Vorlesung Lineare Algebra II

SoSe'20 hhu
K. Halupczok

§2: Normalformentheorie

ℓ6: Hauptvektorenketten

Stichworte: Normalformproblemstellung, diag'bar + trigon'bar, Hauptvektor m-ter Stufe, Kette von HVen, Jordansche Normalform bei Existenz einer Hauptvektorenketten-Basis

---

6.1. Die Normalformentheorie behandelt die Frage, mit welchen möglichst einfachen Matrizen ein Endomorphismus $f \in \mathrm{End}(K^n)$ zu einer Basis $\mathcal{B}$ dargestellt werden kann: Wie kann eine Basis $\mathcal{B}$ so aufgefunden werden, dass die Matrixdarstellung ${}_{\mathcal{B}}[f]_{\mathcal{B}} = S^{-1}AS$ mit $S = (v_1|v_2|\ldots|v_n)$, wenn $\mathcal{B} = (v_1, v_2, \ldots, v_n)$, vgl. LAI, L17.7/L21.1 so einfach wie nur möglich ist? Wir haben diese Fragestellung in LAI schon behandelt und dabei (mit der Eigenwert-/Eigenvektortheorie) die folgenden zentralen Ergebnisse erzielt:

L20.10. Satz: Zu einem Endo f sei $(v_1, \ldots, v_n)$ eine Basis aus EVen von f zu EWen $\lambda_1, \ldots, \lambda_n$ (nicht notwendig verschieden). Dann hat f bzgl. dieser Basis die Matrixdarstellung

$$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{mit } (a_{ij}) = \begin{cases} \lambda_i, & i = j, \\ 0, & \text{sonst}. \end{cases}$$

(f heißt dann diagonalisierbar).

Dies ließ sich umkehren zu:

L20.12. Satz: Jeder diagonalisierbare Endo f von V mit $\dim V = n$ besitzt eine Familie $(v_1, \ldots, v_n)$ von Eigenvektoren, die eine Basis von V bilden.

(Nicht notwendig gehören diese EVen zu verschiedenen EWen. Aber:)

L20.8. Kor.: Besitzt ein Endo f eines n-dim. K-VRs V genau n verschiedene EWe $\lambda_1, \ldots, \lambda_n$, so bildet eine Familie $(v_1, \ldots, v_n)$ zugehöriger EVen eine Basis von V. Bezüglich einer solchen Basis hat f die Matrixdarstellung $\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ als Diagonalmatrix.

6.2. Fazit/Kurzzusammenfassung:

(i) diag'bar $\iff \exists$ Basis aus EVen, (ii) $\exists\, n$ versch. EWe $\Rightarrow$ diag'bar.

(iii) Nicht jede Matrix ist diag'bar! $\rightsquigarrow$ Frage der "Trigonalisierbarkeit"?

6.3 Schauen wir auf folgenden Spezialfall:

1) Betrachten wir den durch die Dreiecksmatrix $A = \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \lambda_1 & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_1 \end{pmatrix}$, $\lambda_1 \in \mathbb{C}$, gegebenen Endo $f: \mathbb{C}^n \to \mathbb{C}^n$, $x \mapsto Ax$.
Das charakteristische Polynom ist $\chi(T) = (\lambda_1 - T)^n$, also ist λ_1 der einzige Eigenwert von f, und offensichtlich ist e_1 ein EV zum EW λ_1 (siehe 1. Spalte von A).

2) Bestimmen wir alle Eigenvektoren zu λ_1:
Für die Einheitsvektoren $e_i = (\delta_{ji})_{j=1,\dots,n} \in K^n$ rechnen wir nach:

$$f(e_i) = Ae_i = \begin{cases} \lambda_1 e_i, & i = 1, \quad (\text{d.h. } Ae_1 = \lambda_1 e_1) \\ \lambda_1 e_i + e_{i-1}, & i = 2,\dots,n. \quad (\leftarrow i\text{-te Spalte von } A \text{ für } i > 1) \end{cases}$$

Ist nun $x = \xi_1 e_1 + \dots + \xi_n e_n$ ein EV zu λ_1 von f, also $\underline{f(x) = \lambda_1 x}$, so ist $f(x) = \xi_1 f(e_1) + \dots + \xi_n f(e_n) = \xi_1 \lambda_1 e_1 + \sum_{i=2}^{n} \xi_i (\lambda_1 e_i + e_{i-1})$

$$= \lambda_1 (\xi_1 e_1 + \dots + \xi_n e_n) + \sum_{i=2}^{n} \xi_i e_{i-1} = \lambda_1 x + \sum_{i=2}^{n} \xi_i e_{i-1} \overset{!}{=} \lambda_1 x,$$

also muss $\sum_{i=2}^{n} \xi_i e_{i-1} = 0$ sein, also sind $\xi_2 = \dots = \xi_n = 0$ da die e_i lin. unabh., es folgt $x = \xi_1 e_1$.

3) Demnach ist e_1 EV zum einzigen EW λ_1 von f, und es gibt keinen dazu unabh. Eigenvektor zu λ_1, d.h. f ist also (für $n > 1$) nicht diagonalisierbar. Dennoch würden wir sagen, die obige Matrix A ist so "einfach" wie geht. Sie liegt in Dreiecksform vor (d.h. $\alpha_{ij} = 0$ für alle $i > j$, vgl. L14.9.)

4) Untersuchen wir das Verhalten von f über den e_i genauer: Es ist $f(e_1) = \lambda_1 e_1 \Leftrightarrow (f - \lambda_1 \mathrm{id})(e_1) = 0$, d.h. e_1 ist nichttriviales El. ($\neq 0$) von $\ker(f - \lambda_1 \mathrm{id})$.
Weiter ist $(f - \lambda_1 \mathrm{id}) e_2 = f(e_2) - \lambda_1 e_2 = \lambda_1 e_2 + e_1 - \lambda_1 e_2 = e_1$,
und somit $(f - \lambda_1 \mathrm{id})^2 e_2 = (f - \lambda_1 \mathrm{id}) \circ (f - \lambda_1 \mathrm{id})(e_2) = (f - \lambda_1 \mathrm{id}) e_1 = 0$,
d.h. e_2 ist zwar $\notin \ker(f - \lambda_1 \mathrm{id})$, aber dafür gilt $e_2 \in \ker(f - \lambda_1 \mathrm{id})^2$.
Dies setzen wir fort und erhalten für $m = 1, 2, \dots, n$ die Glg.

$$(f - \lambda_1 \mathrm{id})^{m-1}(e_m) \neq 0, \quad \text{aber} \quad (f - \lambda_1 \mathrm{id})^m (e_m) = 0.$$

Diese Überlegungen führen uns zu folgender Definition. Dazu sei V ein K-VR.

6.4. **Def.** (Hauptvektoren): Ist λ ein EW eines Endos $f \in End(V)$, so heißt jeder Vektor $x \in V$ mit $(f-\lambda id_V)^{m-1} x \neq 0$ aber $(f-\lambda id_V)^m x = 0$ ein Hauptvektor m-ter Stufe zum EW λ von f, $m = 1, 2, \ldots$

6.5. **Bem.:** Hauptvektoren 1-ter Stufe sind genau die schon behandelten Eigenvektoren. Das Konzept "Hauptvektor" stellt somit eine Verallgemeinerung von "EV" dar.

6.6. Abkürzung: Im folgenden setzen wir $\boxed{f_\lambda := f - \lambda id_V}$.
Für "Hauptvektor" schreiben wir auch HV.

Einige einfache Aussagen über Hauptvektoren:

6.7. Lemma: (1) Ist x_m ein HV m-ter Stufe zum EW λ, und definieren wir rekursiv/induktiv (absteigend) für $i = m-1, m-2, \ldots$ die Vektoren
$$\boxed{x_i := f_\lambda(x_{i+1}),}$$
so ist x_i ein HV i-ter Stufe zu λ, für $i = 1, \ldots, m$.

(2) Es gilt $f(x_i) = \begin{cases} \lambda x_i + x_{i-1}, & i > 1, \\ \lambda x_1, & i = 1. \end{cases}$

(3) Die Familie $(x_1, x_2, \ldots, x_m)$, d.h. diese von einem HV m-ter Stufe (x_m) erzeugte Kette von HVen, ist linear unabhängig.

Bew.: zu (1): Es ist $x_i = f_\lambda^{m-i}(x_m)$ (*), denn dies ist klar für $i = m$, und gilt (*) für ein i, so auch für $i-1$, denn
$$x_{i-1} = f_\lambda(x_i) = f_\lambda(f_\lambda^{m-i}(x_m)) = f_\lambda^{m-i+1}(x_m) = f_\lambda^{m-(i-1)}(x_m).$$
Damit ist $f_\lambda^{i-1}(x_i) = f_\lambda^{i-1}(f_\lambda^{m-i}(x_m)) = f_\lambda^{m-1}(x_m) \neq 0$, da x_m HV m-ter Stufe,
und $f_\lambda^{i}(x_i) = f_\lambda^{i}(f_\lambda^{m-i}(x_m)) = f_\lambda^{m}(x_m) = 0$, ebenfalls da x_m HV m-ter Stufe.

zu (2): $f(x_i) = f_\lambda^1(x_i) + \lambda x_i = \begin{cases} x_{i-1} + \lambda x_i, & i > 1, \\ \lambda x_1, & i = 1. \end{cases}$

zu (3): Sei $\sum_{j=1}^m \alpha_j x_j = 0$. Dann ist für jedes $k \in \{0, 1, \ldots, m-1\}$ wahr, dass
$$0 = f_\lambda^k\Big(\sum_{j=1}^m \alpha_j x_j\Big) = \sum_{j=1}^k \alpha_j \underbrace{f_\lambda^k(x_j)}_{\substack{=0, \text{ da} \\ x_j \text{ HV } j\text{-ter Stufe}, j \leq k}} + \sum_{j=k+1}^m \alpha_j \underbrace{f_\lambda^k(x_j)}_{\substack{\neq 0, \\ \text{ebenso}}}.$$
Es folgt für jedes k: $0 = \sum_{j=k+1}^m \alpha_j f_\lambda^k(x_j)$.
Setzt man $k := m-1, m-2, m-3, \ldots$, so ergibt sich sukzessive $0 = \alpha_m = \alpha_{m-1} = \ldots = \alpha_1$. □

6.8 Sobald eine Basis von V vorliegt, die aus lauter solchen <u>Ketten von HVen</u> besteht, liegt die Matrix, die den Endo bzgl. dieser Basis darstellt, in der Gestalt vor wie oben in <u>6.3</u> betrachtet. Der folgende Satz beschreibt die dadurch vermittelte Matrixdarstellung genau.

6.9 <u>Satz (Jordansche Normalform bei Existenz einer HVenkettenbasis):</u> Sei V ein n-dim. K-VR, sei $f \in \operatorname{End}(V)$. Weiter seien $\lambda_1, \lambda_2, \dots, \lambda_m$ <u>nicht</u> notwendig verschiedene Eigenwerte von f. Gegeben sei eine Basis $\mathcal{B} = (v_1, v_2, \dots, v_n)$ von V mit folgender Eigenschaft ⊛:

⊛ $\left\{\right.$ Es gibt Indizes $0 =: i_0 < i_1 < \dots < i_m = n$ so, dass für jedes $\mu \in \{1, \dots, m\}$ die Basisvektoren $\underbrace{(v_i \,;\, i_{\mu-1} < i \le i_\mu)}_{=(v_{i_{\mu-1}+1},\, v_{i_{\mu-1}+2}, \dots,\, v_{i_\mu})}$ eine Kette von HVen zum EW λ_μ bilden.

($\leadsto$ haben m HV-Ketten!)

<u>Dann</u> hat f bzgl. $\mathcal{B}$ die Dreiecksmatrixdarstellung

$$A = \begin{pmatrix} C_1 & & & 0 \\ & C_2 & & \\ & & \ddots & \\ 0 & & & C_m \end{pmatrix},$$

wobei darin C_μ eine $\nu_\mu \times \nu_\mu$-Matrix der Form

$$C_\mu = \begin{pmatrix} \lambda_\mu & 1 & & 0 \\ & \lambda_\mu & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda_\mu \end{pmatrix}$$

ist, d.h. λ_μ in der Diagonale und 1 in <u>der darüberliegenden Nebendiagonale.</u>

Diese Darstellung von f heißt die <u>Jordansche Normalform</u> (Kurz: <u>JNF</u>)

6.10 <u>Bem.:</u> 1) Vor. ⊛ ist wesentlich für diesen Satz. Über algebraisch abgeschlossenen Körpern, also etwa $K = \mathbb{C}$, werden wir die Existenz dieser Jordanschen Normalform beweisen, d.h. also dass dann eine Basis mit ⊛ gefunden, ja sogar konstruiert werden kann.

Dafür werden wir im Rest des §2 den Beweis geben, indem wir einen allgemeinen Normalformensatz beweisen werden, woraus sich die Existenz der Jordanschen Normalform als Spezialfall für $K = \mathbb{C}$ ableiten lässt.

2) Die oben mit Indizes formulierte Vor. ⊛ kann noch etwas einfacher ausgedrückt werden: Die Basisvektoren liegen blockweise "sortiert" vor in der Form

die λ_μ nicht notw. verschieden!

$$\Big(\underbrace{v_1, v_2, \dots, v_{i_1}}_{\text{HVenkette zu } \lambda_1},\ \underbrace{v_{i_1+1}, \dots, v_{i_2}}_{\text{HVenkette zu } \lambda_2},\ \underbrace{v_{i_2+1}, \dots, v_{i_3}}_{\text{HVenkette zu } \lambda_3},\ \dots\dots,\ \underbrace{v_{i_{m-1}+1}, \dots, v_{i_m}}_{\text{HVkette zu } \lambda_m}\Big)$$

Länge einer HV-Kette

3.) Mit $j_\mu := i_\mu - i_{\mu-1}$ ist also $v_{i_{\mu-1}+j}$ ein HV j-ter Stufe zu λ_μ, wo $j = 1, 2, \ldots, j_\mu$, d.h. es gilt

$$f\left(v_{i_{\mu-1}+j}\right) = \begin{cases} \lambda_\mu v_{i_{\mu-1}+1} & , \ j = 1, \\ \lambda_\mu v_{i_{\mu-1}+j} + 1 \cdot v_{i_{\mu-1}+j-1}, & j > 1. \end{cases}$$

→ Die Koeffizienten ergeben die Spalteneinträge der Jordan-Blockmatrix $C_\mu = \begin{pmatrix} \lambda_\mu & 1 & & \\ & \lambda_\mu & 1 & \\ & & \lambda_\mu \ddots & 1 \\ & & & \lambda_\mu \end{pmatrix}$.
(auch: "Jordan-Kästchen" genannt)

6.11. Bew. des Satzes: • v_{i_μ} ist also HV j_μ-ter Stufe zum EW λ_μ, und die Vektoren des entsprechenden Blockabschnitts bilden eine von v_{i_μ} erzeugte Kette. Ist φ die von der Basis vermittelte Koordinatenabbildung $\varphi: V \to K^m$, $v_i \mapsto e_i$, d.h. $\varphi(x) = {}_{\mathcal{B}}[x]$, gilt für die dadurch gegebene Matrixdarstellung ja gerade, dass die j-te Spalte von A gleich $Ae_j = (\varphi f \varphi^{-1})(e_j) = \varphi(f(v_j))$ ist, vgl. auch LA I, L 17.7. Dies kann hier leicht angegeben werden:

$Ae_1 = \varphi(f(v_1)) = \varphi(\lambda_1 v_1) = \lambda_1 e_1$, d.h. λ_1 ist in der Diagonalen,

ist $i_1 > 1$, so folgt
$$\left.\begin{aligned} Ae_2 &= \varphi(f(v_2)) = \varphi(\lambda_1 v_2 + v_1) = \lambda_1 e_2 + 1 \cdot e_1 \\ &\vdots \\ Ae_{i_1} &= \varphi(f(v_{i_1})) = \varphi(\lambda_1 v_{i_1} + v_{i_1-1}) = \lambda_1 e_{i_1} + 1 \cdot e_{i_1-1} \end{aligned}\right\} \lambda_1 \text{ in der Diagonalen, 1en darüber}$$

Für die Spalten $j = 1, 2, \ldots, i_1$ gilt also die Beh.

• v_{i_1+1} ist wieder EV von f, und zwar zum EW λ_2 (es kann $\lambda_2 = \lambda_1$ sein). Damit ist die nächste Spalte $Ae_{i_1+1} = \varphi(f(v_{i_1+1})) = \varphi(\lambda_2 v_{i_1+1}) = \lambda_2 e_{i_1+1}$, d.h. λ_2 steht dort in der Diagonalen. Ist $i_2 > i_1 + 1$, so kommen wieder Hauptvektoren zu λ_2 höherer Stufe, wie eben folgt

$$Ae_{i_1+2} = \varphi(f(v_{i_1+2})) = \varphi(\lambda_2 v_{i_1+2} + v_{i_1+1}) = \lambda_2 e_{i_1+2} + e_{i_1+1}, \text{ usw.}$$

• Bis zum Index $i_2 + 1$, der dann zum nächsten EW gehört, erhalten wir also wieder λ_2 in der Diagonalen und darüber eine 1, womit die Beh. für die ersten i_2 Spalten richtig sind.

• Fortsetzung dieses Verfahrens führt den Beweis zu Ende. □

6.12. Wir werden noch sehen, dass bei diagonalisierbaren Endos keine echten HV auftreten, also nur EVen = HVen 1. Stufe. Dann haben die in der Jordanschen Normalform vorkommenden Kästchen-Matrizen C_μ alle nur die Dimension 1, d.h. die zugehörige Jordan-Normalform wird dafür insgesamt zu einer Diagonalmatrix. Daher erweist sich Satz L20.10 als ein Spezialfall des Satzes 6.9 über die JNF.

6.13. Bsp.: 1.) Die folgenden Matrizen liegen in JNF vor (nichtgeschriebene Einträge sind $=0$):

$$\underbrace{\begin{pmatrix} 2 & & \\ & \boxed{2} & \\ & & 3 \end{pmatrix}}_{\lambda_1=2,\ \lambda_2=3}, \underbrace{\begin{pmatrix} 2 & 1 & \\ & 2 & \\ & & 3 \end{pmatrix}}_{\lambda_1=2,\ \lambda_3=3}, \underbrace{\begin{pmatrix} 2 & & & \\ & 2 & 1 & \\ & & 2 & \\ & & & 3 \end{pmatrix}}_{\lambda_1=2,\ \lambda_3=3}, \underbrace{\begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}}_{\lambda_1=1,\ \lambda_2=2}, \underbrace{\begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}}_{\lambda_2=2}, \underbrace{\begin{pmatrix} 2 & 1 & & \\ & 2 & 1 & \\ & & 2 & \\ & & & 2 \end{pmatrix}}_{\lambda_2=2}, \underbrace{\begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix}}_{\lambda_1=0}, \ldots$$

2.) Gegeben sei der Endo f, geg. durch die Matrix
$$F = \begin{pmatrix} 17 & 0 & -25 \\ 0 & 3 & 0 \\ 9 & 0 & -13 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Wollen, falls möglich, die JNF bestimmen. Dazu benötigen wir eine Basis aus HVenketten.

1. Bestimmung der EWe: Das charakteristische Polynom lautet

$$\chi_F(\lambda) = \det(F - \lambda I_3) = \begin{vmatrix} 17-\lambda & 0 & -25 \\ 0 & 3-\lambda & 0 \\ 9 & 0 & -13-\lambda \end{vmatrix} = (3-\lambda)\cdot\big(-(17-\lambda)(13+\lambda) - 9\cdot(-25)\big)$$
$$= (3-\lambda)\cdot(\lambda-2)^2.$$

$\uparrow$ Entwickeln hier

Also EWe: $\lambda_1 = 3$, $\lambda_2 = 2$ (doppelt).

2. Bestimmung der EVen/der Eigenräume:

- $\underbrace{\text{Zu } \lambda_1=3}$: Das LGS $(F - 3\cdot I_3)x = o$ hat die Lösung $L(v_1)$ mit genau einem unabh. v_1, d.h. haben als EV $v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, denn $(F - 3\cdot I_3)x = o$ $\iff \begin{pmatrix} 14 & 0 & -25 \\ 0 & 0 & 0 \\ 9 & 0 & -16 \end{pmatrix} x = 0$ hat die Lösungsmenge $L(v_1) = \mathbb{R}v_1$.

- $\underbrace{\text{Zu } \lambda_2=2}$: Für $(F - 2\cdot I_3)x = o \iff \begin{pmatrix} 15 & 0 & -25 \\ 0 & 1 & 0 \\ 9 & 0 & -15 \end{pmatrix} x = o$ existiert wegen $\operatorname{rg}(F - 2I_3) = 2$ (= maximale # lin. unabh. Spalten) genau eine unabh. Lösung $L(v_2)$, nämlich mit $v_2 = \begin{pmatrix} 5 \\ 0 \\ 3 \end{pmatrix}$. Da aber $\underline{v_2 \in \operatorname{im}(F - 2I_3)}$ ist, hat auch das LGS $(F - 2I_3)x = v_2$ eine Lösung,
nämlich: $\iff \begin{pmatrix} 15 & 0 & -25 \\ 0 & 1 & 0 \\ 9 & 0 & -15 \end{pmatrix} x = \begin{pmatrix} 5 \\ 0 \\ 3 \end{pmatrix}$, $\underline{\text{etwa } v_3 = \begin{pmatrix} 1/3 \\ 0 \\ 0 \end{pmatrix}}$.

Dann ist $(F-2I_3)v_3 = v_2 \neq 0$,

aber $(F-2I_3)^2 v_3 = (F-2I_3)v_2 = 0$,

somit ist mit v_3 ein HV 2. Stufe zum EW $\lambda_2 = 2$ von F gefunden.

Insgesamt bilden (v_1, v_2, v_3) eine Basis von $\mathbb{R}^3$,

in der v_2, v_3 eine HV-Kette zu $\lambda_2 = 2$ bilden.

Nach Satz 6.9 hat der von F gegebene Endo $f: \mathbb{R}^3 \to \mathbb{R}^3,\ x \mapsto Fx$, bezüglich dieser Basis somit die Matrixdarstellung

$$F' = \left(\begin{array}{c|cc} 3 & 0 & 0 \\ \hline 0 & 2 & 1 \\ 0 & 0 & 2 \end{array}\right),$$

$$\begin{array}{cccc} & \uparrow & \uparrow & \uparrow \\ \text{zu} & v_1 & v_2 & v_3 \end{array}$$

diese liegt vor in Jordanscher Normalform.

Ü Rechnen Sie nach, dass mit der Matrix $S = (v_1 | v_2 | v_3) = \begin{pmatrix} 0 & 5 & 1/3 \\ 1 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$ gilt: $S^{-1} F S = F'$.

<u>Bem.:</u> Auch $v_3' := \begin{pmatrix} 0 \\ 0 \\ 1/5 \end{pmatrix}$ wäre ein HV gewesen, der mit v_2 eine Kette bildet. Man kann offenbar verschiedene (un.abh.) HVen in die Basis aufnehmen.