\ -8 \end{pmatrix}$$

und die Matrix ist nun:

$$B = \begin{pmatrix} 22 & 1 & -16 & 0 & * \\ 13 & 0 & -10 & 1 & * \\ -18 & 0 & 12 & 0 & * \\ -9 & 0 & 6 & 0 & * \\ 11 & 0 & -8 & 0 & * \end{pmatrix}$$

Schließlich ist der letzte Jordan-Block (Größe 1) an der Reihe.
Man wähle

$$v_3 \in \mathrm{Kern}(A - 3\mathrm{Id}) \setminus \mathrm{Span}\{v_1, u_1, u_2, v_2\}$$

etwa:

$$v_3 := \begin{pmatrix} 2 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Dann ist:

$$B = \begin{pmatrix} 22 & 1 & -16 & 0 & 2 \\ 13 & 0 & -10 & 1 & 0 \\ -18 & 0 & 12 & 0 & 0 \\ -9 & 0 & 6 & 0 & 1 \\ 11 & 0 & -8 & 0 & 0 \end{pmatrix}$$

eine reguläre Matrix mit $J = B^{-1}AB$.