# 1 TESTMATERIAL

Spezialfall: Wie drückt sich der Basiswechsel in den Koordinaten eines Vektor aus? D.h. angenommen, man hat 2 verschiedene Basen $\mathscr{B}_1$ und $\mathscr{B}_2$ in einem (abstrakten oder konkreten) Vektorraum. Wie kann man dann für $\mathbf{v} \in \boldsymbol{V}$ die Umrechnung von $[\mathbf{v}]_{\mathscr{B}_1}$ nach $[\mathbf{v}]_{\mathscr{B}_2}$ bewerkstelligen.

A.Cap verwendet folgende Schreibweise (aus gutem Grunde):

$$[T]^{\widetilde{\mathscr{B}}}_{\widetilde{\mathscr{C}}} = [Id_W]^{\mathscr{C}}_{\widetilde{\mathscr{C}}} * [T]^{\mathscr{B}}_{\mathscr{C}} * [Id_V]^{\widetilde{\mathscr{B}}}_{\mathscr{B}}$$

In der "Feichtinger Notation" (weniger kompakt, aber klarer an das sog. *Domino-Prinzip* erinnernd!) sieht das so aus:

$$[T]_{\widetilde{\mathscr{C}} \leftrightarrow \widetilde{\mathscr{B}}} = [Id_W]_{\widetilde{\mathscr{C}} \leftrightarrow \mathscr{C}} * [T]_{\mathscr{C} \leftrightarrow \mathscr{B}} * [Id_V]_{\mathscr{B} \leftrightarrow \widetilde{\mathscr{B}}}. \tag{1}$$ baschangboth

Wollen wir also konkrete Fälle betrachten:

Einerseits können wir verschiedene Basen nehmen, und dieselbe lineare Abbildung in verschiedenen Formen darstellen, um bessser zu verstehen, wie dieselbe lineare Abbildung durch vollkommen (?) verschiedene Matrizen dargestellt werden kann. Andererseits können kompliziertere Matrizen durch einfacher (alternative) Darstellung besser verstanden werden.

Aufgabe: Es seien die quadratischen Lagrange-Interpolationspolynome für die Folge $-1, 0, 1$ (der einfachste Fall von drei Werten in $\mathbb{R}$) gegeben. Dies sind die Polynome $q_1(x) = (x-0)(x-1), q_2(x) = (x-(-1))(x-1)$ und $q_3 = (x-(-1))(x-0)$, richtig normalisiert, um sicherzustellen dass $L_k(x_k) = 1$ gilt, also $L_k(x) = q_k(x)/q_k(x_k)$:

$$L_1(x) = (x^2 - x)/2, L_2(x) = -x^2 + 1, L_3(x) = (x^2 + x)/2. \tag{2}$$ LagrBasis

Das zugehörige MATLAB Experiment würde wie folgt aussehen (um es besser lesen zu können der Hinweis, dass MATLAB für $\mathscr{P}_2(\mathbb{R})$ mit er Basis $\{x^2, x, 1\}$ arbeitet (also die Monome in absteigender Potenz):

```
V1 = vander([-1,0,1])
V1 =
     1    -1     1
     0     0     1
     1     1     1
>> IV1 = inv(V1)
IV1 =
    0.5000   -1.0000    0.5000
   -0.5000         0    0.5000
         0    1.0000         0
```

Das ist so zu lesen: Die zweite Spalte ist die Folge $[-1, 0, 1]$, dem also das Lagrange-Interpolationspolynom $L_2(x) = -x^2 + 1$ entsprecht (der Koeffizient von $x$ ist 0!), das offensichtlich das eindeutig bestimmte quadratische Polynom auf $\mathbb{R}$ ist mit $L_2(0) = 1, L_2(\pm 1) = 0$, und entsprechend für die Spalten 1 und 3.

Diese drei Polynomfunktionen bilden eine Basis für $\mathscr{P}_2(\mathbb{R})$[1]: Ihre Koeffizienten bilden zusammen die inverse Vandermonde-Matrix, also ein invertierbare Matrix. Ganz allgemein gilt aber: bildet man aus mit Hilfe der Koeffizienten einer invertierbaren Matrix (hier IV1) aus einer Basis (hier $\{x^2, x, 1\}$) eine Familie (hier $L_1(x), L_2(x), L_3(x)$) von "Vektoren", d.h. $n = \dim(\mathscr{P}_2(\mathbb{R})) = 3(!)$ im Vektorraum $\boldsymbol{V} = \mathscr{P}_2(\mathbb{R})$, dann sind diese wieder eine Basis, denn beispielsweise ist es leicht zu zeigen, dass sie wieder linear unabhängig sind (Details dem Leser überlassen). Weil die Anzahl die richtige ist (war haben 3 Funktionen) ist es eine Basis. Einfacher ist vielleicht das folgende Argument: Die Monone selbst können als Linear-Kombinationen der Lagrange Interpolationspolynome gefunden werden. In der Tat, die einfache Tatsache dass $IV1 * V1 = Id_3$ ergibt (es ist ja die inverse) besagt auch, dass die Linearkombination die aus den Lagrange-Interpolationen gebildet werden mit Hilfe der Koeffizienten aus den Spalten der Vandermonde-Matrix (wir nehmen die erste Spalte von $V1$), gerade die Basisvektoren $x^2$ (bzw. $x$ oder 1) sind. Konkret: die erste Spalte von $V1$ ist der Spaltenvektor $[1; 0; 1]$, also gilt

$$1 \cdot L_1(x) + 0 \cdot L_2(x) + 1 \cdot L_3(x) = 1 \cdot x^2 + 0 \cdot x + 0 \cdot 1$$

oder einfach

$$x^2 = L_1(x) + L_3(x),$$

was (im Nachhinein und in dem konkreten Fall) sogar mit "freiem Auge" zu sehen gewesen wäre!

Will man nun die (offenbar linearen) Differentiationsabbildung $T : p(x) \mapsto p'(x)$ in verschiedenen Koordinatensystemen (mit unterschiedlichen Basen) beschreiben, haben wir nun unterschiedliche Möglichkeiten:

- Wenn man sich die Basis aussuchen darf, dann kann man versuchen, eine möglichst einfache (vom Standpunkt der Beschreibung von $T$). Das ist im gebenen Fall wohl die Monomial Basis. Wir haben $T(x^2) = 2x, T(x) = 1, T(1) = 0$, das Nullpolynom. Also ist die Matrix bezüglich der Monomialbasis $\mathscr{M} = \{x^2, x, 1\}$ einfach die Matrix, deren Spalten genau die Koeffizienten der Bilder der Basisvektoren sind, nun in ebenderselben Basis, das ist die Matrix $[T]_{\mathscr{M} \leftarrow \mathscr{M}}$, wir nennen sie kurz $M$ ist gleich:

  $$\begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{3}$$

  Das entspricht auch der Tatsache, dass der Nullraum von $T$ genau der Teilraum aller konstanten Polynomfunktionen ist (die skalaren Vielfachen von 1), und der Rang von $M$ gleich 2 (alle linearen Polynomfunktionen treten als Bilder auf, und die bilden natürlich einen $2-$dimensionalen Teilraum). Weiters hat man für $M * M$:

  $$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix} \tag{4}$$

  und $M^3 := M * M * M$ ist klarerweise die Nullmatrix.

[1] vermutlich kann man eine ganze Reihe von Argumenten dafür finden, die vorgeschlagene Version ist nur im Moment und unter den gegebenen Voraussetzung vermutlich/hoffentlich die einfachste!

- Wollen wir dieselbe Abbildung nun in der Lagrange-Basis $\mathscr{L} := \{L_1(x), L_2(x), L_3(x)\}$ bestimmen, welche in der allgemeinen Systematik die Bezeichnung $[T]_{\mathscr{L}\leftrightarrow\mathscr{L}}$ trägt. Wir schreiben $N$ für diese Matrix. Zuerst müssen wir die Ableitungen der Lagrange Interpolationspolynome bilden, also

$$L_1'(x) = x - 1/2; \quad L_2'(x) = -2x; \quad L_3'(x) = x + 1/2.$$

  Im Sinne der üblichen Monome haben wir also als Matrix für $[T]_{\mathscr{M}\leftrightarrow\mathscr{L}}$ die Matrix $LM$:

$$\begin{pmatrix} 0.0000 & 0.0000 & 0.0000 \\ 1.0000 & -2.0000 & 1.0000 \\ 0.5000 & 0.0000 & 0.5000 \end{pmatrix} \tag{5}$$

  Um daraus die Matrix $[T]_{\mathscr{L}\leftrightarrow\mathscr{L}}$ zu machen, müssen wir so noch (von links) mit der Basis-Wechsel-Matrix von $\mathscr{M}$ nach $\mathscr{L}$ kombinieren, welches aber genau die Vandermonde-Matrix $V1$ ist, also bekommen wir $V1 * LM$, und somit ist $N$ als

$$\begin{pmatrix} -1.5000 & 2.0000 & -0.5000 \\ -0.5000 & 0.0000 & 0.5000 \\ 0.5000 & -2.0000 & 1.5000 \end{pmatrix} \tag{6}$$

  LLbas

  gegeben.

- Die direkte Umsetzung der Formel (baschangboth 1) ergibt direkt

$$N = V1 * M * IV1,$$

  was nach Ausmultiplizieren zur Matrix (LLbas 6) führt. Klarerweise gilt auch

$$LM = M * IV1.$$

Die folgende Rechnung ist nicht ganz überraschend:

```
>> LL*LL  ( =  V1 * M * M * IV1 )
ans =
     1    -2     1
     1    -2     1
     1    -2     1
>> LL*LL*LL
ans =
     0     0     0
     0     0     0
     0     0     0
```

Wir wollen als nächstes ein paar Isomorphismen (d.h. bijektive, linear Abbildungen) von $\boldsymbol{V}$ in sich für einen typischen Speziallfall betrachten, nämlich (wieder) $\mathscr{P}_2(\mathbb{R})$:

Einerseits die Verschiebungsmatrix, andererseits die Streckungsmatrix. Bekanntlich ist mit jeder Polynomfunktion $p(x)$ (vom Grad $r$) auch die verschobene Funktion eine Polynomfunktion $q(x)$ vom gleichen Grad, dasselbe gilt für die Streck-Operation.

Wir wollen also zwei entsprechende lineare Operationen (Abbildungen von $\mathscr{P}_2(\mathbb{R})$ in sich) auf $\boldsymbol{V} = \mathscr{P}_2(\mathbb{R})$ definieren, und zwar die Translation um $z \in \mathbb{R}$:

**Definition 1.**

$$T_z(p(x)) = r(z) := p(x - z) \quad x, z \in \mathbb{R}; \tag{7}$$

transldef1

und die Dehnung (for $\rho \geq 1$, bzw. Drück-Operation für $\rho < 1$):

$$D_\rho(p(x)) = q(x) := p(x/\rho), \quad \rho > 0, x \in \mathbb{R}. \tag{8}$$

dehndef1

Es ist klar (?!), dass diese Operationen linear sind, d.h. der Vektorraumstruktur von $\mathscr{P}_2(\mathbb{R})$ verträglich (genauso könnte man sie auf dem Raum von Polynomfunktionen vom Grad $r, r \in \mathbb{N}$ betrachten, oder auf dem Vektorraum aller stetigen Funktionen auf $\mathbb{R}$, mit Werten in $\mathbb{R}$ oder genauso in $\mathbb{C}$).

Die gewählten Bezeichnungen sind so, wie man es sich bei Betrachtung der Graphen der Funktionen $p(x)$ erwarten würden. Für $z = 5$ werden die Werte von $T_5(p(x))$ genau dieselben Werte sein, die $p(x)$ um 5 Einheiten weiter links hat, also wird der Graph um den Wert $z = 5$ nach rechts *translatiert*. Für den Fall $z = -3$ wird (der Graph von) $p(x)$ um 3 nach *links* verschoben. Ebenso kann man leicht feststellen, dass $D_2$ den Graphen um den Faktor 2 dehnt, d.h. z.B. dass die Nullstellen, lokalen Extremwerte etc. genau um den Faktor gedehnt werden.

Es ist ebenfalls klar, dass die Abbildungen bijektiv sind, denn es ist leicht, die inverse Abbildung anzugeben:

$$T_z^{-1} = T_{-z} \quad z \in \mathbb{R}; \qquad D_\rho^{-1} = D_{1/\rho}, \quad \rho > 0. \tag{9}$$

invtransdil

In bezug auf die Standard Basis von $\mathscr{P}_2(\mathbb{R})$ ist leicht zu erkennen, dass aus $T_z$ (mein denke sich beim ersten Lesen einen konkreten Wert von $z$, sagen wir $z = 2$):

$$T_z(x^2) = (x - z)^2, T_z(x) = (x - z), T_z(1) = 1,$$

somit ist die Matrix $MZ = [T_z]_{\mathscr{M} \leftarrow \mathscr{M}}$ gegeben durch die Dreiecksmatrix

$$\begin{pmatrix} 1 & 0 & 0 \\ -2z & 1 & 0 \\ z^2 & -z & 1 \end{pmatrix} \tag{10}$$

Man kann sich leicht (im Sinne eine Plausibilitätstestes, z.B. durch die Wahl von $z = 1$ davon überzeugen, dass die inverse Matrix dazu genau die Matrix zu $T_{-z}$ ist (die sich lediglich von der obigen Matrix durch einen Vorzeichenwechsel an den beiden Positionen, in denen $z$ auftritt, bemerkbar macht).

```
>> MZ
MZ =
     1     0     0
    -2     1     0
     1    -1     1
>> inv(MZ)
ans =
     1     0     0
     2     1     0
     1     1     1
```

Die Matrix $[D_\rho]_{\mathscr{M}\leftarrow\mathscr{M}}$ ist klarerweise einfach eine Diagonalmatrix zu $[\rho^2, \rho, 1]$. Also für den Fall $\rho = 3$ einfach $D3 = diag([9, 3, 1])$; also

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{11}$$

## 2 Basen für den Dualraum

Wir haben bisher gesehen (kurze Wiederholung), dass zu jeder Basis $\mathscr{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ eines $n-$dimsionalen Vektorraumes $\boldsymbol{V}$ eine duale Basis existiert, d.h. eine Folge von linearen Funktionalen $\phi_1, \ldots, \phi_n$ im Dualraum von $\boldsymbol{V}^*$ zu $\boldsymbol{V}$ (linearen Abbildungen von $\boldsymbol{V}$ in the Grundkörper $\mathbb{K}$), welche jedem Vektor $\mathbf{v} \in \boldsymbol{V}$ (der ja eine eindeutige Basisdarstellung der Form

$$\mathbf{v} = \sum\nolimits_{k=1}^{n} c_k \mathbf{b}_k$$

hat, die eindeutig bestimmten Koeffizienten $c_k$ zuordnet, d.h.

$$\phi(\mathbf{v}) = c_k, 1 \leq k \leq n.$$

Wie wir gesehen haben, lassen sich beliebige lineare Funktionale (die Gesamtheit derselben bildet einen Vektorraum, mit gewöhnlicher punktweiser Addition von Funktionen mit Werten in $\mathbb{K}$!) als Linearkombinationen derselben bilden. Andererseits sind sie linear unabhängig (?warum?, Übung!), und daher eine *Basis* für den Dualraum $\boldsymbol{V}^*$.

Wir wollen das ein wenig konkreter machen. Wie testet man die lineare Unabhängigkeit einer Folge von $\psi_1, \ldots, \psi_k$ von linearen Funktionalen, also in $\boldsymbol{V}^*$?

**Lemma 1.** *Eine Folge $\psi_1, \ldots, \psi_k$ in $\boldsymbol{V}^*$ ist linear unabhängig, wenn für eine (bzw. jede) Basis $\mathscr{B} = \mathbf{b}_1, \ldots, \mathbf{b}_k$ von $\boldsymbol{V}$ die Matrix*

$$a_{j,k} := \psi_j(\mathbf{b}_k), 1 \leq j \leq k, 1 \leq k \leq n$$

*vollen Rang hat, d.h. Rang $r = max(k, n)$ hat.*

*Proof.* Wenn eine Folge $\psi_1, \ldots, \psi_k$ in $\boldsymbol{V}^*$ linear abhängig ist, dann gibt es Koeffizienten $c_1, \ldots, c_k$, sodass

$$\sum_{j=1}^{k} c_j \psi_j = 0, \quad \text{d.h.} \quad \sum_{j=1}^{k} c_j \psi_j(\mathbf{b}_r) = 0, 1 \leq r \leq n, \tag{12}$$ psilindep

aber das bedeutet, dass $\boldsymbol{A} * \mathbf{c} = 0$ gilt, d.h. $\boldsymbol{A}$ hat linear abhängige Spalten, und daher nicht maximalen Rang (Details!?).

Umgekehrt sind die Elemente $\psi_1, \ldots, \psi_k$ in $\boldsymbol{V}^*$ linear unabhängig, wenn aus dem Verschwinden der Linearkombination dieser linearen Funktionale folgt, dass die Koeffizientenfolge $(c_j)_{j=1}^{k}$ die Nullfolge ist. Aber eine Folge von Funktionalen ist gleich das Nullfunktional, *genau dann* wenn dies auf einer beliebigen Basis das Nullfunktional ergibt, was aber wieder gleichbedeutend damit ist, dass $\boldsymbol{A} * \mathbf{c} = 0$ impliziert, dass $\mathbf{c} = \mathbf{n}$ ist. □

Prinzip: Sie $\mathscr{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_n]$ eine (bel.) Basis für $\boldsymbol{V}$ dann gilt:

$$\sum\nolimits_{k=1}^{n} c_j \psi_j = \mathbf{n} \in \boldsymbol{V}^* \Leftrightarrow \sum\nolimits_{k=1}^{n} c_j \psi_j(\mathbf{b}_l) = 0. \tag{13}$$ linindepfunct

Definiert man also die Matrix $\boldsymbol{A} = (a_{j,l}) = (\psi_j(\mathbf{b}_l))$ dann gilt (13) genau dann, wenn gilt:

$$\boldsymbol{A} * \mathbf{c} = \mathbf{n} \in \mathbb{K}^k. \tag{14}$$ linindepfunct

Die lineare Unabhängigkeit der Funktionale bedeutet, dass (13) impliziert, dass $\mathbf{c} = \mathbf{n}$ gilt, was aber gleichbedeutend damit ist, dass die Spalten von $\boldsymbol{A}$ linear unabhängig sind, wegen (14). Insbesondere ist eine Folge $\psi_1, \ldots, \psi_n$ in $\boldsymbol{V}^*$ eine Basis für $\boldsymbol{V}^*$ genau dann, wenn für eine (und daher für jede) Basis $\mathscr{B}$ in $\boldsymbol{V}$ die zugehorige Matrix invertierbar ist.

biorthbasV **Lemma 2.** *Es sei $\Phi = [\phi_1, \ldots, \phi_n]$ eine beliebige Basis in $\boldsymbol{V}^*$. Dann gibt es eine Basis $\mathscr{B}$ von $\boldsymbol{V}$, sodass die $\Phi$ die zu $\mathscr{B}$ duale Basis ist, d.h. jede Basis von $\boldsymbol{V}^*$ ist eine duale Basis. Weiters ist diese Relation reflexiv, d.h. die duale der dualen Basis (in beide Richtungen) ist wieder die ursprüngliche Basis!* [2]

*Proof.* xxx □

dualbasxx **Theorem 1.** *Es sei $\boldsymbol{V}$ ein $n-$dim. Vektorraum. Weiters sei Es sei $\Phi = [\phi_1, \ldots, \phi_n]$ eine Folge in $\boldsymbol{V}^*$ und $\mathbf{b}_1, \ldots, \mathbf{b}_n$ eine Folge von Elementen in $\boldsymbol{V}$. Wir bilden wieder die Matrix*

$$\boldsymbol{C} = (c_{j,k}) = \phi_j(\mathbf{b}_k).$$

*Wenn die Matrix $\boldsymbol{C}$ invertierbar ist, dann ist die Folge $\Phi := (\phi_j)_{j=1}^n$ eine Basis für $\boldsymbol{V}^*$ und $\mathscr{B} := (\mathbf{b}_k)_{k=1}^n$ ist eine Basis für $\boldsymbol{V}$.*

*Die zu $\Phi$ bzw. $\mathscr{B}$ dualen Basen können mit Hilfe der Matrix $\mathbb{H} := \boldsymbol{C}^{-1}$ berechnet werden.*

*Um aus $\Phi$ die zu $\mathscr{B}$ duale Basis $\widetilde{\mathscr{B}}$ zu finden, müssen die Koeffizienten der Zeilen von $\mathbb{H}$ verwendet werden. Umgekehrt enthalten die Spalten von $\mathbb{H}$ die Information, wie man aus $\mathscr{B}$ die zu $\Phi$ biorthogonale Basis $\widetilde{\Phi} = [\widetilde{\phi_1}, \ldots, \widetilde{\phi_n}]$ bestimmt werden können.*

*Wir haben also*

$$\widetilde{\phi_j}(\mathbf{b}_k) = \delta_{j,k} = \phi_j(\widetilde{\mathbf{b}_k}),\ 1 \le j,k \le n.$$

[2]Im klassischen Fall von $\boldsymbol{V} = \mathbb{R}^n$ gilt bekanntlich: das Biorthogonalsystem zu einer Basis besteht aus den Spalten der transponiert konjugierten Inversen Matrix! Man kann sich leicht überlegen, dass die Operation

$$\boldsymbol{A} \mapsto (\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$$

eine Involution auf der Menge der invertierbaren $n \times n$-Matrizen ist, d.h. zweimal angewendet bekommt man wieder die ursprüngliche Matrix, weil einfach $\boldsymbol{B}'' = \boldsymbol{B}$, und $\boldsymbol{C}^{-1^{-1}} = \boldsymbol{C}$ gilt.

# 3 Basiswechsel, Diagonalisierung und Eigenwerte

Als Vorbereitung f.d. Beschreibung der Diagonalisierbarkeit von (symmetrischen bzw. normalen) Matrizen wollen wir vorerst einmal die Frage der Diagonalisierbarkeit von Matrizen diskutieren.

Ganz am Anfang fragen wir nur: was sind die (gemeinsamen) Eigenschaften von verschiedenen Matrizen, die dieselbe lineare Abbildung beschreiben. Erinnern wir uns zuerst an den Fall von Matrix-Multiplikation: Es sei $T(x) = \boldsymbol{A} * \mathbf{x}$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ (wie wir wissen ist das der allgemeine Falle einer linearen Abbildung). Der Einfachheit wollen wir nur den Fall $m = n$ betrachten, und vorerst die gleiche Basis auf beiden Seiten.

In der abstrakten Sichtweise, und unter Verwendung des Symbols $\mathscr{E}$ für die Standardbasis können wir auch schreiben: gegeben ist $[T]_{\mathscr{E}\leftarrow\mathscr{E}}$, kann man sich natürlich fragen, wie die Darstellung derselben linearen Abbildung in einer *anderen Basis* aussehen würde, als z.B. $[T]_{\mathscr{B}\leftarrow\mathscr{B}}$, für eine Basis $\mathscr{B}$, welche aus den Spalten einer Matrix $\boldsymbol{B}$ bestehen möge. Dann gilt

$$[T]_{\mathscr{B}\leftarrow\mathscr{B}} = \boldsymbol{B} * \boldsymbol{A} * \boldsymbol{B}^{-1}. \qquad (15)$$ baschange3

Auch die Begründung ist einfach: Wir müssen nur den Basiswechsel als eine Realisierung der *identischen* Abbildung in $\mathbb{R}^n$ ansehen. Sind die Koeffizienten $\mathbf{c}$ gegeben, so bildet man daraus den Vektor durch Bilden der Linearkombination (d.h. Matrixmultiplikation), umgekehrt bestimmt die inverse Matrix den Übergang von Vektoren (in der Standard-Darstellung) zu Koordinaten, d.h. $[\mathbf{1}]_{\mathscr{E}\leftarrow\mathscr{B}} = \boldsymbol{B}^{-1}$.

**Theorem 2.** *Die Eigenschaften, welche unabhängig von der Basis gelten, sind etwa die folgenden:*

- *Spalten der Matrix spannen den ganzen $\mathbb{R}^n$ auf;*
- *Der Rang der Matrizen ist unabhängig von der Darstellung;*
- *Die Dimension des Nullraumes ist unabhängig von $\boldsymbol{B}$;*
- *(wie wir noch sehen werden:) die Determinante zur Abbildung;*

Es ist nun naheliegend, all diese Matrizen als äquivalent zu betrachten. Formal lautet die Definition so:

**Definition 2.** Zwei Matrizen $\boldsymbol{A}_1$ und $\boldsymbol{A}_2$ heißen ähnlich, wenn es eine invertierbare Matrix $\boldsymbol{B}$ gilt, wenn es eine invertierbare Matrix $\boldsymbol{C}$ gibt, sodass

$$\boldsymbol{A}_1 = \boldsymbol{C} * \boldsymbol{A}_2 * \boldsymbol{C}^{-1} \qquad (16)$$ equivmatrdef1

gilt. Wir schreiben in dem Fall auch $\boldsymbol{A}_1 \simeq \boldsymbol{A}_2$.

Die soeben definierte Relation "$\simeq$" ist tatsächlich eine Äquivalenz-Relation, d.h. symmetrisch, reflexiv und transitiv. [eigentlich zu beweisen: > Übungsaufgabe!]

Ein paar wesentlich Fakten seien hier doch noch zusammengefasst:

**Theorem 3.** *Es sei* $\boldsymbol{A}$ *irgendeine (reelle bzw. komplexe)* $n \times n-$*Matrix, und* $T : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ *die eindeutig assoziierte lineare Abbildung.*

*Dann besteht die Menge der äquivalenten Matrizen genau aus der Kollektion der Matrizen der Form*

$$\{[T]_{\boldsymbol{B}\leftarrow\boldsymbol{B}}, \mid \boldsymbol{B} \text{ ist Basis in } \mathbb{R}^n\}.$$

baschange3

*Proof.* Der Beweis ist ganz einfach eine Folge von (15). □

Insbesondere sind die im vorigen Theorem angeführten Eigenschaften für äquivalente Matrizen (gleichzeitig) zutreffend.

Wir fragen nun als nächstes: Welche Matrizen sind zu Diagonalmatrizen äquivalent. Um diese Diskussion führen zu können, verwenden wir noch die Begriffe **Eigenvektor** und **Eigenwert** zu einer linearen Abbildung.

**Definition 3.** Es sei $T$ eine lineare Abbildung auf $\boldsymbol{V}$, d.h. von $\boldsymbol{V}$ nach $\boldsymbol{V}$. Dann heisst ein Vektor $\mathbf{v}_0$ *Eigenvektor* für $T$ wenn gilt:

$$T(\mathbf{v}_0) = \lambda_0 \mathbf{v}_0, \quad \text{für ein } \lambda_0 \in \mathbb{C}. \tag{17}$$

defeigvect

Der Skalar $\lambda_0$ heisst *Eigenwert* von $T$. Allgemein heisst eine Zahl $\lambda$ Eigenwert von $T$ wenn es einen von Null verschiedenen Vektor $\mathbf{v} \neq 0 \in \boldsymbol{V}$ gibt, sodass $T\mathbf{v} = \lambda\mathbf{v}$ gilt.

Man kann zeigen, dass jede lineare Abbildung von $\boldsymbol{V}$ nach $\boldsymbol{V}$ wenigstens *einen* Eigenvektor hat. O.B.d.A. genügt es, den Fall $\boldsymbol{V} = \mathbb{C}^n$ (bzw. $\mathbb{R}^n$) zu betrachten, also das Problem für Matrizen zu behandeln. Ausgangspunkt (siehe das Buch von S. Axler: "Linear Algebra done right"). Zu jedem Vektor $\mathbf{v} \neq 0$ betrachtet man dazu die Vektoren $T^k(\mathbf{v}), k = 0, 1, \ldots, n$. Wenn $\boldsymbol{V}$ $n$-dimensional ist, ist klar, dass diese Menge von $n+1$ Vektoren in $\boldsymbol{V}$ linear abhängig sein müssen. Die weiteren Details werden wir wohl später betrachten (es ist dies ein eher unüblicher Ansatz!).

**Lemma 3.** *Es sei* $\lambda$ *ein Eigenwert von* $T$*, dann bezeichnen wir mit* $E_\lambda$ *den zugehörigen Eigenraum, d.h.*

$$E_\lambda = \{\mathbf{v} \,|\, T\mathbf{v} = \lambda\mathbf{v}\}.$$

*Seine Dimension wird die geometrische Vielfachheit des Eigenwertes* $\lambda$ *genannt (*$\geq 1$ *nach Vorausssetzung!).*

*Proof.* Es genügt, festzustellen, dass die Definition gleichbedeutend damit ist, dass die Abbildung $T - Id$ einen nicht-trivialen Nullraum hat. In der Tat ist folgendes klar:

$$T\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow (T - Id)\,\mathbf{v} = \mathbf{0}.$$

Daraus ergibt sich der Rest, bzw. die Tatsache dass $E_\lambda = Null(T - Id)$, also ein Teilraum von $\boldsymbol{V}$ ist. □

Wir wissen schon, dass im Falle von $2 \times 2$ bzw. $3 \times 3$-Matrizen die Existenz von nichttrivialen Elementen mit dem "Verschwinden der Determinante" von $\boldsymbol{A} - Id_n$ (dem Nullsein derselben) verknüpft ist, also damit, dass diese Matrix *singulär* ist.[3]

[3]Die "guten" Matrizen sind ja bekanntlich die invertierbaren, die eben oft einfach als *nicht-singulär* bezeichnet werden. Der allg. Fall wird in der kommenden Woche noch diskutiert.

Mit Hilfe dieser Terminologie können wir nun folgenden Satz formulieren, davor aber noch kurz folgende Beobachtung:

**Lemma 4.** *Es seien die Vektoren* $\mathbf{v}_1$ *bzw.* $\mathbf{v}_2$ *Eigenvektoren zu zwei verschiedenen Eigenwerten* $\lambda_1$ *bzw.* $\lambda_2$*. Dann sind sie voneinander linear unabhängig.*

Die Aussage für zwei solche Vektoren ist trivial, weil sie ja nur sagt, dass zwei Vektoren, die beide Vielfache voneinander sein können, sich unter $T$ ja nicht unterschiedlich verhalten können! (Details sind der LeserIn überlassen!).

Etwas allgemeiner: es sei $\mathbf{v} \in E_\lambda$ eine Linearkombination von Vektoren, die alle nicht in $E_\lambda$ liegen.

Es ist wohl auch klar (wirklich trivial), dass ein Vektor nicht im Durchschnitt von zwei Eigenräumen liegen kann!!

**Lemma 5.** *Es liege* $\mathbf{v}$ *in* $E_{\lambda_1}$ *und gleichzeitig in* $E_{\lambda_2}$*, für zwei verschiedene Skalare* $\lambda_1 \neq \lambda_2$*. Dann ist* $\mathbf{v} = \mathbf{0}$*, d.h. definitionsgemäß ist dann* $\mathbf{v}$ *kein Eigenvektor.*

*Proof.* Wir haben nach Voraussetzung

$$\lambda_1 \mathbf{v} = T\mathbf{v} = \lambda_2 \mathbf{v} \quad \Rightarrow \quad (\lambda_1 - \lambda_2)\,\mathbf{v} = \mathbf{0},$$

also wegen $\lambda_1 - \lambda_2 \neq 0$ folgt $\mathbf{v} = \mathbf{0}$. □

Eine andere wesentliche Erkenntnis kann man aus den folgenden äquivalenten Umformungen bekommen:

$$\boldsymbol{A} = \boldsymbol{C} * \mathrm{D} * \boldsymbol{C}^{-1} \quad \Leftrightarrow \quad \boldsymbol{A} * \boldsymbol{C} = \boldsymbol{C} * \mathbf{D} \quad \Leftrightarrow \quad \boldsymbol{B} * \mathrm{D} * \boldsymbol{B}^{-1} = \mathrm{D}. \qquad (18)$$ diageigvec00

In Worten, wenn wir zuerst die erste bzw. letzte Formel betrachten: $\boldsymbol{A}$ ist zu einer Diagonalmatrix äquivalent bzw. in einer geeigneten Basis ($\boldsymbol{B}$) kann $\boldsymbol{A}$ durch eine Diagonalmatrix beschrieben werden. Man spricht hier davon, dass $\boldsymbol{A}$ **diagonalisierbar** ist.

Besonders ist die mittlere Gleichung: Erinnern wir uns, was die RECHTE SEITE bedeutet: $\boldsymbol{C} * \mathbf{D}$ ist die Matrix, in deren $k-$ter Spalte die das $d_k$-Fache der $k-$ten Spalte $\mathbf{c}_k$ von $\boldsymbol{C}$ steht. Andererseits steht in der LINKEN SEITE der Gleichung $\boldsymbol{A} * \boldsymbol{C}$, also nach Definition der Matrix-Multiplikation die Folge von Spaltenvektoren $\{\boldsymbol{A} * \mathbf{c}_1, \ldots, \boldsymbol{A} * \mathbf{c}_n\}$. Die MITTLERE Gleichung ist also gleichbedeutend damit, dass gilt $\boldsymbol{A} * \mathbf{c}_k = d_k \mathbf{c}_k$. Da die Spalten von $\boldsymbol{C}$ eine Basis bilden (nach Voraussetzung) sind sie eine Basis von Eigenvektoren zu den Eigenwerten $d_1, \ldots, d_n$.

**Theorem 4.** *Eine Matrix ist genau dann zu einer Diagonalmatrix äquivalent, wenn es eine Basis von Eigenvektoren gibt, oder anders ausgedrückt, wenn es* $n$ *linear unabhängige Eigenvektoren gibt.*

Wir werden uns in der Folge nicht mit dieser ganz allgemeinen Situation befassen, sondern vor allem die Frage nach der Existenz von Orthonormalbasen befassen, d.h. mit der Fragen, wann es eine unitäre Diagonalisierbarkeit gibt. Das hat zwei Gründe: einerseits ist die allgemeine Frage viel schwieriger zu beantworten, es gibt keine *einfachen* Kriterien, und andererseits sind die *wichtigen* Beispiele (nämlich symmetrische reelle Matrizen bzw. hermitsche komplexe Matrizen) durchaus auch unitär diagonalisierbar! Und der Begriff der *Normalität* liefert auch ein Kriterium.

**Definition 4.** Eine $n \times n$-Matrix heißt **normal**, wenn gilt

$$\boldsymbol{A} * \boldsymbol{A}' = \boldsymbol{A}' * \boldsymbol{A}.$$

Frei nach dem Spruch "Wer ist schon normal" sind fast alle (gewürfelten) Matrizen einerseits nicht-singulär, aber anderseits auch nicht-normal.

Zuvor wollen wir aber noch anhand eines einfachen Beispiels der Nicht-Diagonalisierbarkeit klären. Es wird ja nach Einführung der Begriffe nicht verwundern zu erfahren, dass es auch nicht-diagonalisierbare Matrizen gibt. Als einfaches Beispiel kann man die Matrix $\boldsymbol{A}$ nehmen, welche die Abbildung $T(p(t)) = p'(t)$ auf $\boldsymbol{V} = \mathscr{P}_2(\mathbb{R})$ beschreibt (einfach weil das nicht so abstrakt erscheint wie eine einfach hingeschriebene $3 \times 3$-Matrix). Die Matrix hat die Gestalt.

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{19}$$

Es ist auch klar, das $\boldsymbol{A} * \boldsymbol{A} * \boldsymbol{A} = \mathbf{0}$, d.h. die Nullmatrix ist. Ist also $\mathbf{v} \in \mathbb{R}^3$ ein Eigenvektor zum Eigenwert $\lambda$, dann ist es es auch für $\boldsymbol{A}^2 = \boldsymbol{A} * \boldsymbol{A}$ und höhere Potenzen, zum Eigenwert $\lambda^k$, $k = 2, \ldots,$, denn klarerweise gilt das Induktionsargument

$$T^{k+1}(\mathbf{v}) = T^k(T(\mathbf{v})) = T^k(\lambda\mathbf{v}) = \lambda T^k(\mathbf{v}). \tag{20}$$ powereigvec

Wenn also im konkreten Falle gilt, dass $\boldsymbol{A}^3 = \mathbf{0}$ gilt, muss $\lambda^3 = 0$ gelten, d.h. letztendlich $\lambda = 0$.[4]

Wenn wir aber die Frage stellen, wie der Nullraum von $\boldsymbol{A}$ aussieht, oder konkreter die Eigenraum $E_0$, dann ist auch klar, dass dieser eindimensional ist, und aus den Vielfachen des dritten Einheitsvektors besteht. Anders ausgedrückt, eine Polynom hat genau dann Ableitung konstant gleich Null wenn es eine konstantes Polynom (ein Vielfaches des Monomes $t^0 = 1$ ist). Es ist aber auch klar, dass der Bildraum im Falle von $\mathscr{P}_2(\mathbb{R})$ zweidimensional ist (jede linear Funktion von der Form $t \mapsto a_1 t + a_2$ ist die Ableitung von $p(t) = 0.5 a_1 t^2 + a_2 t$. Also ist klar, dass der einzige Eigenraum (um zwei Dimensionen) zu klein ist, und somit keine Diagonalisierbarkeit gegeben sein kann! Durch Wahl von verschiedenen Dimensionen (z.B. $\mathscr{P}_3(\mathbb{R})$ statt $\mathscr{P}_2(\mathbb{R})$, etc.) kann man alle möglichen Beispiel konstruieren, in denen die (direkte) Summe aller Eigenräume alles von $1-$dim. bis $n-$dimensional sind.

Noch zu machen: $q(T)(\mathbf{v}) = q(\lambda)\mathbf{v}$ für Eigenvektoren. >> Invertierbarkeit, Wurzeln.

[4]Vom Skalar $\lambda$ wird keineswegs! vorausgesetzt, dass er nicht Null sein darf. Ganz im Gegenteil, der Eigenraum zu $\lambda = 0$, also $E_0$, war schon lange in Form des Nullraums der Abbildung in Diskussion!