****

Material zur Einführung in die Lineare Algebra und Geometrie

Hans G. Feichtinger (hgfei)

Sommersemester 2014, update Feb.-Juni 2014

Dokumentversion per datum: 9. Juni 2014

http://www.univie.ac.at/NuHAG/FEICOURS/ss14/MainLinAlg14.pdf

Weitere Information unter `www.nuhag.eu/ss14`

Inhaltsverzeichnis

1 Allgemeine Bemerkungen

- Dieses "Skriptum" wird sich von Woche zu Woche verändern. Am Ende des Semesters soll es alle Informationen enthalten, die notwendig sind, um die schriftlichen Abschlussprüfungen erfolgreich zu beenden. Ein eigener (umfassender) Fragenkatalog wird auch noch bereitgestellt. Es lohnt sich daher nicht, das Skriptum laufend auszudrucken (vielleicht von Zeit zu Zeit, oder *stabile* Abschnitte).

- Dies ist eine Massenvorlesung. Daraus ergeben sich ein paar Probleme. Einerseits ist es nicht so einfach, auf individuelle Anfragen während der Stunde einzugehen, andererseits sollte jede/r mitdenken und im Falle von gröberen Unklarheiten (oder Schreibfehlern auf der Tafel, etc.) durchaus den Mut haben, vom Vortragenden anzufragen.

  Ein Skriptum von Andreas Cap ist unter folgender Adresse zu finden:
  `http://www.mat.univie.ac.at/~cap/files/Linalg.pdf`
  Ein Skriptum von Stefan Haller ist unter folgender Adresse zu finden:
  `http://www.mat.univie.ac.at/~stefan/files/LA/LA+LA1+LA2.Skriptum.pdf`

  Wir empfehlen ausdrücklich, eine möglichst regelmässige Anwesenheit in der Vorlesung, und die Erstellung einer eigenen Mitschrift. Einerseits werden oft Skizzen, geometrische Demonstrationen (an der Tafel, oder mit Stäben, mit Händen und Fingern) durchgeführt werden, die nicht in diesem Protokoll zu finden sind, andererseits ist der unmittelbare Lerneffekt während einer Vorlesung (hoffentlich) viel eindrucksvoller und nachhaltiger als die Erarbeitung von mathematischen Inhalten aus einem Buch oder Skriptum. Trotzdem wird stark empfohlen, auch das eine oder andere Buch zu Rate zu ziehen. Bekanntlich ist ein Object besser zu verstehen, wenn man es aus mehreren Blickwinkeln zu sehen bekommen hat. Genauso wird oft ein- und dieselbe Geschichte von verschiedenen Personen ganz unterschiedlich gesehen. Auch der Einsatz von Farbstiften wird in schriftlicher Form nicht nachvollziehbar sein[1].

- Die Literatur zum Thema der Vorlesung ist umfangreich. Wir setzen jedenfalls einmal die im Buch von Schichl/Steinbauer dargelegten Standards ([6]) voraus. D.h. ich kann darauf verzichten, einige grundlegende Definitionen (wie die eines Körpers, einer Gruppe, etc.) zu wiederholen, sondern berufe mich auf diese Quelle. In einigen Fällen werde ich *eigene* Notationen verwenden. Dabei ist einerseits klar, dass man für einen Begriff alle möglichen Symbole verwenden kann (so wie man eine Geschichte in verschiedenen Sprachen erzählen kann), andererseits werde ich mich um eine einheitliche Darstellung bemühen, um nicht durch dauernd welchselnde Notationen Verwirrung zu stiften. Kurz gesagt, die *Unbekannt* muss nicht immer x heissen!

- Ich sehe die Vorlesung als eine *grosse gemeinsame Expedition* an[2] Verschiedene

[1] Es wird auch empfohlen ein paar farbige Stifte zur Anfertigung der Mitschrift dabei zu haben!
[2] in der Vorlesung gibt es dazu noch viele und ausführliche Kommentare.

Teilnehmer erwarten sich verschiedene Dinge davon. Die einen wollen die Welt kennenlernen, die anderen ein (hier geistiges) Abenteuer erleben, die einen wollen einen groben Überblick bekommen um den Gegenstand dann in der Schule zu unterrichten, die anderen wollen Mathematik als Hilfswissenschaft für andere Wissenschaften (z.B. Physik) ausreichend kennenlernen, und wieder andere sind daran interessiert daran, rasch an die vordere Front der Forschung zu kommen, um ihre eigenen Gipfel zu erstürmen (auch wenn das nur eine kleine Minderheit ist).

Ganz gleich wie, werden die kommenden Wochen und Monate eine prägende Wirkung auf Sie alle haben, weil sie die Sichtweise der Mathematik, inbesondere der Linearen Algebra prägen werden.

Beispielsweise bin ich (im Laufe der Jahre) zu einem anwendungsorientierten Mathematiker geworden, und möchte diesen Aspekt der Mathematik sicher nicht mehr vermissen, weil er spannend und auch "rewarding" ist (man hat das Gefühl echte Probleme zu lösen). Andererseits ist es auch immer wichtig, eine entsprechende theoretische Fundierung zu haben, um nicht in den Details verloren zu gehen. Genau das zeichnet nämlich die MathematikerInnen von den reinen Anwendern (die sich mit anderen Details herumschlagen müssen) aus und hilft oft, gerade in komplizierten Situationen den Überblick zu bewahren.

- Jede Vorlesung befindet sich im Spannungsfeld von **axiomatischer** versus **exemplarischer** Darstellungsweise. Im ersten Fall, geht man von den Axiomen des Vektorraums (über einem allgemeinen Körper) aus, um daraus die ganze (universielle) Theorie aufzubauen. Der Begriff *abstrakt* besagt ja nur, dass man von unnötigen konkreten Details absieht, den 3 Äpfel + 5 Äpfel sind 8 Äpfel, ebenso sind 3 Birnen + 5 Birnen eben 8 Birnen, d.h. das Attrribut Apfel oder Birne ist für das Abzählen nicht wichtig, und kann *abstrahiert* werden. Der exemplarische Zugang (den man eher in Büchern aus dem anglo- amerikanischen Raum findet, oder in Büchern zur Matrix-Rechnung) hilft einerseits, die Rechenfertigkeiten zu stärken, andererseits ist es schwer, auf einer solchen Basis etwa Verallgemeinerungen in Richtung Funktional-Analysis (Theorie der normierten Vektorräume, etc., die nicht mehr endlich-dim. sein müssen) zu betreiben.

  Ich werde mich bemühen eine Mischform zu verwenden, d.h. einmal von der einen Seite her, ein andersmal durch illustrative Beispiele, die Sache zu beleuchten. Feedback von Seiten der Hörerschaft, das Zurateziehen von alternativen Quellen (mehr als nur ein Buch) kann sehr hilfen, Begriffe richtig zu erfassen. Ausserdem werde ich mich auch jeweils bemühen, in der Präsentation an der Tafel anzumerken, in welchem Modus der jeweilige Abschnitt zu verstehen ist.

- Abgesehen von dem Wechsel zwischen Axiomatik und Vorführen von Beispiele und praktischen Erläuterungen gibt es noch andere Kategorien. Manchmal ist es wichtig, eine Idee von grossen Zusammenhängen zu geben (dann sind Details natürlich nicht auf Anhieb zu verstehen), ein anderes Mal sollen alle Details vorgeführt werden, um klarzumachen, was/wie eine bestimmte Technik angewandt werden kann.

  Ich vergleiche das auch wieder mit einer Reise: Manchmal verwendet man die Gondelbahn, um rasch auf den Berg zu kommen, ein andermal geht man Schritt für

Schritt den Berg hinauf. Sie als Hörer sollten nicht wie Halbschuhtouristen plötzlich die Kälte der (Hochschul-)Mathematik spüren, ohne dafür gerüstet zu sein. Aber wie im *wirklichen Leben*, kann man sich Schritt für Schritt daran gewöhnen, und eine Art Kondition aufbauen. Das erfordert eine Anstrengung von Seiten der Hörer (die auch Agierende, Rechnende, Fragende,... sein sollten), aber der Erfolg (Verstehen, eigene Fragen finden, etc.) lohnt die Anstrengung.

Wer sich nur in den Rundfahrtsbus setzt, wird die wichtigsten Sehenswürdigkeiten einer Stadt sehen (von aussen), und kann seinen Freunden sagen, dass er dort war, aber wer zu Fuss unterwegs ist, kennt sich wohl besser aus und kann verborgene Winkel entdecken. Zum Mindesten kann sie/er Kondition tanken, und schreckt dann später vor längeren (und spannenderen) Touren nicht zurück!

- Selbst ein Kochbuch (wie etwas das Grosse Sacher Kochbuch) kommt nicht ohne Definitionen aus (dort heißt es: Küchen-Abc, beginnend mit Ananas, bis Zwetschkenröser) und Maße und Gewichte. Bei uns sind das Definitionen und Symbole, die verwendet werdet.

  Aber genausowenig wie Musikwissenschaften die Lehre von den verschiedenen Notenschriften ist, die Sprachwissenschaften sich nicht auf das Transkribieren in Lautschrift oder die Chemie sich nicht auf das Aufzeichnen von chemischen Formeln beschränkt ist, so ist auch der *Gehalt* der Mathematischen Theorien von den *Techniken und Formeln* zu unterscheiden.

- Es gibt eine Liste von Lernzielen. Insbesondere die praktischen Lernziele, die bei den Zwischentest abgefragt werden, werden noch extra bekannt gegeben werden. Nach meiner Erfahrung ist es meistens kein Problem, diese genau vorgegebenen Ziele zu erreichen.

  Hier nur einmal allererste Stichworte: *Vektorräume über Körpern, Linear- Kombinationen, lineare Abbildungen, Matrizen, Vektoren, Matrix-Vektor Multiplikation, Matrix-Produkt, etc..*

- Eine lange Liste von weiteren Büchern (hauptsächlich englisch-sprachig) sind in

  `http://www.univie.ac.at/nuhag-php/books/`

  zu finden. Diese können im Seminarraum 5.131 (neben dem Büro von hgfei, OMP-1) auch [fast jederzeit] eingesehen werden.

- Auch die Vorlesungsunterlagen von Prof. Andreas Cap (vom Vorsemester) können eine sehr gute Quelle sein, siehe sein homepage:

  `http://www.mat.univie.ac.at/~cap/lectnotes.html`

- Literatur ist umfassend, u.a. auch die Bücher von Muthsam, Fischer, etc. ([2, 4]). Weitere Literaturhinweise folgen noch.

- Die Verwendung von MATLAB (im PC-Labor) oder von OCTAVE(frei) wird sehr empfohlen! Von Nicki Holighaus hab ich eine Routine mat2tex (MATLAB to LATEX) die hilft, MATLAB output direkt (in MATH-mode, nicht nur im Verbatim Mode) in LATEX, d.h. z.B. in Übungsaufgaben zu integrieren. Siehe Web-Seite des Vortragenden:

  `http://www.univie.ac.at/NuHAG/FEICOURS/ws1011/ws1011.htm`

- Das Aufspüren und Mitteilen von Ungenauigkeiten im Skriptum (per Email) ist erwünscht und kann mit *Bonuspunkten* honoriert werden.

2 Literatur

Das empfohlene Lehrbuch (nach dem aber nicht streng vorgegangen wird) ist das Buch von Gilbert Strang ([8]).

Stoff zum Kolloqium sind (mit kleinen Abweichungen) die Kapitel **1 - 4** sowie **7** aus dem Buch von G. Strang.

Dem AUFBAU nach ist die Vorlesung stark an das Buch [7] angelehnt: Elementary LINEAR ALGEBRA: A Matrix Approach 2e, by Spence, Insel and Friedberg. Im dortigen Buch entspricht der Stoff der aktuellen Vorlesung (SS14) den Kapiteln 1,2,4,6 (teilweise) sowie 7.

Exemplare der Bücher [8] bzw. [7] können in der NuHAG Hand-Bibliothek (Raum 5.131) (fast jederzeit) eingesehen bzw. studiert werden (Zimmer neben dem Büro von hgfei, es gibt Mehrfachexemplare).

2.1 References, Books

Eine große Liste von Büchern zur linearen Algebra finden Sie unter

`http://www.univie.ac.at/nuhag-php/bibtex/load_q.php?id=91092`

Fast alle dieser Bücher (die mit dem Buchsymbol gekennzeichneten) finden sich als Teil der NuHAG Bibliothek im Raum 5.131, neben dem Büro von HGFei.

Literatur

an98 [1] H. Anton. *Lineare Algebra. Einführung, Grundlagen, Übungen. Aus dem Amerikanischen von Anke Walz. (Linear algebra. Introduction, foundations, exercises. Transl. from the American by Anke Walz).* Spektrum Akademischer Verlag, Heidelberg, 1998.

fi08 [2] G. Fischer. *Linear algebra. An introduction for beginners. (Lineare Algebra. Eine Einführung für Studienanfänger.) 16th revised and enlarged ed.* Vieweg Studium: Grundkurs Mathematik. Wiesbaden: Vieweg. xxii, 384 p. EUR 19.90, 2008.

fu12 [3] P. Fuhrmann. *A Polynomial Approach to Linear algebra 2nd Ed.* Universitext. New York, NY: Springer. xvi, 2012.

mu06 [4] H. J. Muthsam. *Lineare Algebra und ihre Anwendungen.* Spektrum Akademischer Verlag, Heidelberg, 2006.

olsh06 [5] P. Olver and C. Shakiban. *Applied Linear Algebra.* Pearson Education Inc., 2006.

scst09 [6] H. Schichl and R. Steinbauer. *Introduction to Working Mathematically (Einführung in das Mathematische Arbeiten).* Springer Berlin, 2009.

frinsp00 [7] L. Spence, A. Insel, and S. Friedberg. *Elementary Linear Algebra: A Matrix Approach.* Prentice Hall, 2000.

st03-8 [8] G. Strang. *Lineare Algebra.* Springer-Lehrbuch. Springer, Berlin, 2003.

3 Abkürzungen, MATLAB/OCTAVE Notationen

3.1 Abkürzungen

Abkürzung	ausgeschrieben
o.B.d.A	ohne Beschränkung der Allgemeinheit
l.S.	linke Seite (der Gleichung)
r.S.	rechte Seite (der Gleichung)
":= "	Definitionszeichen: r.S. erklärt l.S.
w.z.z.w	was zu zeigen war
q.e.d.	quod erat demonstrandum
Lin.Komb.	Linear Kombination
Matr. Mult.	Matrix Multiplikation
Wo.	Woche
$\Rightarrow$	impliziert (hat zur Folge)
$\Leftrightarrow$	dann und nur dann
$\exists!$	es existiert genau ein
$\forall$	für alle (... gilt)
ZStf. oder (R)ZSF	(red.) Zeilenstufenform (der Matrix $\boldsymbol{A}$);
allg. Lsg.	allgemeine Lösung
lin. GLS	lineares Gleichungssystem
spez. Lsg.	spezielle Lösung eines lin.GLS
f.d.	für die *oder* für den
lin.Abb.	lineare Abbildung
sog.	sogenannte/r/s
spez.	speziell/er/es/e
Pol.Fkt.	Polynom-Funktionen
ONB	Orthonormal Basis
$\delta_{0,1}$	Kronecker Delta
ZSF	Zeilenstufenform (Treppennormalform)
i.a.	im allgemeinen

3.2 Schreibkonventionen

$\vec{s} = [1; 2; 3; 4]$; bzw. $\vec{s} = [1, 2, 3, 4]$; beschreibt die Eingabe von *Spaltenvektoren* resp. *Zeilenvektoren* in OCTAVE (bzw. MATLAB).

$\boldsymbol{A} = [\vec{\mathbf{a}}, \cdots, \vec{\mathbf{a}}_n]$ or $\boldsymbol{B} = [\vec{\mathbf{b}}, \cdots, \vec{\mathbf{b}}_m]$ beschreibt eine $m \times n$-Matrix $\boldsymbol{A}$ (resp. $m \times p$-Matrix $\boldsymbol{B}$) als Kollektion (geordnete Familie, man könnte sagen nummerierte) von Spaltenvektoren im $\mathbb{K}^m$ (resp. $\mathbb{K}^p$).

3.3 OCTAVE bzw. MATLAB Rechenbefehle

Befehl	Funktionsweise
$\boldsymbol{A} * \mathbf{x}$	$\boldsymbol{A}\mathbf{x}$: Matrix-Vektor Multiplikation
$\boldsymbol{A} * \boldsymbol{B}$	Matrix-Matrix Multiplikation
$\boldsymbol{A}'$	$conj(\boldsymbol{A}^t)$: transponiert und konjugierte Matrix
$\boldsymbol{A}.'$	$\boldsymbol{A}^t$: transponierte Matrix
$\det(\boldsymbol{A})$	Determinante von $\boldsymbol{A}$
$\mathrm{inv}(\boldsymbol{A})$	Inverse Matrix zu $\boldsymbol{A}$;
$\boldsymbol{A}(2,:)$	zweite Zeile von $\boldsymbol{A}$;
$\boldsymbol{A}(:,4);$	vierte Spalte von $\boldsymbol{A}$;
$\mathrm{rref}(\boldsymbol{A})$	row reduced echelon Form (Zeilen-Stufen-Form);
bzw.	beziehungsweise (respektive...)
etc.	und so weiter
z.B.	zum Beispiel (oder: beispielsweise)

3.4 MATLAB bzw. OCTAVE und Matrizen

Der Befehl ";" eröffnet eine neue Zeile, bzw. kann zum trennen von Zeilenvektoren (siehe Def. von A_1 unten) verwendet werden, das Komma "," trennt Eintragungen, die nebeneinander stehen.

```
>> A = [1,3; 2,4]
%  A = [1,3; 2,4];  wuerde Ausgabe unterdruecken!
A =
     1     3
     2     4
>> A = [1 3; 2 4]
A =
     1     3
     2     4
>> a1 = A(:,1)    % Entnehmen der ersten Spalte
a1 =
     1
     2
>> a2 = A(:,2);
>> z1 = A(1,:)
z1 =
     1     3
>> z2 = A(2,:);
>> A1 = [z1; z2]   % Untereinanderstellen von Zeilenvektoren
A1 =
     1     3
     2     4
>> A2 = [a1,a2]   % Nebeneinanderstellen von Spaltenvektoren
A2 =
     1     3
```

```
     2     4

oder auch
>> A = zeros(2); A(:) = 1:4;
% Befuellen einer 2x2 Nullmatrix mit der Folge 1,2,3,4
>> >> v = A(:)
v =
     1
     2
     3
     4
ist die dazu inverse Operation (Auslesen in der "natuerlichen"
 Reihenfolge, erste bis letzte Spalte, von oben nach unten.
```

4 Konventionen and Sprachregelungen

Wir werden von **Vektoren** reden, und zwar einerseits konkrete (Zeilenvektoren, Spaltenvektoren), andererseits von (abstrakten) Vektoren. Das Wort "Vektor" dient dann einfach dazu, um Elemente von Vektorräumen anzusprechen[3], auch wenn sie (sozusagen bei näherer Betrachtung auch Funktionen sind, oder zwei Matrizen, etc., mit noch mehr Struktur).

Es ist angebracht, Vektoren und ihre Koordinaten mit ähnlichen Symbolen zu versehen. Wir verwenden z.B. (in gedruckter Form) boldface Symbole, wie $\mathbf{x}, \mathbf{y}, \mathbf{z}$ für (gewöhnliche) Vektoren im $\mathbb{R}^m$ (oder $\mathbb{R}^k$ oder $\mathbb{R}^n$). Auf der Tafel ist es hingegen oft besser leserlich wenn wir Symbole wie $\vec{x}$ oder $\vec{\mathbf{x}}$ verwenden. Es ist dann naheliegend, die einzelnen Koordinaten gleichlautend zu verwenden, also als $x_1, \cdots x_n$ zu bezeichnen (nun nicht mehr bold-face). Im drei-dim. Euklidischen verwendet man gerne (x, y, z) anstelle von (x_1, x_2, x_3), aber das wird sich aus dem Kontext leicht erschliessen lassen. Anwender (z.B. Nachrichtentechniker) verwenden oft das Symbol $x[3]$ anstelle von x_3, und in MATLAB (dort kann man kein fettgedruckten Namen verwenden, es unterscheidet nur nach Gross- und Kleinschreibung der Variablen-Namen) ruft man die dritte Koordinate von x mit $x(3)$ auf.

Für Matrizen verwenden wir normalerweise Grossbuchstaben, also etwa $\boldsymbol{A}, \boldsymbol{B}$ etc.. Die entsprechenden Koordinaten sind dann oft mit Kleinbuchstaben versehen, d.h. man schreibt oft $\boldsymbol{A} = (a_{j,k})_{j,k}$, oder $\boldsymbol{A} = (a_{j,k})_{1\le j\le m, 1\le k\le n}$, wobei **Z**uerst der **Z**eilenindex j und **Sp**äter der **Sp**altenindex k angegeben wird. In MATLAB hat man z.B. $AA = rand(3)$ (ergibt eine 3×3 -Matrix mit zufälligen Eintragungen aus $[0, 1]$). Dann ist $AA(2, 3)$ das, was man in der üblichen Sprechweise als $a_{2,3}$ bezeichnen wuerde.[4]

Die Elemente mit festen Index z.B. $j = 2$ bilden die zweite Zeile, in MATLAB $z2 = \boldsymbol{A}(2, :)$;. Anderseits sind die Eintragungen mit festem Spaltenindex $k = 3$ (wieder nur beispielsweise) die dritte Spalte bilden, in MATLAB $a3 = \boldsymbol{A}(:, 3)$. Das Symbol ":" steht also für "irgendwas". Es gibt sogar den Befehl $\mathbf{v} = \boldsymbol{A}(:)$, der einfach die Eintragungen von $\boldsymbol{A}$ in Vektorform (als Spaltenvektor) ausgeben. Dabei steckt die *Konvention* dahinter, dass eine $m \times n$-Matrix einfach als eine Kollektion von Spaltenvektoren gesehen wird, mit der ersten Spalte (ganz links) beginnend!

Es gibt auch die "umgekehrte Operation", beispielsweise: $Z = zeros(2, 3)$, $Z(:) = 1 : 6$; ergibt die Matrix

$$\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} \tag{1}$$

[3] So wie eine *Menge* aus *Elementen* besteht, und wenn die Menge aus Personen besteht, sind die "Elemente" eben Menschen ...

[4] Man kann diese Eintragung auch als $AA(8)$ abfragen, indem man die Matrix spaltenweise ausliest, von oben nach unten, beginnend mit der ersten Spalte!

5 Einstiegsbeispiel, Matrizen, Gauss-Elimination

Motivation: Normalerweise beginnt man die Lineare Algebra entweder vom *abstrakten* Ende (d.h. mit der Definition eines Vektorraumes über einem Körper) und studiert dann die Eigenschaften von linearen Abbildungen. Die scheinbar *konkrete* Methode beginnt gleich mit Vektoren und Matrizen und vermittelt zu allererst die Techniken der Matrizen-Rechnung.

Ich wähle einen *anwendungsorientierten* Einstieg, d.h. weder das eine noch das andere steht am Anfang, sondern etwas "in der Mitte". In der Tat ist es so, dass die (numerische) Lineare Algebra in den Ingenieurswissenschaften (insbes. Nachrichtentechnik) eine immer größere Rolle spielt. Sie treten dort aber nur sehr selten direkt als abstrakt mathematische Fragen auf.[5]

Weder die *Vektorraumstruktur* noch die *Linearität* sind sofort als wichtige Struktur-Merkmale des Problems zu erkennen, allerdings stellt sich bei näherer Betrachtungsweise schnell heraus, dass eine abstrakte Sichtweise hilft, den Überblick zu gewinnen und systematisch Methoden (der linearen Algebra) zu entwickeln, die universell (für eine große Klasse ähnlicher Probleme) anwendbar sind.

Ich möchte daher (auch um ein paar Fakten aus der komplexen Analysis kurz in Erinnerung zu rufen) von einem Problem betreffend komplex-wertige Funktionen auf der Zahlengeraden, also für bestimmte Funktionen $f : \mathbb{R} \to \mathbb{C}$ starten.

Bekanntlich kann man für rein imaginäres z, d.h. für $z = it$, mit $t \in \mathbb{R}$, den Zusammenhang zwischen der Exponentialfunktion und den Winkelfunktionen durch die wichtige Gleichung $e^{it} = cos(t) + i\, sin(t)$

$$e^{it} = cos(t) + i\, sin(t), \quad t \in \mathbb{R}, \tag{2}$$ Euler1

beschreiben. Ersetzt man in dieser Gleichung t durch $-t$ erhält man (?warum?)

$$e^{-it} = cos(t) - i\, sin(t), \quad t \in \mathbb{R}, \tag{3}$$ Euler2

Diese beiden Gleichungen (2) und (3) zusammen sind in Wirklichkeit äquivalent zu einem anderen Paar von Gleichungen, die die Winkelfunktionen durch die Exponential-Funktion ausdrücken. Zu diesem Zweck braucht man nur die Summe bzw. die Differenz der beiden Gleichungen bilden. Dividiert man dann jeweils durch den Faktor (2 neben dem Term $cos(t)$ bzw. $2i$ neben dem Term $sin(t)$) so bekommt man

$$cos(t) = \frac{e^{it} + e^{-it}}{2}; \quad sin(t) = \frac{e^{it} - e^{-it}}{2i} = \frac{-i}{2}(e^{it} - e^{-it}); \quad t \in \mathbb{R}, \tag{4}$$ Euler3

Natürlich kommt man, einfach durch Einsetzen, aus dem Gleichungspaar (4) sofort (wieder auf die Ausgangsgleichung zurück) (5) zurück, die auch als

$$cos(t) = \text{real}(e^{it}), \quad sin(t) = \text{imag}(e^{it}), \quad t \in \mathbb{R}, \tag{5}$$ Euler00

beschrieben werden kann (Real- bzw. Imaginärteil von $z \in \mathbb{C}$ sind reelle! Zahlen!).

[5] obwohl auch Vektorräume üeber endlichen Körpern durchaus eine wichtige Rolle in der Kryptographie spielen!

Bekanntlich gilt (siehe Analysis Kurs) für die Exponentialgleichung (auch im Komplexen) das Exponentialgesetz, d.h. die Formel (hier ohne Beweis)

$$e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}\, z_1, z_2 \in \mathbb{R}. \tag{6}$$

exp-gesetz

Daraus lassen sich dann leicht Additionstheorem für die Winkel-Funktionen herleiten (siehe *De Moivre'sche Formeln*), durch Bildung von Real- bzw. Imaginärteil, siehe beispielsweise

`http://de.wikipedia.org/wiki/Moivrescher_Satz`

$$\cos(nt) = \text{real}(e^{int}) = \text{real}((e^{it})^n) = \text{real}[\cos(t) + i\sin(t)]^n \tag{7}$$

DeMoivre1

in Verbindung mit dem binomischen Lehrsatz: $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

Eine bereits abstrakte Sichtweise der bisherigen Rechnungen wäre die Beobachtung, dass (abgesehen von den speziellen Eigenschaften der Funktionen (entweder der Winkelfunktonen oder der Exponentialfunktion) nur folgender Zusammenhang eine Rolle spielt:
Ausgehend von den Funktionen $f_1(t) = \cos(t)$ und $f_2(t) = i \cdot \sin(t)$ (beides komplexwertige Funktionen auf $\mathbb{R}$) haben wir neue Funktionen gebildet:

$$g_1(t) = f_1(t) + f_2(t); g_2(t) = f_1(t) - f_2(t);$$

Geht man genauso vor wie bisher (Addieren bzw. Subtrahieren der beiden Gleichungen voneinander) so kommt man zu der "umgekehrten Relation":

$$f_1(t) = \frac{1}{2}(g_1(t) + g_2(t))\,,\; f_2(t) = \frac{1}{2}(g_1(t) - g_2(t))$$

oder (dasselbe ein wenig anders aufgeschrieben)

$$f_1(t) = 1/2 \cdot g_1(t) + 1/2 \cdot g_2(t)$$

$$f_1(t) = 1/2 \cdot g_1(t) - 1/2 \cdot g_2(t)$$

Gar nicht viel anders verhält es sich, wenn wir aus den beiden Einheitsvektore $\mathbf{e}_1 = [0;1]; \mathbf{e}_2 = [1,0]$ ein neues Koordinatensystem bilden, und zwar

$$\vec{\mathbf{u}}_1 = \vec{\mathbf{e}}_1 + \vec{\mathbf{e}}_2\;; \quad \vec{\mathbf{u}}_2 = \vec{\mathbf{e}}_1 - \vec{\mathbf{e}}_2.$$

Auch in diesem Fall kann man die alten (Standard-) Einheitsvektoren wieder im neuen Koordinatensystem ausdrücken, und zwar als

$$\vec{\mathbf{e}}_1 = (\vec{\mathbf{u}}_1 + \vec{\mathbf{u}}_2)/2\,,\; \vec{\mathbf{e}}_2 = (\vec{\mathbf{u}}_1 - \vec{\mathbf{u}}_2)/2.$$

Daraus folgt sogar, dass ein Vektor $\vec{\mathbf{x}} = x_1\vec{\mathbf{e}}_1 + x_2\vec{\mathbf{e}}_2$ sich natürlich auch in dem $\mathbf{u}$-Koordinatensystem in der Form

$$\vec{\mathbf{x}} = c_1\vec{\mathbf{u}}_1 + c_2\vec{\mathbf{u}}_2$$

darstellen läßt, und zwar durch .. (Excercise!).

Wir können diese Aktionen auch in folgender Weise darstellen, unter Benutzung der Programmiersprache MATLAB^{TM}.

Anmerkung: es wird sehr empfohlen, wenigstens eine alte Version von MATLAB oder OCTAVE zu installieren, weil das praktische Experimentieren mit solchen Programmen sicherlich hilft, den Zusammenhang zwischen abstrakten Prinzipien und numerischer, d.h. rechnerischer Umsetzung herzustellen.

Konkret kann man sagen, dass das, was wir bisher gemacht haben, sich in der Form der *zugehörigen Matrizen*[6] wie folgt ausdrückt.

```
Gemeinsamkeiten dieser Rechnungen (in Farbe!)
A = [1,1; 1,-1]    % Eingabe der Matrix A,  ";"  bedeutet Zeilenende
A =
     1     1
     1    -1
Neue Linearkombination der Zeilen: Summe bzw. Differenz der Zeile von A: B
>> B = [A(1,:)+A(2,:); A(1,:) - A(2,:)]
B =
     2     0
     0     2
>> C = B/2        %  Multiplizieren beider Zeilen von B mit 1/2 > C
C =
     1     0
     0     1
Alternative: (andere Zeilenoperationen, laut Gauss Elimination)
B = [A(1,:); A(2,:) - A(1,:)]    % subtrahiere erste Zeile von der 2-ten
B =
     1     1
     0    -2
>> B(2,:) = B(2,:)/-2      % multipliziere die zweite Zeile mit -1/2
B =
     1     1
     0     1
>> B(1,:) = B(1,:) - B(2,:)  % subtrahiere die zweite von der ersten Zeile
B =
     1     0
     0     1
```

[6]Dabei sei hier zunächst dieser Begriff nur zu verstanden, dass man die bekannten Zahlen in wohlorganisierter Form, in einem quadratische Schema aufschreibt, und nur aus den vorgegebenen Gleichunge sukzessive neue Gleichungen aufbaut, die ebenfalls richtig/gültig sein müssen, wenn die ursprünglichen Gleichungen gelten sollen. Wir leiten also - strenggenommen - nur notwendige Bedingungen ab, die sich aber in den bisherigen Fällen als eine äquivalente Umformung erweisen, weil man ja wieder zu den Ausgangsgleichungen zurückkommt

Dabei sind nur einige wenige Prinzipien verwendet worden: etwa (im ursprünglichen System): Wenn $h_1(t) = h_2(t)$ und $s_1(t) = s_2(t)$ dann gilt auch

$$h_1(t) + s_1(t) = h_2(t) + s_2(t)$$

sowie $\lambda h_1(t) = \lambda h_2(t)$ (für alle $\lambda \in \mathbb{C}$, falls $\lambda \neq 0$ gilt auch die Umkehrung). Kurz gesagt, wir benötigen nur eine *Vektorraumstruktur* !! (auf der Menge der komplex-wertigen stetigen Funktionen auf $\mathbb{R}$).

5.1 Matrizen als kompakte Darstellung

Ausgehend von diesen Beispiele kann man Matrizen als rechteckige Schemata von Körperelementen definieren (s. Einführungsbuch). Jedes Element *muss* besetzt sein, d.h. es gibt keine Matrix mit Löchern[7].

Meistens (beispielsweise in MATLAB) werden nicht bekannt Felder mit der Zahl 0 initialisiert (vordefiniert), aber das ist manchmal ganz schön gefährlich, weil die Gefahr besteht, dass ein 0 für NEIN steht, und 1 für ja (sagen wir in einer Umfrage), und natürlich ist ein NEIN nicht dasselbe wie eine nicht gegebene Antwort[8].

MATLAB Initialisierung: $Z = zeros(4)$ oder $V = zeros(2, 3)$ erzeugt eine entsprechende Null-Matrix vom Format 4×4 bzw. 2×3.

So gesehen ist das Gleichungssystem oben in Matrix Betrachtung so aufzuschreiben:

$$\begin{aligned} x_1 + x_2 &= b_1 \quad (8) \\ x_1 - x_2 &= b_2 \quad (9) \end{aligned}$$

oder aber auch (um klarzumachen dass vor den Zahlen der *Faktor* 1 steht!):

$$\begin{aligned} 1 \cdot x_1 + 1 \cdot x_2 &= b_1 \quad (10) \\ 1 \cdot x_1 - 1 \cdot x_2 &= b_2 \quad (11) \end{aligned}$$

Daher ist es naheliegend, dieses Gleichungssytem in drei Teile zu zerlegen, die *Variablen* x_1, x_2, die *Systemmatrix* $\boldsymbol{A}$ und die rechte Seite $\vec{\mathbf{b}} = [b_1, b_2]$, mit

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (12)$$

Der *VORTEIL* dieser Darstellung ist, dass wir alle Rechnungen dann nur mit der System-Matrix machen können. Zieht man die rechte Seite noch mit ein, so kommt man zur

[7] Wenn man aber das Matrizen-Schema wie ein Excel-File zur Speicherung von Daten benützt, oder zur Speicherung von Roh-Bild-Information, wie dies in einer Digital-Kamera der Fall ist, muss auch der Wert **NaN**: Not a Number, erlaubt sein, um nicht laufend Fehlermeldungen, z.B. beim Display etc. zu produzieren; solche Konventionen sind praktisch wichtig und sinnvoll, aber nicht innerhalb der streng mathematischen Sichtweise. Details dazu wohl noch später!

[8] Nichtwähler bilden sozusagen eigene eigene Partei, nicht zu wählen ist eine Entscheidung für die Partei der Nichtwähler, wenn man z.B. Wählerstrom- Analysen machen will!

erweiterten System-Matrix, die ich (im Gegensatz zu Gilbert Strang, der sie mit $\boldsymbol{A}'$ bezeichnet das ist aber ganz schön gefährlich, weil dieses Symbol in MATLAB die Funktion des Transponierens einer reellen Matrix realisiert!) mit $\tilde{\boldsymbol{A}}$ bezeichnen möchte:

$$\tilde{\boldsymbol{A}} = \begin{pmatrix} 1 & 1 & b_1 \\ 1 & -1 & b_2 \end{pmatrix} \tag{13}$$

Um von dieser Art der "kompakten Schreibweise" praktikabel zu machen, **definiert** man ein Produkt zwischen der System-Matrix $\boldsymbol{A}$ und dem Vektor $\vec{\mathbf{x}} = [x_1; x_2]$ eben genau so, dass man als $\boldsymbol{A} * \vec{\mathbf{x}}$ genau die linke Seite des Gleichungssytems bekommt.

Wir kommen also zu folgender, allgemeinen Definition eines *Matrix-Vektor Produktes*:

matr-vec **Definition 1.** Es sei $\boldsymbol{A}$ eine reelle $m \times n$-Matrix, und $\vec{\mathbf{x}} \in \mathbb{R}^n$ ein Spaltenvektor, dann ist das Matrix-Vektor-Produkt, welches als $\mathbf{y} = \boldsymbol{A}\mathbf{x}$ oder wie in MATLAB $\mathbf{y} = \boldsymbol{A} * \mathbf{x}$ geschrieben wird, der Spaltenvektor $\vec{\mathbf{y}} = [y_1; \cdots; y_m] \in \mathbb{R}^m$, mit[9]

$$y_r = \sum_{k=1}^{n} a_{r,k} x_k = a_{r,1} x_1 + \cdots a_{r,n} x_n. \tag{14}$$ matvdef

Anmerkung: Man "realisiert das" (händisch) indem man den k-ten Zeilen-Vektor mit dem linken Zeilenvektor entlangfährt, während gleichzeitig der Zeigefinger der rechten Hand den Spaltenvektor entlangfährt, und entsprechende Eintragungen zuerst ausmultipliziert und dann aufsummiert werden: siehe Vorlesung, und !Übungen!

Man sieht sofort, dass man durch diese Multiplikation eine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ definiert hat[10]. In der obigen Definition ist daher auch der Fall $m = 1$ erlaubt, der eine spezielle Bedeutung hat(und noch ausführlich diskutiert werden wird).

alar-Produkt **Definition 2.** Das *Skalarprodukt von zwei Vektoren* $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^n$ ist gegeben durch

$$\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle := x_1 y_1 + \cdots x_n y_n \sum_{k=1}^{n} x_k y_k.$$

Zwei Vektoren heißen zueinander **orthogonal** (oder *stehen senkrecht aufeinander*), wenn:

$$\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle = 0 = \quad \text{(und dann gilt ja auch} \quad \langle \vec{\mathbf{y}}, \vec{\mathbf{x}} \rangle = \overline{(0)} = 0!).$$

Es benötigt in MATLAB keinen eigenen Befehl sonder ist einfach mit Hilfe des **Transpositions-Befehls** und der Matrix-Vektor Multiplikation realisiert: $spxy = \vec{\mathbf{y}}' * \vec{\mathbf{x}}$, wobei man in der math. Literatur für $\vec{\mathbf{y}}'$ einfach $\vec{\mathbf{y}}^t$ schreibt ("t" wie Transponieren, oder "*transpose*" im Englischen)[11].

[9] Der Strichpunkt bedeutet in MATLAB, dass man einen Spaltenvektor bildet. Mit einfachem Beistrich würde man einen Zeilenvektor beschreiben.

[10] Manchmal wird als störend empfunden, dass eine $m \times n$-Matrix von $\mathbb{R}^n$ nach $\mathbb{R}^m$ führt, wir werden aber sehen, dass das logisch sinnvoll ist, und andererseits damit zu tun hat, dass man die Matrix-Multiplikation von links schreibt; im Prinzip könnte man auch mit Rechts-Matrix-Multiplikation arbeiten! Das Transponieren führt das eine ins andere über.

[11] Unsere Schreibweise ist direkt mit MATLAB/OCTAVE verträglich.

euklnorm **Definition 3.** Für $\vec{\mathbf{x}} \in \mathbb{C}^n$ wird die **Euklidische Länge** bzw. **Norm** (in der Schule: **Absolutbetrag** des Vektors) definiert durch

$$\|\vec{\mathbf{x}}\| :=^{+}\sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}.$$

Die Transposition ist auch für Matrizen anwendbar, und macht aus einer $m \times n$-Matrix eine $n \times m$-Matrix, indem man einfach Zeilen und Spalten in ihren Rollen vertauscht. Anders ausgedrückt: die Zuweisung $\boldsymbol{B} = \boldsymbol{A}'$ macht aus der ersten Zeile von $\boldsymbol{A}$ die erste Spalte von $\boldsymbol{B}$ etc., oder (alternative betrachtet) aus der ersten Spalte von $\boldsymbol{A}$ die erste Zeile von $\boldsymbol{B}$, oder formaler ausgedrückt:

$$b_{k,j} := a_{j,k}, \quad \text{mit} \quad 1 \leq k \leq m, 1 \leq j \leq n. \tag{15}$$

transpmatr

Um gleich für die entsprechenen Begriffe im Falle komplexer Vektoren vorbereitet zu sein, wollen wir gleich anmerken, dass die richtige Verallgemeinerung von $\mathbf{y}$ zu $\mathbf{y}'$ der Übergang von einem Zeilen- oder Spaltenvektor zum *konjugierten* (!) Spalten bzw. Zeilenvektor bedeutet. Das ist auch sinnvoll, wie wir später sehen werden. Insbesondere haben wir (das betrifft nun nur das Skalar-Produkt, aber NICHT die Matrix- Multiplikation!!!)

lar-ProduktC **Definition 4.** Das **Skalarprodukt von zwei Vektoren** $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{C}^n$ ist gegeben durch

$$\langle \vec{\mathbf{x}}, \vec{\mathbf{y}}\rangle := x_1\overline{y_1} + \cdots x_n\overline{y_n} = \sum_{k=1}^{n} x_k\overline{y_k}.$$

Eine für das spätere wichtige Konsequenz dieser Definition ist die Tatsache, dass $\langle \vec{\mathbf{z}}, \vec{\mathbf{z}}\rangle$ immer nichtnegativ ist, ja sogar dass gilt: $\langle \vec{\mathbf{z}}, \vec{\mathbf{z}}\rangle > 0$ falls $\vec{\mathbf{z}} \in \mathbb{C}^n, \vec{\mathbf{z}} \neq \vec{\mathbf{0}}$, weil ja gilt

$$\Leftrightarrow \quad \langle \vec{\mathbf{z}}, \vec{\mathbf{z}}\rangle = \sum_{k=1}^{n} |z_k|^2 > 0 \quad \Leftrightarrow \quad z_k \neq 0 \text{ für wenigstens ein } k, 1 \leq k \leq n. \tag{16}$$

skalpos0

Übungsaufgabe: Man zeige, dass gilt

$$\forall \vec{\mathbf{x}} \in \mathbb{C}^n, \lambda \in \mathbb{C}: \quad \|\lambda \cdot \vec{\mathbf{x}}\| = |\lambda| \cdot \|\vec{\mathbf{x}}\|. \tag{17}$$

normhomog1

Da die Konstruktion des Skalarproduktes aber nicht für allgemeine Körper (sondern nur für $\mathbb{R}$ und $\mathbb{C}$) brauchbar ist wollen wir vorerst diese Linie nicht so stark weiterverfolgen, sondern lieber einige anderen Beispiele von abstrakten und konkreten Vektorräumen, von Linearkombinationen und Konsequenzen behandeln.

Einschub (SS2014) Es ist vielleicht wichtig, darauf hinzuweisen, dass es (zu Recht!) in MATLAB keinen eigen Typ von Vektoren gibt, und keinen Befehl für das Bilden von Skalar-Produkten. Vielmehr sind die in MATLAB vorhandenen Begriffe einzig und alleine Matrizen sowie das Symbol $*$ zur Bildung des Matrix-Produktes. Weiters gibt es in MATLAB (bzw. OCTAVE) den Befehl $\boldsymbol{A} \mapsto \boldsymbol{A}'$ (die konjugierte, transponierte Matrix). Wenn man also das Skalarprodukt von zwei Spaltenvektoren $\vec{\mathbf{x}}$ und $\vec{\mathbf{y}}$ bilden will, also den mathematischen Ausdruck $\langle \vec{\mathbf{x}}, \vec{\mathbf{y}}\rangle$, dann kann/muss man das durch Berechnung des Ausdruckes $\vec{\mathbf{y}}' * \vec{\mathbf{x}}$ machen. Auf diese Weise ist die Konjugation (die im Skalarprodukt

nach math.(!) Konvention, bei den Physikern ist es umgekehrt!) auf den Komponenten von *vecby*, allerdings muss auf diese Weise die Reihenfolge umgestellt werden.

Als Konsequenz kann man den Test, ob ein Kollektion von Spaltenvektoren ein Orthonormalsystem darstellt auf die Überprüfung der sogenannten Gram-Matrix zurückführen, d.h. die Betrachtung der Matrix $\mathbf{U}' * \mathbf{U}$ (wenn die Spalten der Matrix $\mathbf{U}$ die gegebenen Vektoren $\mathbf{u}_1, \cdots, \mathbf{u}_n$ sind). Genau dann wenn diese Gram-Matrix die Einheitsmatrix ist, sind bilden die Spalten ein System von Vektoren der Länge 1 (normiert), welche paarweise zueinander senkrecht stehen. Zur Kontrolle: Wenn $\mathcal{G} := \mathbf{U}' * \mathbf{U}$ ist, dann is $g_{2,3}$ das Skalarprodukt von zweiter Zeile von $\mathbf{U}'$ (also zweiter Spalte von $\mathbf{U}$) mit der dritten Spalte von $\mathbf{U}$, aber wenn/weil im Komplexen auf die Konjugation zu achten ist, die sich auf den Vektor $\mathbf{u}_2$ bezieht, müssen/sollten wir es als $\langle \mathbf{u}_3, \mathbf{u}_2 \rangle$. interpretieren.

Insbesondere ist unser erstes Ziel, das *Arbeitspferd* der Linearen Algebra, das sogenannte *Gauss'schen Eliminationsverfahren* bzw. den Gauss-Jordan Algorithmus zu beschreiben und in Matrix- Formulierung zu beschreiben. Das ergibt auch umfassendes Material für die Übungen (incl. Proseminartest!).

5.2 Das Gauss'sche Eliminationsverfahren

In Matrix Form ist das Ziel, **aus einer** *(erweiterten) System-Matrix* **durch systematische Umformen in die (reduzierte) Zeilenstufenform zu kommen.** Dabei sind die folgenden drei Operationen erlaubt:

1. **Multipliziere eine Zeile mit einer Zahl ungleich Null;**
2. **Addiere (das Vielfache) eine(r) Zeile zu einer anderen Zeile**[12]**;**
3. **Wenn nötig, können auch verschiedene Zeilen vertauscht werden.**

Das sind sehr einfache und schnell zu realisierende Prozeduren, die aus einer Matrix schrittweise eine "bessere" machen, aber was ist das (perfekte?) Ziel? Betrachten wir ein *trt*-Gleichungssystem, mit den Unbekannten/Variablen (x, y, z). Dann wäre eine modifizierte Systemmatrix von der Form einer sogenannten *Einheitsmatrix* bzw. einer *Diagonalmatrix* mit lauter Einsen auf der *Hauptdiagonale*, d.h. konkret

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{18}$$

ideal, weil das entsprecht dann $[x, y, z] = [b_1, b_2, b_3]$, d.h. die Lösung hat man dann sofort vor Augen.

Ein Alternative Betrachtungsweise (sozusagen Gauss Elimination ohne den Jordan Zusatz) beendet das Eliminationsverfahren sobald man in der linken "unteren Ecke" (markiert durch die Pivot Elemente) nur Nullen stehen hat, d.h. im Idealfall, sobald man die

[12] Achtung: man muss dabei die "operierende Zeile" im System belassen!! Erst in einem weiteren Schritt kann dann - wenn erwünscht - die neue Zeile wieder auf der alten zur Wirkung gebracht werden

Matrix auf obere Dreiecksgestalt gebracht hat, weil dann die Lösung des Systems durch **Rückwärtssubstituieren** erfolgen kann.

Da dieser Idealfall nicht immer erreichbar ist (sicherlich nicht dann, wenn die Ausgangsmatrix etwas rechteckig ist!) macht man folgenden Kompromiss: Eine Matrix ist in *! Zeilenstufenform* [13], wenn jede von Null verschiedene Zeile mit einer Eins beginnt, die Position der führenden Einsen (der sog. **Pivot**-Elemente der Matrix) von Zeile zu Zeile mindestens um eins nach rechts rückt, und somit letztendlich alle Zeilen die nur aus Nullen bestehen am unteren Ende der Matrix zu finden sind.

Eine Matrix ist in **reduzierter** Zeilenstufenform, wenn die Spalten auch oberhalb der Pivot-Elemente nur Nullen enthalten (das wird normalerweise als der Jordan-Schritt der *Gauss-Jordan* Elimination gesehen!).

Ein typisches Beispiel einer solche Form ist

$$\begin{pmatrix} 1.0000 & 0.0000 & 1.0000 & 0.0000 & -0.3333 \\ 0.0000 & 1.0000 & 1.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.3333 \end{pmatrix} \tag{19}$$

Aber warum sind diese 3 Schritte sinnvoll und wirksam (in Bezug auf die Herstellung des Idealzustandes)??

Dazu betrachten wir die Bedeutung jedes einzelnen Schrittes im Rahmen von Gleichungssystemen der Form (beispielsweise)

$$\begin{array}{llll} 1x & +4y & +7z = 12 \\ 2x & +5y & +8z = 15 \\ 3x & +6y & +9z = 18 \end{array} \tag{20}$$

Wir haben drei Fälle zu betrachten. Am einfachsten ist die Permutation der Zeilen zu erklären. Offenbar kommt es bei der Lösung eines Gleichungssytems, d.h. bei der Suche nach einem Tripel (x, y, z) sodass in allen drei Gleichungen Gleichheit gilt, *nicht auf die Reihenfolge der Gleichungen an.* Ebenso ändert sich in Wirklichkeit nichts an einer Gleichung wenn sie (für sich, und die ganze Zeile) mit einer festen Zahl multipliziert wird. Genau wenn diese Zahl ungleich Null ist hat sie im Körper $\mathbb{K}$ ein (multiplikativ) inverses Element, und daher ist dieser Schritt auch umkehrbar. Addiert man zwei gültige Gleichungen, so erhält man natürlich auch eine gültige Gleichung, und weil man ja vorher und nacher mit einer Zahl ungleich Null multiplizieren darf ist es also auch möglich, ein *Vielfaches* einer (anderen, der operierenden) Zeile zu einer anderen zu Addieren.

Soweit so gut. Lösungen des ursprünglichen Gleichungssytems sind also auch Lösungen des modifizerten Gleichungsystem. Es gilt aber auch **die Umkehrung!**. Lösung des modifizerten Gleichungssystems zu sein ist eine (notwendige) Bedingung für eine Lösung des gegebenen Systems. Aber ist sie auch hinreichend[14]?

In der Tat, denn jeder Schritt ist umkehrbar! Man vertauscht nochmals die betroffenen Zeilen, man *multipliziert !* durch den Kehrwert der von Null verschiedenen Zahl, oder

[13] `rref(A)` in MATLAB liefert die *reduced row echelon form*

[14] Man denke an das Beispiel: Um zu zeigen, dass $3 = 5$ gilt, schliesse man, dass dann ja auch $5 = 3$ gelten müsse, da aber $3 + 5 = 5 + 3$ ist, hat man etwas Richtiges bekommen! Also hat man damit bewiesen, dass die Gleichung $3 = 5$ gilt??

man *subtrahiert* das entsprechende Vielfache der operierenden Zeile, und kann so wieder zum Ausgangssystem zurückgewinnen.

Die einzelnen Schritte der Gauss-Elimination und somit die ganze Umformungskette bewahren also die Lösungsmenge, und man kommt zu dem Schluss, dass ein Gleichungssystem genau dann lösbar ist, wenn es in der reduzierten Z-ST-Form lösbar ist. Man spricht dann oft von der **Konsistenz** des inhomogenen linearen Gleichungssytems.

In unserem Falle ist die erweiterte Systemmatrix $\widetilde{\boldsymbol{A}}$ von der Form

$$\begin{pmatrix} 1 & 4 & 7 & 12 \\ 2 & 5 & 8 & 15 \\ 3 & 6 & 9 & 18 \end{pmatrix} \tag{21}$$

und die reduzierte Matrix $\mathrm{rref}(\widetilde{\boldsymbol{A}})$ hat folgende Form

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{22}$$

und das ist OK. Es ist sogar so, dass noch Freiheitsgrade sind. Wir können z beliebig festsetzen und bekommen dafür immer noch eine Lösung.

Normalerweise führt man eine *freie Variable* z.B. λ ein, und bekommt aus der zweiten Zeile: $y = 3 - 2\lambda$ und daraus dann $x = \lambda$. Das Gleichungssystem hat also unendlich viele Lösungen, die durch $\lambda \in \mathbb{R}$ *parametrisiert* sind. Die Lösungsvektoren sind also von der Form

$$[x, y, z] = [\lambda, 3 - 2\lambda, \lambda] = [0, 3, 0] + \lambda \cdot [1, -2, 1].$$

das ist eine Gerade im $\mathbb{R}^3$ mit Richtungsvektor $\vec{\mathbf{x}} = [1, -2, 1]$, durch den Punkt $[0, 3, 0]$. Beispielsweise bekommen wir für $\lambda = 1$ den Punkt $[1, 1, 1]$. Wir werden noch später sehen, welche spezielle Bedeutung diese *spezielle Lösung* des Gleichungssystems hat.

Man hätte auch eine andere spezielle Lösung finden können, indem man einfach $x_3 = 0$ gesetzt hätte. Dann wäre in dem obigen Gleichungssystem die "unnötige Variable" (die irgendwie zuviel ist) einfach verschwunden, und man hatte eine *andere spezielle* Lösung gefunden, nämlich $\vec{\mathbf{x}} = [0, 3, 0]$. Natürlich sieht man schnell, dass die Differenz der beiden speziellen Lösungen gerade wieder ein Vielfaches des Richtungsvektors $[1, -2, 1]$ ist.

Wie wir sehen werden, ist das genau typisch für den allgemeinen Fall. *Weil es 2 Pivot-Elemente gibt*, gibt es nur zwei wirklich relevante Gleichungen, also *einen* $(3 - 2)$ *freien Parameter*!

Was ein Problem gewesen wäre, wenn wir z.B. in der letzten Zeile eine Gleichung der Form $0 = 7$ gehabt hätten ($>$ *Inkonsistenz* des inhom. lin. Gleichungssystems!).

5.3 Relationen zwischen Spalten

NEUES MATERIAL 10.4.2014

Die Interpretation der einzelnen Schritte der Gauss(-Jordan) Elimination können (siehe !??) durch Multiplikation der (erweiterten) Systemmatrix mit Elementarmatrizen von

links auf (red.) Zeilenstufenform gebracht. Wichtig ist dabei, dass jeder Schritt umkehrbar ist (äquivalent formuliert: jede der verwendeten Elementarmatrizen ist invertierbar, und de facto ist die inverse Matrix wieder eine Elementarmatrix von der gleichen Bauart[15]) und daher ist die Lösungsmenge des zugehörigen Gleichungssystems unverändert. Am Ende der Reduktion sieht man auch, dass de facto von den n Variablen durch r (Rang von $\boldsymbol{A}$) Gleichungen (denn nur soviele bleiben am Ende des Verfahrens übrig) ein $n-r$-dimensionaler Lösungsraum des *homogenen* Gleichungssystems übrigbleibt. Ein entsprechendes inhomogenes GLS hat dann also $n-r$ frei Variable oder es ist inkonsistent. Algebraisch bedeutet das, wenn wir mir $\mathbf{R}A$ die (Row reduced echelon form, d.h. eine) *Zeilenstufenform von $\boldsymbol{A}$ bezeichnen, dass gilt*

$$\mathbf{RA} * \vec{\mathbf{x}} = \vec{\mathbf{b}} \quad \Leftrightarrow \quad \boldsymbol{A} * \vec{\mathbf{x}} = \vec{\mathbf{b}}. \tag{23}$$ linrelRA1

Wir werden diese Aussage gleich noch verwerten.

Als zweites Faktum ist klar, dass in jedem Schritt nur immer wieder neue Linear-Kombinationen von Zeilen der Matrix gebildet werden. Das heißt, dass die jeweils neue Matrix nur aus Linear-Kombinationen der Zeilen der vorhergehenden Matrix besteht, d.h. dass sich die gesamt Prozedur (alle jemals auftretenden Zeilenvektoren sind ja somit Linearkomb. der Zeilen der ursprünglichen Matrix!) im Zeilenraum von $\boldsymbol{A}$ abspielt. Da aber jeder Schritt umkehrbar ist, und die Zeilen der Zeilenstufenform schrittweise wieder zu den Zeilen der Matrix $\boldsymbol{A}$ kombiniert werden können, gilt auch die umgekehrte Enthaltensrelation. Mit anderen Worten, der Zeilenraum $Z(\boldsymbol{A})$ von $\boldsymbol{A}$ stimmt mit dem Zeilenraum von $\mathbf{RA}$ überein, und da die Zeilen von $\mathbf{RA}$ offenbar linear unabhängig sind[16] *sind die ersten r Zeilen von $\mathbf{RA}$ eine Basis für den Zeilenraum von $\boldsymbol{A}$.*

Aber die reduzierte (!) Zeilenstufenform enthält auch noch andere Informationen, welche sich aus der allgemeinen Gültigkeit von (23) ergibt. linrelRA1

Zwei konkrete Konsequenzen können wie folgt verbalisiert werden:

- *Wenn eine Kollektion von Spalten von $\mathbf{RA}$ linear unabhängig bzw. linear abhängig sind, dann gilt dasselbe für die entsprechenden Spalten von $\boldsymbol{A}$. Daraus folgen insbesondere die beiden folgenden Konsequenzen:*
- *Die Pivotspalten von $\boldsymbol{A}$ sind linear unabhängig (d.h. wenn man die Spalten $p_1, \cdots, p_r$ in $\mathbf{RA}$ identifiziert hat, in denen sich die Pivot-Elemente von $\mathbf{RA}$ finden, d.h. die führenden Einsen, so kann man sicher sein, dass die zugehörigen Spaltenvektoren in $\mathbf{a}_{p_1}, \cdots, \mathbf{a}_{p_r}$ von $\boldsymbol{A}$ eine linear unabh. Menge im Spaltenraum $\mathrm{Sp}(\boldsymbol{A})$ sind.*
- *In der red. ZSF $\mathbf{RA}$ ist andererseits sofort erkennbar, dass jede Non-Pivot-Spalte $\vec{\mathbf{q}}$ ein Linearkombination der vorhergehenden Spalten von $\mathbf{RA}$, wobei die Koeffizienten direkt ablesbar sind, weil ja der Beitrag der k-ten Pivotspalte, die ihren führenden Einser genau in der k-ten Zeile hat, genau die k-te Komponente von $\vec{\mathbf{q}}$ ist!*

[15] Empfehlung an die LeserIn: Bitte für jede der drei Arten überprüfen, ob die Behautpung auch eigenständig nachweisbar ist, durch Angabe der Elementarmatrizen und deren jeweiligen Inverse!

[16] es genügt, sie auf die Pivotspalten einzuschränken, wo sie zu einer $r \times r$ Einheitsmatrix werden, daher linear unabhängig sind, und wenn ein lineares (!) Bild von Vektoren linear unabh. ist, dann müssen es auch die Vektoren selber sein! In der Vorlesung haben wir einen direkten Beweis durchgeführt!

In der Tat finden sich in jeder Non-Pivot-Spalte von $\mathbf{RA}$ *nur in den ersten Zeilen von Null verschiedene Eintragungen, sodass man sagen kann, dass die (beispielsweise) dritte Non-Pivotspalte, die sich etwa hinter der fünften Pivotspalte findet, eben eine Linear-Kombination der* 5 *Pivotspalten davor ist, und die Koordianten* $q_1, \cdot, q_5$ *sagen genau, wieviel von den Pivotspalten benötigt wird, um die Spalte* $\vec{\mathbf{q}}$ *als Linearkombination der Pivotspalten von* $\mathbf{RA}$ *zu schreiben. Das ist aber trivial, weil die Pivotspalten von* $\mathbf{RA}$ *in unserem Fall genau die ersten* 5 *Einheitsvektoren im* $\mathbb{R}^m$ *sind. Konkret muss ja* $\vec{\mathbf{q}}$ *die* 8*-te Spalte von* $\mathbf{RA}$ *sein, denn es gibt* 2 *Non-Pivot und* 5 *Pivotspalten in unserer Situation. Wir können also insgesamt sagen, in der konkreten Sitution gilt für die* 8*-te Spalte von* $\boldsymbol{A}$ *(diese entspricht von der Position her der Spalte* $\vec{\mathbf{q}}$ *die wir jetzt genau diskutiert haben, und die eine Non-Pivot-Spalte ist!):*

$$\mathbf{a}_8 = \sum_{j=1}^{5} q_j \mathbf{a}_{p_j}.$$

Mit anderen Worten, die Werte der Non-Pivotzeilen geben Auskunft darüber, welche linearen Relationen zwischen der 8-ten Spalte von $\boldsymbol{A}$ *und den davorliegenden Pivotspalten von* $\boldsymbol{A}$ *besteht. Im Sinne der üblichen Beschreibung von linearer Abhängigkeit des Systems kann man diese Gleichung natürlich in ein homogenes Gleichungssystem umschreiben, indem man einfach die Spalte* $\mathbf{a}_8$ *(dann mit Koeffizient* -1*) auf die andere Seite bringt.*

Hinweis: Es könnte hilfreich sein, sich selbst, allenfalls mit der Hilfe von OCTAVE/MATLAB eine derartiges Beispiel zusammenzustellen!

Somit enthält die red. ZSF $\mathbf{RA}$ *all nötigen Informationen, im aus den Pivotspalten von* $\boldsymbol{A}$ *die gesamte Matrix zu rekonstruieren.*

6 Vektorräume über allgemeinen Körpern

Hier werden wir nochmals die Körperaxiome kurz wiederholen:

defkoerper

Definition 5. Eine Menge $\mathbb{K}$ ist ein **Körper**, wenn auf $\mathbb{K}$ zwei Operationen, genannt *Addition* bzw. *Multiplikation* gegeben sind, sodaß $(\mathbb{K}, +)$ eine abelsche (d.h. kommutative) Gruppe ist (wir schreiben $-x$ für das additive Inverse), und $(\mathbb{K} \setminus \{0\}, \cdot)$ ebenfalls eine kommutative Gruppe ist, wobei das neutrale Element bzw. der Multiplikation mit dem Symbol 1 bezeichent wird. Weiter wird vorausgesetzt, dass es zwischen den beiden Operationen (Addition bzw. Multiplikation) die entsprechenden *Distributivgesetze* gelten, wie z.B. $x \cdot (y + z) = x \cdot y + x \cdot z$.

Hinweis: Im Englischen wird oft das Symbol $\mathbb{F}$ *(für "field" == Körper) verwendet, wir verwenden das Symbol* $\mathbb{K}$ *(typischerweise* $\mathbb{R}$ *oder* $\mathbb{C}$*) in den konkreten Anwendungen. Die meisten Aussagen sind unabhängig davon, welcher Körper verwendet wird (erst im zweiten Teil, wenn es um Wurzeln des charakteristischen Polynoms geht, wird es gravierende Unterschiede geben, dann sollte man besser* $\mathbb{C}$ *verwenden).*

Wie in den bekannten Fällen ($\mathbb{K} = \mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$) üblich schreiben wir für das multiplikative inverse Element zu x einfach x^{-1} bzw. $1/x$, und für $x \cdot 1/y$ auch verkürzt x/y. Anstatt von vier Grundrechnungsarten zu sprechen bevorzugen wir es meistens, von zwei Grundrechnungsarten zu sprechen. Division ist also beispielsweise die Multiplikation mit dem inversen Element (multiplikative Gruppe). Hat man also einmal bewiesen, dass $(x/y)^{-1} = y/x$ gilt, ist klar dass die "Division" durch x/y besser/äquivalent als Multiplikation mit dem "Kehrwert", also mit y/x realisiert werden kann/sollte.

Als nächstes folgt die Definition eines Vektorraumes über einem Körper $\mathbb{K}$:

defvektraum

Definition 6. Eine Menge $\boldsymbol{V}$ heißt **Vektorraum** über einem Körper $\mathbb{K}$ wenn einerseits $(\boldsymbol{V}, +)$ eine kommutative Gruppe[17] bezüglich einer als *Addition* bezeichneten Operation ist, anderseits aber auch eine (damit verträgliche) *skalare Multiplikation* definiert ist, d.h. eine *äußere Wirkung von Skalaren* (d.h. Elementen von $\mathbb{K}$), die wie eine Multiplikation angeschrieben wird. Genauer formuliert: Es gibt eine Abbildung von $\mathbb{K} \times \boldsymbol{V}$ nach $\boldsymbol{V}$, die einem Paar $(\lambda, \mathbf{v})$ ein Elemente $\lambda \cdot \mathbf{v}$ bzw. kurz $\lambda\mathbf{v} \in \boldsymbol{V}$ zuordnet. Diese Wirkung ist *assoziativ* und in doppeltem Sinne *distributiv* (d.h. es mögen die üblicherweise zu erwartenden Verträglichkeitsbedingungen zwischen der Addition in $\boldsymbol{V}$ und der Multiplikation und Addition in $\mathbb{K}$ gelten, mit folgenden Eigenschaften (siehe [6], Def. 7.4.30, p.460)[18]

- **(NG)**[19] $\mathbf{1v} = \mathbf{v}$ (wo $\mathbf{1}$ multiplikatives Einselement in $(\mathbb{K} \setminus \{\mathbf{0}\}, \cdot)$ ist);
- **(AG)** $\mu(\lambda\mathbf{v}) = (\mu\lambda)\mathbf{v}$; (Assoziativ-Gesetz);
- **(DG1)** $\lambda(\mathbf{v}_1 + \mathbf{v}_2) = \lambda\mathbf{v}_1 + \lambda\mathbf{v}_2$; Distributiv Gesetz;
- **(DG2)** $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$, Distributiv Gesetz;

Ausgehend von den Definitionen kann/sollte man eigentlich zeigen (mit Hilfe von vollständiger Induktion), dass scheinbar "offensichtliche" Aussagen gemacht werden können, wie etwa:

In jedem Vektorraum sind nicht nur Summen und Skalarprodukte wohldefiniert, sondern auch endliche Linear-Kombinationen der Form $\sum_{k=1}^{n} \lambda_k \mathbf{v}_k$, welche ausserdem nicht von der Reihenfolge der Summanden abhängen. Hat man das gezeigt, kann man (hoffentlich leicht!?) selber verifizieren, dass gilt: Für $k \in \mathbb{N}, \mathbf{v} \in \boldsymbol{V}$ gilt: $k\mathbf{v} = \sum_{j=1}^{k} \mathbf{v}$.

zerotvec1

Lemma 1. *In jedem Vektorraum gilt $0 \cdot \mathbf{v} = \mathbf{0}$.*
In Worten: Multipliziert man einen Vektor mit der "Zahl" Null (d.h. dem neutralen Element bzgl. der Addition im Körper $\mathbb{K}$) dann erhält man den Nullvektor in $\boldsymbol{V}$, d.h. das neutrale Element in der additiven Gruppe $(\boldsymbol{V}, +)$.

[17] Man darf sich hier nicht verwirren lassen, das "$+''$-Symbol hier ist ein anderes als das bei der Körper Addition, aber es wird jeweils klar sein, z.B. beim Distributivgesetz (DG2) unten, welche Art der Addition gerade verwendet wird, nämlich die in $\mathbb{K}$ auf der linken, und die in $\boldsymbol{V}$ auf der rechten Seite!

[18] Man beachte dass die Multiplikation zwischen μ und λ innerhalb von $\mathbb{K}$ stattfindet, während die anderen Punkte die (externe) skalare Multiplikation, d.h. die Wirkung von $\mathbb{K}$ auf $\boldsymbol{V}$ ansprechen.

[19] Das soll im Prinzip nur die triviale Wirkung $(\lambda, \mathbf{v}) \mapsto \mathbf{0}$ ausschliessen!

Proof. Unter Verwendung der verschiedenen Axiome (insbesondere (NG) und des (DG2)) erhalten wir:

$$\mathbf{v} = 1 \cdot \mathbf{v} = (1+0) \cdot \mathbf{v} = 1 \cdot \mathbf{v} + 0 \cdot \mathbf{v} = \mathbf{v} + 0 \cdot \mathbf{v},$$

d.h. man kann den Vektor $0 \cdot \mathbf{v}$ zu $\mathbf{v}$ hinzuaddieren, ohne dass sich etwas ändert.

Addiert man nun zu dieser Gleichung das additive Inverse von v, d.h. den Vektor $-\mathbf{v}$, so bekommt man:

$$\mathbf{0} = 0 \cdot \mathbf{v},$$

und der Beweis ist fertig. □

Hinweis: Eine etwas umständlichere aber durchaus naheliegende Art der Argumentation (ebenfalls korrekt/akzeptabel) für den letzten Teil wär: Es ist zu zeigen, dass für *jedes*(!) Element $\mathbf{v}' \in \boldsymbol{V}$ gilt: $\mathbf{v}' + 0 \cdot \mathbf{v} = \mathbf{v}'$. Das folgt aber leicht aus dem bewiesenen Spezialfall ($\mathbf{v}' = \mathbf{v}$) durch:

$$\mathbf{v}' + 0 \cdot \mathbf{v} = (\mathbf{v}' - \mathbf{v}) + \mathbf{v} + 0 \cdot \mathbf{v} = (\mathbf{v}' - \mathbf{v}) + \mathbf{v} = \mathbf{v}'.$$

Damit ist der (nun korrekte) Beweis fertig (!mögliche Prüfungsfrage).

6.1 Eine lange Liste von Beispielen

Im Prinzip können wir bei jedem der nun aufgelisteten Vektorräumen (nachdem im Einzelfall die Axiome gecheckt worden sind, entweder durch akribisches Nachprüfen, manchmal durch "Hinschauen", manchmal auch etwas länger!) etwa darauf untersuchen, welche Dimension sie haben!

1. $\mathbb{Q}^n, \mathbb{R}^k, \mathbb{C}^m$, d.h. die geordneten $n/k/m$-Tupel von Elementen aus dem jeweiligen Grundkörper, bilden einen Vektorräum über eben diesen Körper. Dabei ist es gleichgültig, ob man sich die geordneten n-Tupel als Spalten (in der Literatur bevorzugt) oder als Zeilenvektoren schreiben will, oder nur geometrisch interpretiert werden (Pfeile im Raum);

   In MATLAB macht man folgenden Unterschied: $\vec{s} = [1;2;3;4]$; ergibt einen Spaltenvektor, hingegen ist $\vec{z} = [1,2,3,4]$, ein Zeilenvektor.[20]

2. Wir haben schon gesehen, dass Zeilenraum, Spaltenraum, Nullraum (bzw. die Lösungsmenge eines homogenen, linearen GLS) Teilräume von $\mathbb{K}^s$ sind (s passend), d.h. dort ist die "abstrakte Vektoraddition" noch einfach die vom großen Raum, d.h. von $\mathbb{K}^s$ aller Vektoren ererbte Addition.

3. Für ein festes paar natürlicher Zahlen $m, n > 0$ ist die Menge $\mathscr{M}_{m,n}$ aller $\mathbb{K}$-wertiger $m \times n$-Matrizen ein Vektorraum über $\mathbb{K}$ (Addition und Multiplikation koordinatenweise!).

$$2 \cdot \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 6 \\ 4 & 8 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 3 & 5 \end{pmatrix}$$

[20] Wobei z auch angezeigt wird (weil das Kommando mit "," endet), hingegen s nicht (weil es mit einem ";" endet). Probieren Sie auch die Eingabe $[1;2;3;4]$; bzw. $[1,2,3,4]$; oder $[1,2,3,4]'$.

4. Für eine beliebige Menge X sei

$$\mathscr{F}(X) := \{f : x \mapsto f(x) \in \mathbb{R} \quad \text{or} \quad \mathbb{C}\},$$

die Menge der *reell-wertigen* (resp. komplex-wertigen) Funktionen auf X. Die naheliegende Vektorraumstruktur ist durch folgende Festsetzung gegeben:

$$(f+g)(x) = f(x) + g(x), (\lambda f)(x) = \lambda f(x), \quad x \in X, \lambda \in \mathbb{K}. \tag{24}$$

addfunct1

Geogebra erlaubt es, das Rechnen mit diesen Funktionen genau so zu realisieren, wie es "abstrakt" erklärt wird. Um es einzuüben schlage ich vor, folgende Befehle einzugeben:

```
f:  2 sin(x)
g:  3 cos( 2.7 x)
h: f+g
[dann plotten und die drei Funktionen mit Farbe ausstatten!]
```

Auch das Beispiel in der Einführung diese Skriptums benutzt diese Struktur! Je nach Sichtweise handelt es sich um reelle Linearkombinationen der komplexwertigen Funktionen $\cos(x)$ und $i\sin(x)$ oder um komplexe Linearkombinationen der (reellen) Funktionen cos bzw. sin und andererseits der entsprechenden komplexen Exponentialfunktionen $x \mapsto e^{\pm ix}$.

5. Die Menge der Polynomfunktionen (siehe unten) auf $\mathbb{R}$, oder $\mathbb{C}$, oder auf einem Intervall $I = [a, b]$, etc.. Speziell die quadratischen und kubischen Polynomfunktionen werden wichtig *halb-abstrakte* Beispiele sein, die wir mit $\mathscr{P}_2(\mathbb{R})$ bzw. $\mathscr{P}_3(\mathbb{R})$ bezeichnen werden. Gerade zu diesen Vektorräumen werden wir viele konkrete Überlegungen anstellen. Ebenso könnte man die Menge der *stetig differenzierbaren* Funktionen auf $\mathbb{R}$ betrachten, etc..

6. Die Menge der *stetigen Funktionen* auf $I = [a, b]$ (siehe Analysis Vorlesung), oft als $C([a, b])$ bezeichnet[21].

7. Die Menge der reellwertigen Stufenfunktionen, d.h. der Funktionen, die nur endlich viele reelle Werte auf Teilintervallen $I_j = [a_j, b_j)$ von $\mathbb{R}$ annehmen. Übungsaufgabe: man verifiziere, dass dies Menge bzgl. der Addition von Funktionen abgeschlossen ist, und klarerweise ebenso bzgl. skalarer Multiplikation. (Zusatzbemerkung: Sogar die punktweise Multiplikation, definiert durch

$$(f \cdot g)(x) := f(x) \cdot g(x)$$

ist möglich und macht diese Menge zu einer multiplikativen Halbgruppe.

8. In der Funktionalanalysis betrachtet man z.B. den Vektorraum $\boldsymbol{\ell}^\infty$ aller komplexwertigen Folgen $\mathbf{x} = (x_k)_{k\in\mathbb{Z}}$ welche beschänkt sind, ebenfalls mit der *koordinatenweisen* Addition und Skalaren Multiplikation (!Exercise!). Die Menge

$$\boldsymbol{c}_0 := \{\mathbf{x} \in \boldsymbol{\ell}^\infty \mid \lim_{k\to\infty} x_k = 0\}$$

[21] "C" steht für die englische Bezeichnung von Stetigkeit, stetig = continuous!

ist ein Teilraum, d.h. auch selber wieder eine Vektorraum (der "unendlichdimensional" ist, d.h. der *keine endliche Basis* hat, weil es zu jedem $K \in \mathbb{N}$ mehr als K linear unabhängige Folgen gibt, z.B. die "Einheitsvektoren" $\mathbf{e}_k$, definiert in der naheliegenden Weise!)

9. Für zwei Vektorräume $\boldsymbol{V}, \boldsymbol{W}$ über demselben Körper ist die Menge der linearen Abbildungen (oft auch Menge der *linearen Operatoren*, d.h.

$$\mathcal{L}(\mathbf{V},\mathbf{W}) := \{T \,|\, T : \boldsymbol{V} \to \boldsymbol{W}, T \,\text{linear}\} \tag{25}$$

linopdef1

selbst wieder ein Vektorraum mit den (naheliegenden) Operationen

$$[T_1{+}T_2](\mathbf{v}) := T_1(\mathbf{v}) + T_2(\mathbf{v}) \quad \text{and} \quad [\lambda{\cdot}T](\mathbf{v}) = \lambda \cdot (T(\mathbf{v}))$$

10. Ein wichtiger Spezialfall ist gegeben durch $\boldsymbol{W} = \mathbb{K}$. In diesem Fall spricht man von *linearen Funktionalen* auf $\boldsymbol{V}$, d.h. linearen Abbildungen $\mathbf{v}^* : \boldsymbol{V} \to \mathbb{K}$. Anstatt $\mathcal{L}(\mathbf{V},\mathbb{K})$ verwendet man normalerweise das einfachere Symbol $\mathbf{V}^*$, und nennt diesen Raum den **Dualraum** zu $\boldsymbol{V}$. Später werden wir noch sehen, dass der Dualraum die gleiche Dimension hat wie der ursprüngliche Raum, und dass es zu jeder Basis $\mathscr{B}$ für $\mathbf{V}$ eine eindeutig bestimmte *duale Basis* gibt (manchmal spricht man auch von einer *Biorthogonalität* der beiden Basen)[22].

11. Ein andere Vektorraum von Funktionen (Funktionenraum) ist etwas $\boldsymbol{C}_b(\mathbb{R})$, der als Menge definiert ist als:

$$\boldsymbol{C}_b(\mathbb{R}) := \{f : \mathbb{R} \to \mathbb{C}, f \,\text{stetig}\,, \exists\, C > 0 : |f(x)| \leq C \;\forall\; x \in \mathbb{R}\}$$

Es ist dieser Vektorraum in dem etwa die Eulersche Formel (siehe Eingangsbeispiel) realisiert wird.

12. Für jedes n ist die Menge der *unteren Dreicksmatrizen*, d.h. der Matrizen $\boldsymbol{A}$ mit $a_{j,k} = 0$ für $k > j$, einen Teilraum aller $n \times n$ Matrizen. Die Dimension ist ?? Natürlich gelten analoge Aussagen für die *oberen Dreiecksmatrizen* (mit $a_{j,k} = 0$ für $j > k$).

Dieser Teilraum ist natürlich identifizierbar mit dem Raum aller *symmetrischen, rellen $n \times n$-Matrizen* $\boldsymbol{B}$, indem man $b_{j,k} = a_{j,k}$ setzt für $k \leq j$ und $b_{j,k} = a_{k,j}$ für $k > j$ setzt (erzeugt einen linearen Isomorphismus, Beweis?). Klarerweise ist dann eine mögliche Basis dieser Matrizen die Kollektion von Matrizen welche genau einen 1-er auf der Hauptdiagonale haben, *oder* zwei symmetrisch verteilte 1-er abseits der Hauptdiagonale, also beispielsweise ($n = 4, j = 4, k = 2$ bzw. $k = 4, j = 2$):

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{26}$$

[22] Das konkrete Beispiel werden die Punktauswertungen, sagen wir auf $\mathscr{P}_3(\mathbb{R})$ sein welche biorthogonal zur entsprechenden Familie von Lagrange-Interpolations-Polynomen ist. Umgekehrt gibt es zu jeder Basis des Dualraums, d.h. zu jeder Kollektion von linear unabhängigen linearen Funktionalen mit $n = \dim(\mathbf{V})$ Elementen auch eine Basis des Vektorraumes $\mathbf{V}$ sodass die beiden Basen *biorthogonal zueinander* sind!

Sobald man ein abstraktes Konzept von Vektorräumen hat, ist es auch klar, wie andere Begriffe in den allgemeinen Kontext zu übertragen sind.[23]

[23] In der Tat, FAST alles was in der Vorlesung (2014) bis Ostern gemacht wurde, kann gleich nochmals in dieser Allgemeinheit durchgegangen und wiederholt werden und behält seine Gültigkeit, weil nur an ganz wenigen Stellen konkrete Eigenschaften benutzt wurden. Dies betrifft vor allem die Rolle der Einheitsvektoren $(\mathbf{e}_k)_{k=1}^n$ im $\mathbb{K}^n$, für die wir zunächst keine unmittelbares Analogon im abstrakten Setting haben. Da jedoch die Polynomfunktionen eine "natürliche" Basis haben, nämlich die Monome, werden wir dieses Beispiel als nächstes unter die Lupe nehmen!

6.2 Teilräume, Basen, lineare Abbildungen

Die bisher im Kontext von klassichen Vektoren (geordnete n-Tupel in $\mathbb{K}^n$, oft als Spaltenvektoren betrachtet) entwickelten Konzepte lassen sich - soferne nicht ganz konkret die Einheitsvektoren ($\mathbf{e}_k$) in $\mathbb{K}^n$ verwendet wurden weitgehend wort- bzw. inhaltsident auf allgemeine Vektorräume übertragen. Dabei wollen wir schnell diejenigen Begriffe betrachten, die nicht (einmal) von einer "endlichen Basis" (d.h. der Endlich- Dimensionalität des Vektorraumes) Gebrauch machen.

Wir haben folgende *Begriffe* bzw. weiterhin (mit den gleichen Argumenten wie bisher) gültige Aussagen.

Ein *Teilraum* $\boldsymbol{W}$ eines Vektorraumes $\boldsymbol{V}$ ist eine Teilmenge, welche bzgl. der (bereits in $\boldsymbol{V}$ definierten) Addition bzw. skalarer Multiplikation abgeschlossen ist. Es gilt auch wieder: Eine Menge M ist ein Teilraum genau dann wenn mit jeder endlichen Folge (m_k) in M und Skalar-Folge (c_k) in $\mathbb{K}^n$ gilt: $\sum_{k=1}^{n} c_k m_k \in M$.

Es gilt auch jetzt noch/wieder: Hat man eine Menge $M \subseteq \boldsymbol{V}$, so gibt es einen kleinsten Vektorraum $\boldsymbol{V}_M$ in $\boldsymbol{V}$, welcher M enthält. Da auch jetzt gilt: eine Linear-Kombination von Linear-Kombinationen von Elementen von M kann (!Matrix-Produkt) als Linear-Kombination von Elementen von M geschrieben werden, ist klar, dass dieser kleinste M umfassende Teilraum von $\boldsymbol{V}$ genau die sog. *lineare Hülle* von M ist, d.h. die Menge der endlichen Linear-Kombinationen von M (wir schreiben weiterhin $\mathbf{LH}(M)$.

Eine Menge $M \subseteq \boldsymbol{W}$ heisst *Erzeugendenmenge* für $\boldsymbol{W}$ wenn gilt: $\boldsymbol{W} = \mathbf{LH}(M)$. Eine Menge M heisst *linear unabhängig* wenn gilt: Es bestehen keine linearen Relationen zwischen den Elementen der Menge M, d.h. falls eine Linearkombination $\sum_{k=1}^{n} c_k m_k = 0$ erfüllt, dann muß $\vec{\mathbf{c}} = \mathbf{0}$ gelten. Eine Menge $\mathscr{B}$ heisst *Basis* von $\boldsymbol{V}$ (oder einen Teilraum $\boldsymbol{W}$) wenn sie eine linear unabhängige Erzeugendenmenge ist.

Je zwei Basen von $\boldsymbol{V}$ haben gleich viele Elemente, woraus sich ergibt: Falls $\boldsymbol{V}$ eine endliche Basis von hat (man sagt: falls $\boldsymbol{V}$ endlich-dimensional ist), dann haben alle möglichen Basen von $\boldsymbol{V}$ gleich viele Elemente. Ihre Anzahl wird als die *Dimension* von $\boldsymbol{V}$ bezeichnet. ACHTUNG: Dies ist im Moment nur eine Behauptung, die in der Vorlesung am 9.April bewiesen wurde, aber an dieser Stelle noch nicht im Skriptum ausgeführt wurde [Notiz vom 26.4.2014]

Eine Abbildung T von $\boldsymbol{V}_1$ nach $\boldsymbol{V}_2$ ist linear genau dann wenn sie Linear-Kombinationen respektiert. Das Bild eines Teilraums $\boldsymbol{W}_1$ von $\boldsymbol{V}_1$ ist ein Teilraum von $\boldsymbol{V}_2$, ebenso wie das Urbild $T^{-1}(\boldsymbol{W}_2)$ eines Teilraums $\boldsymbol{W}_2$ von $\boldsymbol{V}_2$ ein Teilraum von $\boldsymbol{V}_1$ ist. (> Exercises)

Falls T eine bijektive und lineare Abbildung ist, so ist die inverse Abbildung $S = T^{-1}$ automatisch linear. Man spricht dann von einem *Isomorphismus* von Vektorräumen. Falls $\boldsymbol{V}_1 = \boldsymbol{V}_2$ ist, nennt man T einen *Automorphismus*.

Analog zur Bildung von $\mathbb{C}$ aus zwei Kopien von $\mathbb{R}$ kann man zeigen, dass die Produktmenge $\boldsymbol{V}_1 \times \boldsymbol{V}_2$ von zwei Vektorräumen auf "natürliche Weise" selbst wieder ein Vektorraum ist (koordinatenweise Operationen).

Standard-Aussagen und Prinzipien

Da die folgenden Aussagen (im Text davor oder danach) nicht von der konkreten Art der Vektorraumstruktur bzw. der "Natur ihrer Elemente" abhängig war, sind sie auch in dieser Allgemeinheit gültig:

1. Das Bild einer linear abhängigen Menge M_1 unter einer linearen Abbildung ist linear abhängig;
2. Das Bild einer linear unabhängigen Menge M_2 unter einem Isomorphismus ist selbst wieder linear unabhängig; In der Tat genügt es, dass die lineare Abbildung injektiv ist, d.h. einen trivialen Kern hat.
3. Das Bild einer Erzeugendenmenge M_1 unter einer surjektiven linearen Abbildung T ist wieder eine Erzeugendenmenge.
4. Es sei M_1 eine Menge in $\boldsymbol{V}_1$, und T sei eine (linearer) Isomorphismus zwischen $\boldsymbol{V}_1$ und $\boldsymbol{V}_2$. Dann ist $T(M_1)$ eine Basis für $\boldsymbol{V}_2$ genau dann wenn M_1 eine Basis für $\boldsymbol{V}_1$ ist. Insbesondere haben isomorphe Räume gleich viele Basis-Elemente, d.h. sie haben die gleiche Dimension.

Aussagen, die sich erst jetzt sinnvoll formulieren lassen (und daher noch genauer ausgeführt werden sollen) sind dann von folgender Art:

isomdim01 **Theorem 1.** *Zwei Vektorräume $\boldsymbol{V}_1$ bzw. $\boldsymbol{V}_2$ über dem gleichen Körper $\mathbb{K}$ sind genau dann isomorph wenn sie die gleiche Dimension haben.*

Proof. (i) Wenn beide Vektorräume n-dimensional sind, dann sind sie beide (dank der Wahl einer Basis mit jeweils n Elementen) *isomorph zu* $\mathbb{K}^n$, und somit (Transitivität und Symmetrie des Isomorphie-Begriffes!) auch zueinander isomorph[24].

(ii) Umgekehrt gilt im Falle von verschiedener Dimensionalität n_1 bzw. n_2, dass die Räume eben zu $\mathbb{K}^{n_1}$ bzw. $\mathbb{K}^{n_2}$ isomorph sind, mit $n_1 \neq n_2$. Aber dann können die beiden Räume nicht isomorph sein, weil ja sonst (wieder unter Zuhilfenahme der Eigenschaften des Isomorphie-Begriffes) folgen würde dass $\mathbb{K}^{n_1}$ zu $\mathbb{K}^{n_2}$ (linear) isomorph wäre. Da aber eine der beiden Matrizen, die diesen Isomorphismus bzw. seine Umkehrabbildung beschreiben notwendigerweise mehr Spalten als Zeilen hätte, muss er eine linear abhängige Spalte enthalten (Gauss-Elimination: es gibt auch non-pivot-Spalten!), im Widerspruch zur Injektivitaet der entsprechenden linearen Abbildung. □

Ein ebenfalls *neuer Begriff* (in der nun vollen Allgemeinheit) ist der Begriff der Matrixdarstellung einer solchen linearen Abbildung in Bezug auf zwei Basen. Hat man eine Basis $\mathscr{B}$ in $\boldsymbol{V}_1$ und eine Basis in $\mathscr{C}$ in $\boldsymbol{V}_2$, dann können wir jede lineare Abbildung T durch eine $m \times n$-Matrix[25] beschreiben, einfach indem wir als k-te Spalte dieser $m \times n$-Matrix jeweils die $\mathscr{C}$-Koordinaten von $T(\mathbf{b}_k)$ eintragen.

[24] Falls diese knappe Formulierung zu kurz oder unklar sein sollte, bitte selbt die Details ausarbeiten bzw. es versuchen!

[25] Wir machen die Standard-Annahmen, und bezeichnen die Dimension von $\boldsymbol{V}_1$ mit n und die von $\boldsymbol{V}_2$ mit m.

matrbasrepr2 **Definition 7.** Für eine lineare Abbildung $T \in \mathcal{L}(\mathbf{V},\mathbf{W})$ zwischen zwei Vektorräumen, ausgestattet mit den Basen $\mathscr{B}$ bzw. $\mathscr{C}$, bezeichnen wir die Matrix $[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ als die Matrix, die der linearen Abbildung (dem Operator) T bzgl. der Basen $\mathscr{B}$ (in $\boldsymbol{V}$) bzw. $\mathscr{C}$ (in $\boldsymbol{W}$) zugeordnet wird.

In der allgemeinen Diskussion werden wir meistens die Dimension von $\boldsymbol{V}$ mit n und die Dimension von $\boldsymbol{W}$ mit m bezeichnen. Dann hat die Matrix $\boldsymbol{A} = [T]_{\mathscr{C}\leftarrow\mathscr{B}}$ das Standard Format $n \times m$. In anderen Büchern und Skripten werden für diese Matrix (gleicher Begriff!) verschiedenste Symbole verwendet.

Eine begrifflich durchaus "komplexe" Aussage ist dann die folgende:

isomatrlin1 **Lemma 2.** *Es seien $\mathscr{B}$ bzw. $\mathscr{C}$ jeweils Basen für $\boldsymbol{V}_1$ bzw. $\boldsymbol{V}_2$. Dann besteht durch die Abbildung $T \mapsto [T]_{\mathscr{C}\leftarrow\mathscr{B}}$ eine Isomorphismus zwischen der Menge der linearen Abbildungen von $\boldsymbol{V}_1$ nach $\boldsymbol{V}_2$ und der Menge der $m \times n$-Matrizen (über $\mathbb{K}$).*

Proof. Zunächst ist klar, dass die Abbildung $T \mapsto T[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ bei festen Basen $\mathscr{B}$ bzw. $\mathscr{C}$ in einem n- bzw. m-dimensionalen Raum die Menge der linearen Abbildungen, d.h. $\mathcal{L}(\mathbf{V},\mathbf{W})$ in die Menge der $m \times n$-Matrizen abbildet.

Weiters ist klar, dass diese Abbildung injektiv ist, weil zwei lineare Abbildungen T_1 und T_2 denen dieselbe Matrix zugeordnet wird, nach Konstruktion auf der Basis $\mathscr{B}$ übereinstimmen. Aber zwei lineare Abbildungen die auf irgendeiner Basis übereinstimmen stimmen auf allen Linearkombinationen, also auf ganz $\boldsymbol{V}$ überein.

Um noch zu zeigen, dass die Abbildung surjektiv ist, genügt es zu zeigen, dass bei vorgegebener $m \times n$-Matrix $\boldsymbol{A}$ eine (eindeutig, nämlich konstruktiv gegebene) gibt, deren zugehörige Matrix gerade $\boldsymbol{A}$ ist. Da die $\mathscr{C}$-Koordinaten der Bilder der Einheitsvektoren vorgegeben sind, muss gelten $T(\mathbf{b}_k) = \sum_{j=1}^{m} a_{j,k}\mathbf{c}_k$ (wenn die Basis Elemente mit $(\mathbf{c}_k)_{k=1}^m$ bezeichnet wurden). Somit muss T für allgemeines $\mathbf{v} = \sum_{k=1}^{n} x_k\mathbf{b}_k$ gelten:

$$T(\mathbf{v}) = \sum_{k=1}^{n} x_k T(\mathbf{b}_k) = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{j,k}x_k \right) \mathbf{c}_k,$$

d.h. in $\mathscr{C}$-Koordinaten geschrieben gilt für das Bild von $\mathbf{v}$ unter T, dass diese durch Matrix-Multiplikation der $\mathscr{B}$-Koordinaten des input Vektors $\mathbf{v}$ mit der Matrix $\boldsymbol{A}$ realisiert wird!

Da die $\mathscr{B}$ Koordinaten der Basis Vektoren "trivialerweise" gerade die Einheitsvektoren sind, und die Anwendung der Matrix-Multiplikation gerade die Spalten von $\boldsymbol{A}$ ergeben, ist klar dass die zur Abbildung T zugeordnete Matrix gerade die Ausgangsmatrix $\boldsymbol{A}$ ist.

Zuletzt ist zu bemerken, dass diese beiden Zuordnungen zueinander invers sind, also eine Bijektion liefern, und ausserdem (einfache Übungsaufgabe!) linear sind (d.h. mit der Bildung Lin.Komb. von linearen Abbildungen bzw. der entsprechenden Bildung von Lin.Komb. von Matrizen verträglich ist). Somit ist der behauptete Isomorphismus bewiesen. □

6.3 Polynomfunktionen auf $\mathbb{R}$

Dieses Kapitel wird noch später sehr viel ausführlicher kommentiert und aus/aufgeschrieben werden (Stand: Ostern 2014).

Es geht darum, die Menge der *reellen Polynomfunktionen* näher zu studieren. Wir definieren diese Menge durch die Bauart
$\mathscr{P}(\mathbb{R}) := \{p(t) \,|\, p(t) = a_0 + a_1 t + \cdots + a_n t^n$, wobei $(a_i)_{i=0}^n$ eine Folge in $\mathbb{R}$ sei.$\}$
Analog kann man (als abstrakte Objekte jedenfalls) für $k \in \mathbb{N}$ die Menge der *Polynome vom Grad* vom Grad $\leq k$ definieren (einfach nun mit Koeffizienten aus $\mathbb{K}$). Üblicherweise (vlg. Cap-Skriptum) wird diese Menge mit $\mathbb{K}_k[x]$ bezeichnet.

Bekanntlich sagt man, dass $p(t)$ *vom Grad* n ist, wenn $a_n \neq 0$ gilt. Die Anzahl der bestimmenden Koeffizienten ist dann $n+1$, man nennt das die *Ordnung des Polynoms*. Natürlich kann man Polynome bzw. Polynomfunktionen auch mit komplexen Koeffizienten bzw. mit komplexen Argumenten betrachten (das werden wir auch tun). So ist etwas das Symbol $\mathscr{P}(\mathbb{C})$ zu verstehen.

Der (wie wir sehen werden "unendlichdimensionale") Vektorraum $\mathscr{P}(\mathbb{R})$ hat aber im Gegensatz zu den meisten anderen (abstrakten bzw. konkreten) Vektorräumen reichliche zusätzliche Struktur. Beispielsweise ist er *invariant* bezüglich Translation und Streckung (allgemeiner affiner Transformation!) sowie eine *Algebra*! Das bedeutet, dass mit zwei Polynomfunktionen $p(z), q(z)$ auch die Produktfunktion $r(z) := p(z) \cdot q(z)$ eine Polynomfunktion ist (!Ausmultiplizieren). Die Bestimmung der Koeffizienten von $r(z)$ aus den Koeffizienten der Faktoren ist eng verwandt mit der Multiplikation von (großen) natürlichen Zahlen. Man nennt die (bilineare) Abbildung die der Multiplikation auf der Koeffizienten-Ebene entspricht das *Cauchy-Produkt* der entsprechenden Koeffizienten.

In MATLAB/OCTAVE kann man das mit dem Befehl `c = conv(a,b);` realisieren. Es ist auch interessant, anzumerken, dass es bei diesem Befehl nicht auf die Reihenfolge der Monome (in der allg. math. Literatur aufsteigend, mit der nullten Potenz beginnend, in MATLAB absteigend, mit der höchsten Potenz beginnend!) ankommt.

polyrekonst

Theorem 2. *Jedes (reelle oder komplexe) Polynom $p(t)$ vom Grad $\leq n$ d.h. mit Ordnung $\leq n+1$ ist eindeutig durch die Werte an $n+1$ Stellen in $\mathbb{R}$ bzw. $\mathbb{C}$ gegeben.*

Proof. Wir müssen zeigen, dass wir $p(t)$ (d.h. die Koeffizenten $(a_i)_{i=0}^n$) aus den Werten von $n+1$ Werten $p(z_j), 0 \leq j \leq n$ bestimmt werden können, wenn nur $z_j \neq z_k$, für $k \neq j$ gilt.

Wir gehen also davon aus, dass wir die Werte $p(z_j) = d_j$ haben, d.h. einen Datenvektor $(d_j)_{j=0}^n$. Wenn wir das Polynom ausschreiben, ergibt das ein Gleichungssystem für die unbekannten Koeffizienten $(a_i)_{i=0}^n$.

Jede einzelne Gleichung dieses Systems von $n+1$ Gleichungen für $n+1$ unbekannte Koeffizienten sieht dann so aus für $z = z_j$, für ein $j \in \{0, \ldots n\}$:

$$1 \cdot a_0 + z \cdot a_1 + z^2 \cdot a_2 \cdots + z^{n+1} a_{n+1}. \tag{27}$$

Vander-row1

Daraus ergibt sich eine System-Matrix, die sogenannte *Vandermonde Matrix*.

Wir werden hier nur den einfachen Fall von quadratischen Polynomen aufschreiben (der allgemeine Fall ist durch vollständige Induktion auf der Basis genau desselben Ansatzes zu realisieren, siehe Buchliteratur, z.B. [5], Chap.4.4., p.206, mittels LU-Zerlegung!).

Für den Fall, dass $z_1 = 2, z_2 = 3, z_3 = 4$ ist, sieht sie folgendermassen aus:

$$\begin{pmatrix} 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix} \tag{28}$$

d.h. wir haben in der ersten Spalte lauter Einser, in der zweiten die Abtastwerte, und in den folgenden Spalten die Potenzen (d.h. von der nullten bis zur n-ten Potenz).

$$\begin{pmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{pmatrix} \tag{29}$$

woraus man durch Subtrahieren der ersten Zeile von den weiteren Zeilen leicht

$$\begin{pmatrix} 1 & z_1 & z_1^2 \\ 0 & z_2 - z_1 & z_2^2 - z_1^2 \\ 0 & z_3 - z_1 & z_3^2 - z_1^2 \end{pmatrix} \tag{30}$$

macht, und weil $z_1 - z_2$ ebenso wie $z_1 - z_3$ nach Voraussetzung nicht gleich Null sind, und man daher die jeweiligen Zeilen durch das Pivotelement dividieren kann, bekommt man sofort

$$\begin{pmatrix} 1 & z_1 & z_1^2 \\ 0 & 1 & z_2 + z_1 \\ 0 & 1 & z_3 + z_1 \end{pmatrix} \tag{31}$$

und zuletzt durch Subtraktion der zweiten von der letzten Zeile:

$$\begin{pmatrix} 1 & z_1 & z_1^2 \\ 0 & 1 & z_2 + z_1 \\ 0 & 0 & z_3 - z_2 \end{pmatrix} \tag{32}$$

und letztendlich durch Division mit $z_3 - z_2$ (auch das ist nicht Null!)

$$\begin{pmatrix} 1 & z_1 & z_1^2 \\ 0 & 1 & z_2 + z_1 \\ 0 & 0 & 1 \end{pmatrix} \tag{33}$$

Durch Rückwärts-Substituieren können wir also in jedem Fall die (so eindeutig bestimmten) Koeffizienten für $p(t)$ bestimmen! [26] □

Ziel: Einführung von Matrix-Matrix-Multiplikation, Zusammenhang zu Linearen Abbildungen... (daraus Assoziativität der Matrix-Multiplikation, nicht nur durch Index-Jonglieren!).

[26] Achtung: das hier für kubische Polynome verwendete Argument ist nicht ganz zu einfach auf allgemeine Polynome vom Grad $n \geq 3$ anzunehmen.

6.4 Konkretes Material zur Gauss-Elimination

Das Ziel der Gauss-Elimination ist es, die vorgegebene Matrix, die ein lineares Gleichungssystem darstellt, in ein lösungsäquivalentes lineares Gleichungssytem umzuformen, dessen Lösung (praktisch betrachtet) "einfach" ist.

vollZeilR **Lemma 3.** *Der Zeilenraum einer Matrix $\boldsymbol{A}$ ist genau der ganze $\mathbb{R}^n$ wenn alle n Einheitsvektoren, $(\vec{\mathbf{e}}_1, \cdots, \vec{\mathbf{e}}_n)$ im Zeilenraum liegen.*

Daraus folgt: *Wenn* der Zeilenraum einer quadratischen Matrix $\boldsymbol{A} \in \mathcal{M}_{n,n}(\mathbb{C})$ der ganze $\mathbb{C}^n$ ist, dann ist jedes lineare, inhomogene Gleichungssystem für jede rechte Seite lösbar! (Umkehrung folgt später).

elimzeilraum **Lemma 4.** *[Eliminations-Schritte erzeugen bel. Elemente des Zeilenraumes]*
Mit Hilfe der Operation (die bei der Gauss-Elimination erlaubt sind):

Addiere das Vielfache einer Zeile zu einer anderen

sowie

Multipliziere einen Zeilenvektor mit $\lambda \neq 0$

kann man aus einer gegebenen Kollektion von Zeilenvektoren
jede mögliche Linearkombination der Zeilenvektoren *gewinnen.*

Im Prinzip (insbesondere im Falle von 2×2-Gleichungssystem (d.h. dem Schnitt von 2 Geraden in der Ebene, als ypischem Spezialfall) sind diese drei Schritte einfach zu erklären, mit folgenden Argumenten: Das Multiplizieren einer (einzelenen) Gleichung mit einem Faktor ändert die Geraden nicht (im 3×3-Fall sind es Ebenen, es ist nur eine andere Darstellung derselben Ebene, wobei man idealerweise einen "normierten" Normalvektor verwenden sollte).

Andererseits gibt es keine "natürliche Reihenfolge" für die Aufzählung der Gleichungen. Ein typisches $n \times 3$ Gleichungssystem bedeutet: Man ist an Durchschnitt von n Ebenen im $\mathbb{R}^3$ interessiert, und da gibt es keine Reihenfolge. Man kann das natürlich nutzen, um eine für die Berechnunge einfache Reihenfolge herzustellen! In anderen Fällen *muß* man eine Zeilenvertauschung vornehmen, weil der führende Koeffizient Null ist und somit nicht durch Multiplikation/Division auf den Wert 1 gebracht werden kann. Man bringt also typischerweise eine weiter unten stehende Zeile nach oben (natürlich ohne die Lösungsmenge zu ändern).

Zuletzt der (praktische gesehen) wichtigste Schritt. **Unter Festhaltung einer Zeile** addiert man ein Vielfaches dieser Zeile zu einer anderen, die dann dafür weggelassen werden kann! Klar ist wieder, dass die Linearkombination (zu verändernde Zeile, typischerweise darunterliegend) ebenfalls gültig ist, wenn die beiden involvierten Gleichungen gelten! Andererseits ist die Operation ja *umkehrbar* (durch eine gleichartige Elementarmatrix, statt zu Subtrahieren kann man das Vielfache ja wieder addieren und umgekehrt!) und somit hat das neue System *genau dieselben* Lösungen[27].

[27] Recall: Lösungen des neuen Systems sind nach dem Argument auch Lösungen des Systems im Schritt davor, und umgekehrt!

Nebenbemerkung: Da jede der Elementarmatrizen durch eine Elementarmatrix gleicher Bauart invertierbar ist (die inverse Elementarmatrix ist sowohl Links- als auch Rechts-Inverse, also Inverse im gruppentheoretischen Sinn!) ist auch das Produkt der Elementarmatrizen selber wieder eine invertierbare Matrix.

Weiters sei angemerkt, dass in dem Fall, in dem keine Permutation benötigt wird, das Produkt der Elementarmatrizen, die benötigt werden, um eine quadratische Matrix auf (obere) Dreiecksgestalt zu bringen, eine untere Dreiecksgestalt hat. Dies führt zum Begriff der LU-Zerlegung einer Matrix! (wobei L für "lower triangular" und U für "upper triangular" matrix steht). Entsprechender MATLAB Befehl ist:

```
[L,U] = lu(A);
```

Ziel der Prozedur ist es, die gegebenen (erweiterte) Systemmatrix in (obere) Dreiecksgestalt zu bringen, *denn eine quadratische, obere Dreiecksmatrix* mit nicht-verschwindenden Diagonalelementen (nach Multiplikation mit deren Kehrwerten dann also eine obere Dreiecksmatrix mit lauter Einsen in der Hauptdiagonale!) sind invertierbar. Diese im Prinzip aus der Schule bekannte Beobachtung wollen wir in ein Lemma packen.

uppertriang1

Lemma 5. *(Invertierbarkeit von Dreiecksmatrizen) Es sei $\boldsymbol{A}$ eine $n \times n$-Matrix mit $a_{j,k} = 0$ für $k \leq j$, und $a_{j,j} \neq 0$ für $1 \leq j \leq n$. Dann ist $\boldsymbol{A}$ invertierbar!*

Proof. Diese Form eines linearen Gleichungssytems ist ja bekanntlich deshalb so erstrebenswert, weil man durch einfache Rücksubstitution, von unten beginnend (d.h. von der letzten Variablen) für jede mögliche rechte Seite $\vec{\mathbf{b}}$ auf eindeutige, ja sogar *konstruktive* und algorithmische effiziente Art eine Lösung $\vec{\mathbf{x}}$ finden kann. Damit ist klar, dass die zugehörige lineare Abbildung $T : \vec{\mathbf{x}} \mapsto \boldsymbol{A} * \vec{\mathbf{x}}$ nicht nur linear sondern auch bijektiv ist, somit auch eine Umkehrabbildung hat, deren Matrix natürlich $\boldsymbol{A}^{-1}$ ist (vgl. Invertierbarkeitskriterium).

Nebenbei bemerkt ist die Inverse einer oberen bzw. unteren Dreiecksmatrix wieder von derselben Bauart! □

Dazu gibt es (mindestens) zwei naheliegende Varianten.

1. Man bringe das Gleichungssytem auf (obere) Dreiecksgestalt. Dann kann man - von unten beginnend - durch *Rückwärts- Substitution* das Gleichungssystem von "hinten" aufrollen.

2. Noch besser ist es, wenn die Matrix in einer Form kommt, wo auf der linken Seite jede Variable genau einmal *alleine* vorkommt, d.h. das Gleichtungssyteme auf eine Form der Gestalt

   $$2y = 6; 3x = 7; z = 3$$

   kommt. Da man ja Zeilen untereinander vertauschen kann (die Reihenfolge der Gleichungen spielt keine Rolle) und auch mit den Zahlen $1/2, 1/3, 1$ multiplizieren kann, ohne das Ergebnis zu aendern, bedeutet das aber gleichzeitig, dass man die Matrix auf die Form einer Einheitsmatrix bringen kann.

Proof. Nennen wir die Originalmatrix $\boldsymbol{A}$ sei $\boldsymbol{B}$ eine gleich grosse Matrix, wobei in jeder Zeile von $\boldsymbol{B}$ eine passende Linear-Kombination der Zeilenvektoren von $\boldsymbol{A}$ steht. Wir wollen *zeigen*, dass wir $\boldsymbol{B}$ aus $\boldsymbol{A}$ durch einfache Anwendung der 3 Typen von elementaren Zeilenoperationen (die wir f.d. Gauss-Elimination verwenden dürfen) in einer endlichen Anzahl von Schritten konvertiert werden kann.

Fangen wir mit der letzen Zeile von $\boldsymbol{B}$ an. Nach Voraussetzung ist es eine Linearkombination der Zeilen von $\boldsymbol{A}$. Es ist klar, dass wir nur die letzte Zeile von $\boldsymbol{A}$ entsprechend multiplizieren müssen, um durch Aufaddieren der früheren Zeilen von $\boldsymbol{A}$ die letzte Zeile von $\boldsymbol{B}$ zu bekommen.

Um den Rest des Beweises nicht zu wortreich zu gestalten, wollen wir folgende Bezeichnungen einführen: Wir schreiben $\vec{\mathbf{w}}_1 \cdots \vec{\mathbf{w}}_n$ für die Eintragungen der ursprünglichen Vektoren und $\vec{\mathbf{u}}_1 \cdots \vec{\mathbf{u}}_n$ fuer die von $\boldsymbol{B}$.

Dann haben wir bisher gezeigt, dass wir $\vec{\mathbf{w}}_n$ durch $\vec{\mathbf{u}}_n$ ersetzen können. Gehen wir nun zur vorletzten Zeile. Dann können wir dort die gewünschte Linearkombination herstellen (nämlich den Vektor $\vec{\mathbf{u}}_{n-1}$) indem wir zu (einem geeigneten Vielfachen von) $\vec{\mathbf{w}}_{n-1}$ die entsprechenden Vielfachen der darüberliegenden (noch in ursprünglichen Form vorliegenden) Zeilen hinzuaddieren, aber ebenso ein Vielfaches des (nicht mehr vorhandenen) ursprünglichen letzen Zeile. Entweder ist der skalare Faktor Null (d.h. die Zeile kommt nicht wirklich vor, und alles ist erledigt), oder er ist nicht gleich Null, aber dann können wir ihn durch passende Multiplikation auf 1 bringen, also nehmen wir an, dass in der Linearkombination $\vec{\mathbf{w}}_n$ vorkommt. Um $\vec{\mathbf{w}}_n$ hinzuzuzählen, müssen wir nur $\vec{\mathbf{w}}_n$ durch $\vec{\mathbf{u}}_n$ ausdrücken, aber diese Differenz ist nur eine Linearkombination der früheren Spalten. Wir können also auch in der vorletzten Zeile zuerste die richtige Vielfachheit der beiden letzen (ursprünglichen) Zeilenvektoren herstellen, und dann - unter Berüchsichtigung der mitgebrachten Beiträge der früheren (noch im ursprünglichen Format vorliegenden) Zeilen entsprechend korrigieren.

Dieselbe Überlegung kann man nun rekursiv auf immer höhere Zeilen anwenden, bis man oben gelandet ist. □

Corollary 1. *Die Menge der Linear-Kombinationen von Zeilenvektoren ist genau dann der ganze $\mathbb{R}^m$ wenn es möglich ist, die Einheitmatrix (mit Einsen in der Hauptdiagonale und Nullen sonst) durch Gauss- Elimination herzustellen.*

Proof. Es sind zwei Implikationen zu beweisen.

Einerseits sind im Falle, dass der Zeilenraum der ganze $\mathbb{R}^m$ ist, alle Einheitsvektoren $\vec{\mathbf{e}}_k, 1 \leq k \leq m$ im Zeilenraum, und somit aufgrund des obigen Lemmas darstellbar.

Umgekehrt kann man leicht schliessen, dass jeder Vektor $\vec{\mathbf{x}} \in \mathbb{R}^m$ sich als $\sum_{k=1}^{m} x_k \vec{\mathbf{e}}_k$ darstellen lässt. Sind also die Einheitsvektoren als Linearkombinationen der ursprünglichen Zeilenvektoren darstellbar, dann ist auch jeder der Vektoren $\vec{\mathbf{e}} \in \mathbb{R}^m$ als Linearkombination (von Linearkombinationen, aber das spielt wegen des Lemma (7) keine Rolle) darstellbar, somit ist die Menge der Linearkombinationen der Zeilen (oft auch der **Zeilenraum der Matrix $\boldsymbol{A}$** genannt) □

lincombcomb

ausszeilpres

Lemma 6. *1. Jede der Zeilenoperationen, die im Laufe der Gauss-Elimination durchgeführt werden ist umkehrbar;*

2. *Die Gausselimination bewahrt den Zeilenraum der Matrix;*

3. *Jeder der drei erlaubten Schritte im Rahmen der Gauss-Eliminationkann durch Links-Multiplikation mit einer Elementar-Matrix realisiert werden, welche durch Anwendung des elementaren Schrittes (z.B. Vertauschen von Zeilen, etc.) auf die Einheitsmatrix vom Format $m \times m$ entsteht.*

Proof. Es genügt, für jede einzelen prototypische Operation die Aussage zu verifizeren. Vertauscht man Zeilen, so kann man das durch Nochmaliges Vertauschen derselben Zeilen wieder rückgängig machen. Multipliziert man mit einer von Null verschiedenen Zahl $\lambda \neq 0$ dann kann man durch Multiplikation mit dem Kehrwert $1/\lambda$ die Operation wieder rückgängig machen. □

lincombcomb **Lemma 7.** *[Linearkombinationen von Linearkombinationen] Hat man eine n Linear-Kombination von m Vektoren in einem Vektorraum, und bildet man von diesen neuerlich eine Anzahl von k Linearkombinationen, so können auch diese (neuen) Linear-Kombinationen als Linear-Kombinationen der ursprünglichen Vektoren geschrieben werden.*

Situation: Hat man eine $m \times n$ Matrix $\boldsymbol{A}$ und eine $k \times m$ Matrix $\boldsymbol{B}$, dann ist $\mathbf{C} := \boldsymbol{B} * \boldsymbol{A}$ eine $k \times n$ Matrix.
In Worten: Sei $\boldsymbol{A}$ die Matrix von Koeffizienten vom Format $m \times n$, die beschreibt wie aus einer Kollektion von m Vektoren $\mathbf{v}_1 \cdots \mathbf{v}_m$ in $\boldsymbol{V}$ n Linearkombinationen der Form

$$\mathbf{u}_k := \sum_{j=1}^{m} a_{j,k}\mathbf{v}_j, \quad 1 \leq k \leq n$$

bilden. Weiters sei $\boldsymbol{B}$ die Koeffizientenmatrix, die beschreibt, wie man aus diesen Vektoren $\mathbf{u}_1, \cdots \mathbf{u}_n$ eine weitere Kollektion von k Linearkombinationen der Form

$$\mathbf{w}_r := \sum_{l=1}^{n} b_{l,r}\mathbf{u}_l, \quad 1 \leq r \leq k,$$

bildet. Dann können auch die (neuen) Vektoren $(\mathbf{w}_r)_{r=1}^{p}$ als Linearkombinationen der ursprünglichen Vektoren $\mathbf{v}_1, \cdot, \mathbf{v}_m$ geschrieben werden, wobei die entsprechenden Koeffizienten genau diejenigen der Produktmatrix sind:

$$\mathbf{w}_r := \sum_{s=1}^{m} c_{s,r}\mathbf{v}_s, \quad 1 \leq r \leq p. \tag{34}$$ classdefmat

Der Beweis, der auch als **eine** der wesentlichen Rechtfertigungen für die (auf den ersten Blick vielleicht eigenartig anmutende) Form der Matrix-Multiplikation gibt, nämlich die **Definition** durch:

defmatrmult **Definition 8.** [MATRIX-Multiplikation]
Hat man eine $m \times n$ Matrix $\boldsymbol{A}$ und eine $p \times m$ Matrix $\boldsymbol{B}$, dann ist $\mathbf{C} := \boldsymbol{B} * \boldsymbol{A}$ die $p \times n$ Matrix, deren Eintragungen gegeben sind durch:

$$c_{r,s} = \sum_{l=1}^{m} b_{r,l}a_{l,s}\ , \quad 1 \leq r \leq p,\ 1 \leq s \leq n.$$

Das bedeutet, dass die Matrix Multiplikation nichts anderes ist als das Loslassen der links stehenden Matrix (in unserer Definition ist das $\boldsymbol{B}$) auf die einzelnen Spalten von $\boldsymbol{A}$, d.h.

$$\boldsymbol{B} * \boldsymbol{A} = [\boldsymbol{B} * \mathbf{a}_1, \dots \boldsymbol{B} * \mathbf{a}_n]; \tag{35}$$

matrmultcols

Man kann im Falle von $\mathbb{K} = \mathbb{R}$ die Eintragungen auch als Skalarprodukte zwischen den Zeilen von $\boldsymbol{B}$ und den Spalten von $\boldsymbol{A}$ interpretieren!

7 Matrix-Vektor Multiplikation und Matrix-Matrix

Zunächst sehe ich die *Berechnung der linken Seite* eines linearen Gleichungssystems aus den beiden Teilen, nämlich der System-Matrix $\boldsymbol{A}$ und dem (potentiellen) Lösungsvektor $\vec{\mathbf{x}} = [x_1, \cdots, x_n]$ als Motivation dafür an, wie die Matrix-Vektor-Multiplikaton einer $m \times n$-Matrix $\boldsymbol{A}$ mit einem Spaltenvektor der Länge (= Höhe n) realisiert wird: $\mathbf{y} = \boldsymbol{A} * \mathbf{x}$ definiert den Vektor, dessen Koordinaten durch folgende Gleichung gegeben sind, und zwar für $1 \leq j \leq m$:

$$y_j = a_{j,1}x_1 + \cdots + a_{j,n}x_n = \sum_{k=1}^{n} a_{j,k}x_k.$$

Die Wirkung einer $m \times n$-Matrix $\boldsymbol{A}$ auf eine $n \times k$-Matrix $\boldsymbol{B}$ wird über die einzelnen Spalten von $\boldsymbol{B}$ erklärt, d.h. man betrachtet die $n \times k$ -Matrix als Kollektion von Spaltenvektoren $[\vec{\mathbf{b}}_1, \ldots, \vec{\mathbf{b}}_n]$ und definiert $\boldsymbol{A} * \boldsymbol{B}$ einfach als die Kollektion von Spaltenvektoren der Form

$$[\boldsymbol{A} * \vec{\mathbf{b}}_1, \ldots, \boldsymbol{A} * \vec{\mathbf{b}}_n]$$

Die einzelnen Vektoren $\boldsymbol{A} * \vec{\mathbf{b}}_r, 1 \leq r \leq k$ sind wohldefinierte Vektoren im $\mathbb{K}^m$, also bekommt man eine $m \times k$ (!changed/corrected) Matrix!

Die Format-Überlegungen können im sogenannten (hgfei) *Domino- Prinzip* beschrieben werden. Das Format der Produkt Matrix $\mathbf{C} = \boldsymbol{A} * \boldsymbol{B}$ ist $[m, k]$ (!), weil $[m, n][n, k]$ im Domino-Spiel den (offenen) Anfang m und das Ende k aufweist (das ist das Format der Produkt-Matrix).

Theorem 3. *Die so eingeführte Matrix-Vektor Multiplikation (Zeilen der Matrix $\boldsymbol{A}$ werden im wesentlichen im Sinne des Skalarproduktes mit dem Spaltenvektor $\mathbf{x}$ zusammengeführt, und man bekommt im output-vector genausoviele Koordinaten wie man Zeilen in $\boldsymbol{A}$ hat) ist gleich einer spaltenweisen Sichtweise, wobei $\boldsymbol{A} = [\mathbf{a}_1; \ldots; \mathbf{a}_n]$ sei:*

$$\boldsymbol{A} * \mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{a}_k \tag{36}$$

matv-sp

Einige einfache Beobachtungen

Die Matrix-Vektor und daher die Matrix-Matrix Multiplikation sind beispielsweise distributiv (in beiden Faktoren), d.h.

$$(\boldsymbol{A}_1 + \boldsymbol{A}_2) * \mathbf{x} = \boldsymbol{A}_1 * \mathbf{x} + \boldsymbol{A}_2 * \mathbf{x};$$

$$\boldsymbol{A} * (\mathbf{x}_1 + \mathbf{x}_2) = \boldsymbol{A} * \mathbf{x}_1 + \boldsymbol{A} * \mathbf{x}_2.$$

Wenn die Formate passen, ist die Matrix-Multiplikation auch *assoziativ*, aber das ist entweder mühselig (buchhalterisch) oder später zu diskutieren.

Andererseits ist (nur für quadratische Matrizen sinnvoll) die Matrix-Multiplikation *nicht* kommutativ, d.h. man hat sehr oft $\boldsymbol{A} * \boldsymbol{B} \neq \boldsymbol{B} * \boldsymbol{A}$!, selbst für $\boldsymbol{B} = \boldsymbol{A}^t$: [28]

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} * \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix} \quad , \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} * \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix}$$

[28] Falls $\boldsymbol{A} * \boldsymbol{A}' = \boldsymbol{A}' * \boldsymbol{A}$ gilt, heißt $\boldsymbol{A}$ *normal*. Der Begriff spielt aber erst später eine Rolle.

8 Quadratische und Kubische Polynomfunktionen

In diesem Abschnitt soll das Rechnen mit Polynomen, sauberer gesprochen mit Polynomfunktionen auf $\mathbb{R}$ bzw. über $\mathbb{C}$ behandelt werden.

Einstiegsbeispiel: Man wähle die (rellen) Parameter λ und μ so, dass die Linearkombination $r(t) := \lambda p(t) + \mu q(t)$, mit

$$p(t) = 2t - 5t^3 \quad , \quad q(t) = 7t - t^3$$

die beiden folgnden Gleichungen erfl̈lt:

$$r(1) = 1 \quad , \quad r'(0) = -5$$

Das Beispiel soll zeigen, dass eine solche Aufgabe nach ein paar kleinen Umformungen zu einem kleinen 2×2-Gleichungssystem führt, das einfach gelöst werden kann. Andererseits benutzt man, dass man Linearkombinationen von Polynomfunktionen bilden kann, ohne die Menge der quadratischen Polynomfunktionen zu verlassen, und dass auch z.B. das Ableiten nicht aus den Polynomen nicht rauskommt (und den Grad um 1 reduziert!)[29].

Ebenso interessant ist es festzustellen, dass auch die Verschiebung einer Polynomfunktion wieder zu einem Polynom gleichen Grades führt, d.h. mit $p(t)$ ist auch $p(t-z)$ für jedes $z \in \mathbb{C}$ oder $\mathbb{R}$ ebenfalls ein Polynom (man muss nur $(t-z)^k$ als Polynom schreiben!).

Wohl am besten: Sei $p(t) = a_0 + a_1 t + a_2 t^2$ eine quadratisches Polynom, dann ist auch $q(t) = p(t-s)$ für jedes $s \in \mathbb{R}$ ein Polynom vom gleichen Grad. Gleiches gilt auch für Polynome von beliebigem festen Grad und auch über $\mathbb{C}$.

Übungsaufgaben: Gegeben 4 Stellen (allgemein: Anzahl verschiedener Stellen, mit der Zahl der Werte = Ordnung des Polynoms!) und 4 reelle Datenwerte, sagen wir $\vec{\mathbf{d}} = [d_1, \ldots, d_4]$ Man zeige, dass es ein eindeutig bestimmtes kubisches (etc...) Polynom $p(t)$ gibt, mit $p(t_i) = d_i \quad \text{for} \quad i = 1, ...4$.

Das kann/sollte man durchaus mit Hilfe der *Lagrange Interpolationsfunktionen* machen. Vielleicht sogar noch besser anhand quadratischer Funktionen diskutieren!

vgl. G. Strang, Beispiel mit der Vandermonde-Matrix für die Abtaststellen $a, b, c \in \mathbb{R}$ (Buch, ca. Seite 254, siehe Index), natürlich ohne Determinanten zu verwenden!

Die *Invertierbarkeit* einer allgemeinen Vandermonde-Matrix, welch zu der Punktfolge $\{z_1, \cdots, z_n\}$ gehört, ist genau dann gegeben, wenn die Punkte paarweise verschieden sind, weil dann die sog. *Vandermode Determinante*, die sich als

$$det(\boldsymbol{V}) = \prod_{1 \le j < k \le n} (z_j - z_k)$$

schreiben lässt, von Null verschieden ist![30]

[29] Bemerkung: Obwohl die Polynome $p(t)$ und $q(t)$ in $\mathscr{P}_3(\mathbb{R})$ liegen, bilden sie nur einen zweidimensionalen Teilraum, und deshalb ist es möglich, mit nur zwei Angaben die Lösung des Gleichungssystems zu berechnen. [Begriffe werden alle noch erklärt]

[30] Weil wir in unserem Aufbau den Umgang mit Determinanten (vor allem für allgemeine $n \times n$ Matrizen) aber erst später erlernen werden, ist das im Moment nur eine Erwähnung.

8.1 Beschreibung von linearen Operatoren auf $\mathscr{P}_n(\mathbb{R})$

```
a = rand(5,1);  % Zufallskoeffizienten f. Polynom p(t), Grad 4
bas1 =  0:4     % Auswertung an den Stellen 0,1,2,3,4
bas1 =      0     1     2     3     4
TR = transmat(4)    % Matrix f.d. Verschiebunsoperator
                    % T_{_1}:  p(t)  >>   p(t+1),
                    % verschiebt den Graphen von p(t) nach links!
% TRANSMAT.M ist ein selbst geschriebenes MATLAB file ($>$ App.)
% MATLAB/OCTAVE verwenden die Basis der Monome in fallender
% Ordnung, also  B = { t^4,t^3,t^2,t^1,t^0 }
% die folgende Matrix enthaelt in den jeweiligen Spalten
% die BILDER!! von diesen Basis-Vektor unter der linearen Abb.
% p(t) >> q(t) := p(t-1)
% Die Werte fuer q(t) sind dieselben Werte wie die von p(t)
% allerdings um eine Einheit weiter links, d.h. der Graph von
% q(t) ist der Graph von p(t), um eins nach RECHTS verschoben.
TR =
     1     0     0     0     0
    -4     1     0     0     0
     6    -3     1     0     0
    -4     3    -2     1     0
     1    -1     1    -1     1
% Man beachte das Auftreten des Pascal'schen Dreiecks!
% in dieser Matrix (mit wechseldem Vorzeichen) (t-1)^k

y1 = polyval(a,bas1-1)
    % Werte von p(t) an den gegebenen Stellen, um 1 = eins
    % nach links verschoben, d.h. an -1,0,1,2,3
y1 =     1.0021    0.0971    1.1581   12.8486   60.7772

y2 = polyval(transmat(4)*a, bas1)
%  Auswertung des verschobenen Polynoms!!
y2 =     1.0021    0.0971    1.1581   12.8486   60.7772
%  man koennte also auch definieren: b = transmat(4)*a,
%  dann hat man die Koeffizienten von q(t) explizit
%  und  y2= polyvl(b, bas1);
%  ergaebe die Auswertung von q(t) an den Stellen 0,1,2,3,4.

y3 = polyval(inv(transmat(4))*a, bas1)
  % inverse Matrix, d.h. Rechts-Shift  % p(t) >>  p(t-1)
y3 =     1.1581   12.8486   60.7772  187.4977  452.5090
y4 = polyval(a,bas1+1)
  % Auswertung von p(t) an den Stellen: 1,2,3,4,5
y4 =     1.1581   12.8486   60.7772  187.4977  452.5090
```

8.2 Operatoren auf $\mathscr{P}_3(\mathbb{R})$

Die hier vorgestellten Operationen sind nur examplarisch zu verstehen. Wir verwenden $\mathscr{P}_3(\mathbb{R})$ weil dies die Polynomfunktionen sind die aus der Schule am besten bekannt sind (und sie nicht so ßchlichtßind wie "lineare" oder quadratische Funktionen).

Für jedes $\alpha \neq 0$ ist die Abbildung $p(t) \mapsto q(t) := p(\alpha t)$ eine (!) lineare Abbildung, und ihre Matrix (Beispiel $\alpha = 2$) ist

```
diag(2.^(3: -1: 0))
```

was ausgeschrieben ergibt:

$$\begin{pmatrix} 8 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{37}$$

dilmat2

Ein anderes Beispiel ist das Differenzieren, hier als Abbildung von $\mathscr{P}_3(\mathbb{R})$ nach $\mathscr{P}_3(\mathbb{R})$ beschrieben, also gibt es eine 4×4-Matrix, konkreter (OCTAVE!):

```
D = zeros(4); D(2,1) = 3; D(3,2) =2; D(4,3) =1
```

was ergibt (jede Spalte beschreibt ein differenziertes Monom, und auf beiden Seiten wird die Monimialbasis $\mathscr{M}^3 := \{t^3, t^2, t^1, t^0 = 1\}$ nach MATLAB-OCTAVE Konvention verwendet!)

$$\mathbf{DM} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{38}$$

$$\mathbf{DM^2} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \end{pmatrix} \tag{39}$$

$$\mathbf{DM}^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \end{pmatrix} \tag{40}$$

und natürlich $\mathbf{DM}^4 = \mathbf{0}$ (die Null-Matrix). Eine solche Matrix heisst *nilpotent* (ein Potenz ist Null!).

Es ist eine gute Übung, dieselbe Abbildung bzgl. der "aufsteigenden" Monomialbasis $\mathscr{M}_3 := \{1, t, t^2, t^3\}$ zu behandeln. Klarerweise (fast) bleiben alle Eigenschaften von Matrizen erhalten.

FRAGE: Kann man auch behaupten, dass die reduzierten Zeilenstufenform aller möglichen Matrizen, die eine *gegebene lineare Abbildung* T bzgl. verschiedener Basen darstellt, alle gleich sind? Kann man das experimentell feststellen (mit Hilfe von MATLAB/OCTAVE!?). Wenn ja, wie?

Integration von Polynomfunktionen ist eine lineare Prozedur, weil (für festes $a < b$ in $\mathbb{R}$) die Abbildung $p(t) \to \int_a^b p(t)dt$ *linear*! ist (das lernt man in der Analysis Vorlesung, bitte selbst davon überzeugen), auf $\mathscr{P}_n(\mathbb{R})$ (für jedes $n \in \mathbb{N}$).

Also hat auch diese Abbildung bei fester Basis eine Matrixdarstellung, und zwar offenbar in Form eines Zeilenvektors in $\mathbb{R}^{k+1}$.

Zur Illustration wollen wir den Fall $n = 3$ betrachten, mit der *aufsteigenden* Monomial-Basis $\mathscr{M}_3 := \{1, t, t^2, t^3\}$ und den Fall $[a, b] = [0, 1]$.

Die bekannten Integrationsregeln ergeben dann

$$\int_0^1 t^j = 1/(j+1),$$

somit ergibt sich der Vektor $\vec{\mathbf{h}} = [1, 1/2, 1/3, 1/4]$. Die Darstellungs-Aussage für diese lineare Abbildung konkretisiert sich dann zu folgendem Prinzip:

Ist ein kubisches Polynom von der Form

$$p(t) = a + bt + ct^2 + dt^3$$

gegeben, so ist das Integral üeber $[0, 1]$ gerade das Matrix-Produkt des Zeilenvektors $\vec{\mathbf{h}}$ mit dem Spaltenvektor $[a; b; c; d]$ (!man beachte die Reihenfolge der Monome, die passen muss), was aber auch als Skalarprodukt der beiden Quadrupel gesehen werden kann, oder konkreter, es gilt:

$$\int_0^1 p(t)dt = a + b/2 + c/3 + d/4,$$

ein Resultat das man natürlich auch durch die übliche Methode (aus der Schule bekannt) zustande gekommen wäre.

In Wirklichkeit ist nicht nur das Ergebnis gleich, sondern sogar die Methode: Die Integration erfolgt auch bei der Schulmethode "Term für Term", und Faktoren (wie a, b, c, d) werden herausgehoben, d.h. es wird genau die Linearität des Integrals zuhilfe genommen.

Ein weiteres Beispiel einer Matrix die im Zusammenhang mit kubischen Polynomen interessant sein kann, ergibt sich aus der Fragestellung der Hermite-Interpolation. Die typische Fragestellung lautet: Gegeben die Werte $p(0), p'(0), p(1), p'(1)$, sagen wir gleich $[1, 2, 3, 4]$. Gibt es dann auch ein (?eindeutig bestimmtest) *kubisches Polynom* das diesen Vorgaben entspricht. Aufstellen des entsprechenden 4×4 Gleichungssystems in den Variablen a, b, c, d und Anwendung der Gauss-Elimination kann die Antwort auf diese Frage liefern.

9 Lagrange Interpolation und Vandermonde Matrix

lagrintpol1

Lemma 8. *Die Existenz der Lagrange Interpolationspolynome (für beliebige endliche Mengen von komplexen Zahlen $Z = \{z_1, \ldots, z_{k+1}\}$!) impliziert,dass die Einschränkung von $\mathscr{P}_k(\mathbb{C})$ auf jede beliebige Teilmenge von $\mathbb{C}$ welche mindestens $k+1$ verschiedene Punkte enthält, die Monome (bis zum Grad k; bzw. eben deren Einschränkung) als Basis enthält.*

Das Argument ist das folgend: Es ist möglich, eine Anzahl von $k+1$ Linear-Kombinationen in $\mathscr{P}_k(\mathbb{C})$ zu bilden, welche *Lagrange Interpolations-Polynome* $L_r(t), 1 \leq r \leq k+1$, heissen für die welche die folgende Kronecker-Bedingung erfüllen:

$$L_r(z_j) = \delta_{r,j} = \begin{cases} 1 & \text{if } r = j \\ 0 & \text{if } r \neq j \end{cases}$$

Die Koeffizienten dieser $(k+1) \times (k+1)$-Matrix können in eine Matrix $\boldsymbol{B}$ geschrieben werden.

Die Auswertung dieser Koeffizienten an den Stellen $Z = \{z_1, \ldots, z_{k+1}\}$ ergibt gerade die Einheitsmatrix. D.h. die Menge $M = \{L_r, 1 \leq r \leq k = 1\}$ wird durch die lineare Abbildung $p(t) \mapsto [p(z_1); \ldots; p(z_{k+1})]$ auf die Menge der Einheitsvektoren abgebildet. Da diese linear unabhängig ist, gilt dasselbe für die Menge M, d.h. für jedes ZA wie oben beschrieben, das aus $k+1$ verschiedenen Punkten besteht, ist die Menge der zugehörigen Lagrange Interpolatoren[31] ein linear unabhängige Menge mit $k+1 = \dim(\mathscr{P}_k(\mathbb{R}))$ Elementen. Also sind sie eine Basis.

Da für jedes Polynom $q(t) \in \mathscr{P}_k(\mathbb{R})$ gilt also

$$q(t) = \sum_{r=1}^{k+1} q(z_r) L_r(t)$$

und die Koeffizienten in dieser Lagrange-Basis sind genau die Abtastwerte der Polynomfunktion, d.h. die Folge $(q(z_r))_{r=1}^{k+1}$.

Insbesondere stimmen zwei Polynome $q_1(t)$ und $q_2(t)$ aus $\mathscr{P}_k(\mathbb{R})$ (ebenso $\mathscr{P}_k(\mathbb{C})$) überein (d.h. haben die gleichen Koeffizienten in der üblichen monomialen Basis) genau dann wenn sie an mindestens $k+1$ Stellen übereinstimmen!

[31] Details sind in WIKI und in fast jedem Buch zu finden, werden auch in der Vorlesung gemacht.

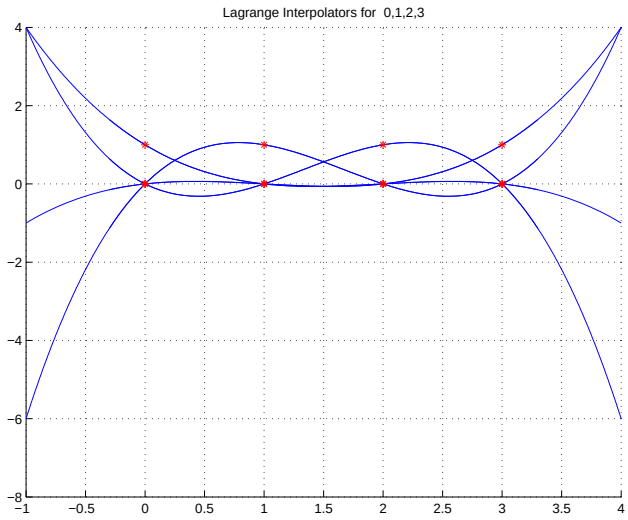


Abbildung 1:

9.1 Weitere Beispiele: Hermite-Interpolation

Wir wollen diese Frage nun mit Hilfe von Matrizenrechnung beantworten. Die Behauptung lautet positiv formuliert: *Jedes kubische Polynom* $p(t) \in \mathscr{P}_2(\mathbb{R})$ *ist eindeutig aus den Werten* $[p(0), p'(0), p(1), p'(1)]$ *zu bestimmen.*

Um dies zu zeigen, genügt es, dass die zugehörige lineare Abbildung von $\mathscr{P}_3(\mathbb{R})$, versehen mit der Standard-Basis, d.h. der Monomialbasis $\{1, t, t^2, t^3\}$ nach $\mathbb{R}^4$ (Standardbasis) durch eine *invertierbare Matrix* $\boldsymbol{A}$ beschrieben wird:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix} \tag{41}$$

Die inverse Matrix ist dann :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -3 & -2 & 3 & -1 \\ 2 & 1 & -2 & 1 \end{pmatrix} \tag{42}$$

Die Spalten dieser Matrix enthalten die Koeffizienten, welche die Fundamentallösungen ergeben, d.h. diejenigen kubischen Polynome, für die man Einheitsvektoren bekommt, also beispielsweise $p_3(t) = 3t^2 - 2t^3$ erfüllt $p_3(0) = 1, p_3'(0) = 0, p_3(t) = 1$ und $p_3'(1) = 0$.

Beispiel von entsprechendem MATLAB Code für 4 Punkte, um zu demonstrieren, dass man solche Beispiele (mit ein wenig Geschick) leicht *programmieren* und vor allem verwenden kann, ohne von Hand die entsprechende 8×8-Matrix invertieren zu muessen. Bei der Eingabe wird die Matrix *zeilenweise!* von unten aufgebaut (und nicht spaltenweise, wie sonst üblich), und aus diesem Grund liefern auch die *Zeilen* der inversen Matrix die Koeffizienten der Koeffizienten. Wir arbeiten mit der Basis $\mathscr{M}^8 = \{t^7, t^6, \cdots, t, 1\}$ und geben die Daten in der Form von Zeilenvektoren ein:

$$[p(-1), p(0), p(1), p(2), p'(-1), p'(0), p'(1), p'(2)]$$

```
HM4 = zeros(8); HM4(8,:) = [ones(1,4), zeros(1,4)];
for k=2:8; HM4(9-k,:)=[(-1:2).^(k-1), (k-1)* HM4( 10-k, 1:4)]; end;
disp(HM4); bas2 = linspace(-1.2,2.2,501); HM4I = inv(HM4);
for jj=1:8;  POLS4(:,jj) =  polyval(HM4I(jj,:),bas2);  end;  figfei;
plot(bas2, POLS4);   grid; % ax20; plotax
hold on;plot(-1:2,ones(1,4),'k*');plot(-1:2,zeros(1,4),'k*');hold off
title(' Hermite interpolation over  4 equidistant points'); shg
```

Der so erzeugte Plot der 8 Funktionen sieht folgdermaßen aus:

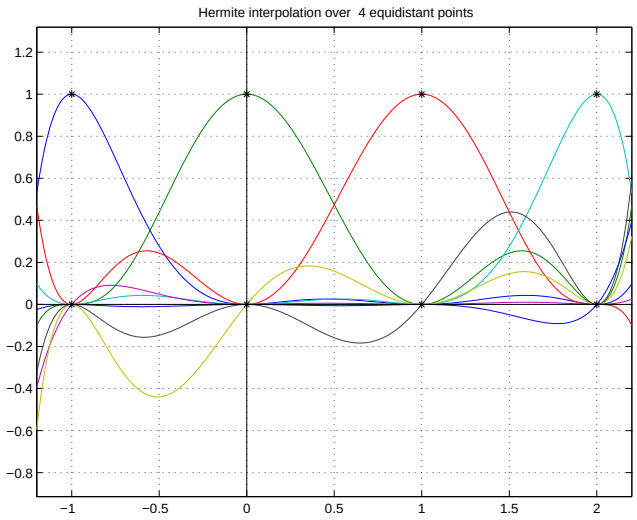


Abbildung 2:

9.1.1 Von Punktwerten zum Integral, ein Beispiel

Wie kann man aus den Werten eines quadratischen Polynomes $p(t)$ an den Stellen $0, 1/2, 1$ das Integral $\int_0^1 p(t)dt$ berechnen?

Wir betrachten es als das Problem, aus den Werten das Polynom $p(t) = a_1t^2+a_2t+a_3$ mit $\vec{\mathbf{a}} = [a_1, a_2, a_3] \in \mathbb{R}^3$ (d.h. $\vec{\mathbf{a}} = [p(t)]_{\mathscr{M}^2}$) aus dem Datenvektor $\vec{\mathbf{d}} := [p(0); p(1/2); p(1)] \in \mathbb{R}^3$ zu bestimmen.

Da alle Zusammenhänge offensichtlich linear bzgl. der Vektorraumstruktur sind (und jedes quadr. Polynom eindeutig durch drei Werte festgelegt ist) suchen wir also eine Zeilenvektor $\vec{\mathbf{u}}$ (bzw. eine 1×3-Matrix), sodass

$$\int_0^1 p(t)dt = \vec{\mathbf{u}} * \vec{\mathbf{f}}, \quad \forall p(t) \in \mathscr{P}_2(\mathbb{R}).$$

Dazu stellen wir die Vandermonde Matrix auf, die die Abbildung $S : \vec{\mathbf{a}} \mapsto \vec{\mathbf{d}} = [p(0), p(1/2), p(1)]$ von $\mathbb{R}^3$ nach $\mathbb{R}^3$ beschreibt ($\mathbf{d}$ steht für Datenvektor). Zuerst wertet man die Basis-Elemente $t^2, t, 1$ an den drei Stellen aus, und bekommt so

$$\mathbf{V} = \begin{pmatrix} 0.0000 & 0.0000 & 1.0000 \\ 0.2500 & 0.5000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 \end{pmatrix} \tag{43}$$

Die Abbildung vom Datenvektor $\vec{\mathbf{d}}$ zum Koeffizientenvektor $\vec{\mathbf{a}}$ ist durch die inverse Vandermonde-Matrix geben, d.h. $\vec{\mathbf{a}} = \mathbf{V}^{-1} * \mathbf{d}$.

Die Abbildung von den Polynomen zu deren Integral $\int_0^1 p(t)dt$ is ebenfalls eine lineare Abbildung, ist also durch eine 1×3-Matrix, d.h. einen Zeilenvektor $\vec{\mathbf{h}} \in \mathbb{R}^3$ gegeben, dessen Koordinaten einfach die Integrale der Basisvektoren sind: $\int_0^1 t^k dt = 1/(k+1), 0 \leq k \leq 2$, also ist die entsprechende Matrix (man beachte die Reihenfolge der Basis-Vektoren in MATLAB Konvention, nämlich $(t^2, t, 1)$: $\vec{\mathbf{h}} = [1/3, 1/2, 1]$. Somit bekommen wir das Integral in der Form

$$\vec{\mathbf{h}} * \vec{\mathbf{a}} = (\vec{\mathbf{h}} * \boldsymbol{V}^{-1}) * \vec{\mathbf{d}}.$$

Konkret ausgeführt hat man

$$\boldsymbol{V}^{-1} = \begin{pmatrix} 2 & -4 & 2 \\ -3 & 4 & -1 \\ 1 & 0 & 0 \end{pmatrix} \tag{44}$$

und somit ergibt sich der Wert

$$\vec{\mathbf{u}} = \begin{pmatrix} 0.8333 & -2.0000 & 1.5000 \end{pmatrix} = \begin{pmatrix} 5/6 & -2 & 3/2 \end{pmatrix}. \tag{45}$$

9.2 Substitution von Polynome in Polynomen

Gegeben sei eine (beispielsweise quadratische) Polynomfunktion $p(y)$, und eine *Substitution* etwa von der Form $y = x^2 + 2x + 1$. Dann ist auch $p(y) = q(x) \in \mathscr{P}_4(\mathbb{R})$.

Um die Matrix dieser (!offensichtlich? linearen) Abbildung $S : p(y) \mapsto q(x)$ zu beschreiben, verwenden wir die entsprechenden (absteigenden) monomialen Basen: $\mathscr{M}^n := \{t^n, t^{n-1}, \ldots, t, 1\}$, für $m = 2$ und $n = 4$, und bekommen so die 5×3-Matrix

$\boldsymbol{A} := [S]_{\mathscr{M}^4 \hookleftarrow \mathscr{M}^2}$ welche als dritte Spalte (wegen $S(1) = 1$) einfach den fünften Einheitsvektor $\vec{\mathbf{e}}_5 \in \mathbb{R}^5$ hat, weil 1 das fünfte Basis-Element on $\mathscr{M}^4$ ist. Andererseits ist klar dass $S(t^1) = y = x^2 + 2x + 3$ ist, also muss die zweite/mittlere Spalte von der Form $\mathbf{a}_2 = [0; 0; 1; 2; 3]$ sein (keine kubischen und höheren Termine). Nur $S(t^2)$ ist ein wenig aufwändiger, aber

$$y^2 = (x^2 + 2x + 3)^2 = x^4 + 4x^3 + 10x^2 + 12x + 9$$

impliziert, dass die erste Spalte von der Form

$\mathbf{a}_1 = [1; 4; 10; 12; 9]$ ist. Insgesamt gilt also:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 0 & 0 \\ 10 & 1 & 0 \\ 12 & 2 & 0 \\ 9 & 3 & 1 \end{pmatrix} \tag{46}$$

Es ist leicht einzusehen, dass diese drei Spalten linear unabängig sind (streicht man die Zeilen mit Nr.2 und Nr. 4 so ist der Rest eine invertierbare Dreicksmatrix, woraus sich die lineare Unabhängigkeit auch im $\mathbb{R}^5$ ergibt!), also bildet die Menge aller Polynomfunktion der Form $q(x) = p(y)$ mit $y = x^2 + 2x + 1$ einen linearen Teilraum (das Bild von $\mathscr{P}_2(\mathbb{R})$ unter S ist ein Teilraum).

Wenn also die Frage gestellt wird, ob ein gegebenes Polynom $r(x) \in \mathscr{M}^4$ von dieser Form ist (d.h. durch diese spez. Substitution aus einer quadratischen Polynomfunktion entstanden ist), ist diese *genau dann positiv zu beantworten* wenn die Koeffizenten von $r(x)$ (man könnte sie funktional auch als $[r(x)]_{\mathscr{M}^4}$ bezeichnen, oder rein pragmatisch als Quintupel $r_1, \cdots, r_5 \in \mathbb{R}$) im Bildbereich der Matrix-Multiplikation

$$\vec{\mathbf{v}} \mapsto \boldsymbol{A} * \vec{\mathbf{v}} = \sum_{k=1}^{3} v_k \mathbf{a}_k$$

liegt, oder mit anderen Worten (und konkreter), wenn das Gleichungssystem $\boldsymbol{A} * \vec{\mathbf{v}} = \vec{\mathbf{r}}$ konstistent ist, was prinzipiell (wie auch praktisch) mittels Gauss-Elimination überprüft werden kann.

9.3 Die diskrete Fourier Transformation:

Die diskrete Fourier-Transformation entspricht der linearen Abbildung von $\mathbb{C}^n$ nach $\mathbb{C}^n$, die man auf folgende Weise verbal beschreiben kann: Man bilde die Matrix, die den Übergang von den Koeffizienten (in $\mathbb{C}$) von $\mathscr{P}_{n-1}(\mathbb{R})$ auf die Werte diese Polynome an den n-ten Einheitswurzeln, beginnend by 1, im Uhrzeigersinn, realisiert.

In MATLAB ist diese lineare Abbildung in Form der Routine `FFT` realisiert, die sog. Fast Fourier Transform (schnelle FT). Um die Linearität zu testen, muss man nur die Verträglichkeit mit Linear-Kombinationen verifizieren (genaugenommen nur experimentell). Nimmt man einen beliebigen Spaltenvektor $\mathbf{c} \in \mathbb{C}^n$ und zwei beliebige Matrizen $\boldsymbol{A}$ und $\boldsymbol{B}$ vom Format $n \times r$ und $r \times s$ mit komplexen Koeffizienten, so bildet man mit $\boldsymbol{A} * \boldsymbol{B}$ s zufällige Linearkombinationen der r Spalten von $\boldsymbol{A}$ in $\mathbb{C}^n$.

Nun muss man die Konvention kennen, dass $\mathrm{fft}(XX)$ eine spaltenweise Anwendung der Fourier-Transformation realisiert. Wenn man also $\mathrm{fft}(\boldsymbol{A}) * \boldsymbol{B}$ bildet, so bildet man die Linearkombinationen der Bildvektoren, die im Falle der Linearität mit mit $\mathrm{fft}(\boldsymbol{A} * \boldsymbol{B})$ übereinstimmen sollte (was durch ein kleines numerisches Experiment bestätigt wird!).

Die zugehörige Basis kann also durch Anwenden der Routine auf die Kollektion von Einheitsvektoren im $\mathbb{C}^n$ bestimmt werden, also durch $\mathrm{fft}(\mathrm{Id}_n)$, realisiert als $F = \mathrm{fft}(eye(n))$.

Einzig und alleine die allgemeinen (historisch bedingten) Konventionen zeigen, dass diese Matrix nur bis auf die Reihenfolge der Spalten identisch mit der Vandermonde-Matrix zu der Folge von Punkten $\omega^0, \omega^1, \ldots \omega^{n-1}$ ist (weil diese die MATLAB-Ordnung der Monomial-Basis verwendet, d.h. $\{t^{n-1}, t^{n-2}, \ldots, t, 1\}$). Konkret sieht die FFT-Matrix für $n = 4$ so aus:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix} \tag{47}$$

Diese Matrix (auch für allgemeines n) gilt, dass sie symmetrisch ist, d.h. $F = F^t$, aber nicht $F = F'$ (transponiert), vielmehr ist $F' = conj(F)$ (elementweise Konjugation). Die Inverse Matrix ist als $conj(F)/n$, weil die Länge der Spalten (alle Spalten haben gleiche euklidische Länge) eben $= \sqrt{n}$ ist.

Man kann durchaus sagen, dass die FFT an der Basis des Bild-Kompressions- Verfahrens **JPEG** steht (Cosinus-Transformation), ebenso wie ein Teil des **MP3** Audio-Kompressions-Schemas ist (Kurz-Zeit Fourier Transformation, Spektrogramm), siehe auch WIKIPEDIA oder andere Quellen.

10 Linearkombinationen von Linearkombinationen

In der Vorlesung wird eine kleines Beispiel vorgerechnet, wie man durch Bilden einer Linearkombination von zwei Vektoren in einem allgemeinen Vektorraum, welche ihrerseits wieder von zwei anderen Vektoren aus ebenfalls zwei Vektoren sind, auch *direkt* als *Linearkombination* der ursprünglichen Vektoren beschrieben werden können. Kurz gesagt: Linearkombinationen von Linearkombination sind wieder Linearkombinationen.[32]

Die allgemeine Situation kann so beschrieben werden. Gehen wir davon aus, dass man zunächst n Elemente $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \boldsymbol{V}$ (in einem allgemeinen Vektorraum), aus denen m Linearkombinationen gebildet werden, die wir mit $\mathbf{v}_1, \ldots, \mathbf{v}_m$ bezeichnen wollen. Es bietet sich an, die Koeffizieten für diese Linearkombinationen mit $a_{k,j}$ zu bezeichnen, d.h. aus der Schreibweise

$$\mathbf{v}_k = a_{k,1}\mathbf{u}_1 + \cdots + a_{k,n}\mathbf{u}_n,\ 1 \leq k \leq m \tag{48}$$ lincombs1

eine kompakte Sammlung aller auftretenden Koeffizienten in Form einer (möglicherweise) rechteckigen Matrix zu machen, die wir natürlich mit $\boldsymbol{A}$ bezeichnen.

In ähnlicher Weise sei eine $h \times m$ Matrix $\boldsymbol{B}$ die Matrix aller Koeffizienten ("nicht vorhandene Koeffizienten sind einfach Null-Eintragungen in $\boldsymbol{B}$), mit denen weitere h Linearkombinationen $\mathbf{w}_1, \ldots, \mathbf{w}_h$ der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ gebildet werden.

Es gilt dann die einfache und wichtige Aussage und Formel.[33]

Auch die h neuen Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_h$ sind ebenfalls als Linear-Kombinationen der ursprünglichen Vektoren $\mathbf{u}_1, \ldots, \mathbf{u}_n$ darstellbar, wobei die zugehörige $h \times n$ Matrix C ist, welche durch Matrix- Multiplikation gegeben ist

$$\mathbf{C} = \boldsymbol{B} * \boldsymbol{A}.$$

Der Beweis dieser Aussage ist ein typisches Beispiel für eine logisch im Prinzip einfach zu durchschauende Überlegung, die man anhand eines konkreten Beispiels (eventuell unter Verwendung von MATLAB oder OCTAVE) leicht plausibel zu machen ist.

Gibt man einen sauberen, streng formalen Beweis, ist dies vielleicht weniger überzeugend, obwohl logisch betrachtet mehr wert als 100 interessante Beispiele.

In der VORLESUNG (siehe Mitschrift) wird daher ein Mittelweg gegangen

HINWEIS: Die Assoziativität der Matrix-Multiplikation ist von ähnlicher Natur. Iteriertes Substituieren hängt offenbar nicht von der Reihenfolge der Substitutionen ab! Trotzdem soll die Assoziativität der Matrix-Multiplikation (weil auch sehr wichtig) formal hergeleitet werden.

matmultassoc **Lemma 9.** *Seien $\boldsymbol{A}, \boldsymbol{B}, \mathbf{C}$ drei Matrizen über $\mathbb{K}$, vom Format $m \times n, n \times p$ bzw. $p \times q$. Dann gilt:*

$$(\boldsymbol{A} * \boldsymbol{B}) * \mathbf{C} = \boldsymbol{A} * (\boldsymbol{B} * \mathbf{C}). \tag{49}$$ matmulassocf

[32] Man stelle sich das ganz mit 2 Vektoren im $\mathbb{R}^3$ vor, die eine Ebene aufspannen (welche den Nullpunkt $(0, 0, 0$ enthält). Die zuerst gebildeten Linearkombinationen sind dann alle Vektoren in dieser Ebene, und jede weitere Linearkombination liefert dann nichts Neues, sondern eben wieder nur Linearkombinationen aus derselben Ebene!

[33] ... die noch eine andere Rechtfertigung für die zunächst willkürlich erscheinende Form der Matrix-Matrix-Multiplikation gibt!

In anderen Worten, das Symbol $\boldsymbol{A}*\boldsymbol{B}*\mathbf{C}$ is wohldefiniert, unabhängig davon, in welcher Reihenfolge die Multiplikation realisiert wird.[34]

Proof. 2010/11: TO BE GIVEN LATER, siehe Formel (71)! matr-assoc1 □

Kommentar: Rein logisch genügt es, sich auf die Darstellungsformel für lineare Abbildungen zu berufen, die (selbstverständliche) Gültigkeit der Assoziativität (es ist ja nur ein Aneinanderketten von Abbildung) für Abbbildungen[35] für Abbildungen (ganz allgemein) zu berufen.

Dennoch wollen wir nun die übliche, formale Beweisführung machen, und durch genau Überprüfung der Berechnung der Koeffizienten nachprüfen, dass die Assoziativität für allgemeine Matrix-Multiplikation gilt.

AB HIER EINSCHUB PER MÄRZ 2014:

Nun brauchen wir noch das Assoziativ-Gesetz der Matrix-Multiplikation (Details folgen hier noch, mit genauer Index-Buchhaltung! Elegantere, aber vielleicht nicht so leicht nachvollziehbare abstrakte Argumente dafür folgen noch später an anderer Stelle!).

Wir werden zwei bis drei Interpretationen der Situation geben, einerseits durch Bestimmung der Eintragungen in der Produkt-Matrix (unabhängig von der Klammernsetzung!), andererseits eine Deutung durch "Linearkombinationen von Linearkombinationen".

Proof. Wir haben zu zeigen, dass für drei Matrizen $\boldsymbol{A}, \boldsymbol{B}, \mathbf{C}$ welche aufgrund ihres Formates zusammenpassen (also eine $m \times n, n \times p$ bzw. eine $p \times q$ Matrix) gilt:

$$(\boldsymbol{A} * \boldsymbol{B}) * \mathbf{C} = \boldsymbol{A} * (\boldsymbol{B} * \mathbf{C}), \tag{50}$$ matr-assoc1

d.h. die Matrix-Multiplikation ist *assoziativ*.

Um dem Beweis eine klarere Struktur zu geben, ist es vorteilhaft, für die verschiedenen auftretenden Matrizen eigene Namen zu vergeben. Wir bezeichnen die linke Seite der Gleichung mit $\mathbf{U}$, und die rechte Seite mit $\mathbf{V}$, und setzen $\boldsymbol{G} = \boldsymbol{A} * \boldsymbol{B}$ und $\boldsymbol{H} = \boldsymbol{B} * \mathbf{C}$, d.h. wir haben zu zeigen, dass $\mathbf{U} = \boldsymbol{G} * \mathbf{C} == \boldsymbol{A} * \boldsymbol{H} = \mathbf{V}$ gilt[36]. Vom Format her (Domino-Prinzip!) ist alles in Ordnung, denn $\boldsymbol{G}$ ist eine $m \times p$ Matrix, die mit der $p \times q$ Matrix $\mathbf{C}$ multipliziert werden kann, ebenso passt es bei $\boldsymbol{A} * \boldsymbol{H}$, sodass sowohl $\mathbf{U}$ als auch $\mathbf{V}$ Matrizen vom Format $m \times q$ sind.

[34] Es ist interessant zu beobachten, dass die Reihenfolge der Multiplikation einen Einfluss auf den Rechenaufwand hat. Beispielsweise kann im Fall von grossen Matrizen die Realisierung von $\boldsymbol{A} * \boldsymbol{B} * \mathbf{x}$ aufwändiger, d.h. langsamer sein, als die von $\boldsymbol{A} * (\boldsymbol{B} * \mathbf{x})$, weil MATLAB als interpretierende Sprache zuerst $\boldsymbol{A} * \boldsymbol{B}$ realisiert. Aufwandsüberlegungen dieser Art spielen in der Numerischen Mathematik und für die Entwicklung von Algorithmen, etwa in der Numerischen Linearen Algebra eine große Rolle!

[35] Ob ich erkläre, dass ich von Wien über St. Pölten nach Linz, und von dort nach Salzburg gefahren bin, oder ob ich von erzähle, dass ich von Wien nach Salzburg gefahren bin, beschreibe ich doch in jedem Fall die Strecke Wien → St. Pölten → Salzburg →, also genau denselben Weg (bitte selbst ins Abstrakte zurückuebersetzen, also $h \circ (g \circ f) = (h \circ g) \circ f$).

[36] Matrizen sind genau dann gleich wenn alle Eintragungen gleich sind, oder - gleichwertig - wenn die zugeordneten linearen Abbildungen gleich sind! In MATLAB wird ergibt die Abfrage $>> A == B$ eine 0/1-Matrix: 1 an den Stellen (j,k) wo $a_{j,k} = b_{j,k}$ gilt, und 0 sonst. $\mathbf{U} == \mathbf{V}$ soll also nicht bedeuten: $\mathbf{U}$ wird mit $\mathbf{V}$ gleichgesetzt, d.h. die linke Seite wird durch die rechte definiert, wie bei Zuweisungen üblich, sonder es soll in Kurz-Schreibweise aussagen: ! die beiden SIND gleich!

Die Hauptproblem für den formalen (allgemeinen) Beweis ist eine *vernünftige Wahl der Indizierung*. Es ist naheliegend (damit auch die Summierungen über passende Index-Symbole erfolgen kann) die Eintragung der Matrizen mit $(a_{i,j}), (b_{j,k})$ bzw. $(c_{k,l})$ zu bezeichnen, woraus sich in natürlicher Weise die Index-Bezeichnungen $(g_{i,k}), (h_{j,l})$ bzw. $(u_{i,l})$ und $(v_{i,l})$ ergeben. Dabei laufen die entsprechenden Indices (konsistenterweise mit den gleichen Buchstaben bezeichnet, damit es "passt") über die Vielfachheiten die beim "Domino-Prinzip an den Stoßstellen auftreten", d.h. $1 \leq j \leq n, 1 \leq k \leq p$.

Einfach der *Definition des Matrix-Produktes* folgend, haben wir

$$g_{i,k} = \sum_{j=1}^{n} a_{i,j} b_{j,k} \quad , \quad h_{j,l} = \sum_{k=1}^{p} b_{j,k} c_{k,l}$$

bzw.

$$u_{i,l} = \sum_{k=1}^{p} g_{i,k} c_{k,l} \quad , \quad v_{i,l} = \sum_{j=1}^{n} a_{i,j} h_{j,l} \, ,$$

bzw. nach Einsetzen der ersten Gleichungen

$$u_{i,l} = \sum_{k=1}^{p} \sum_{j=1}^{n} a_{i,j} b_{j,k} c_{k,l} \qquad v_{i,l} = \sum_{j=1}^{n} a_{i,j} \sum_{k=1}^{p} b_{j,k} c_{k,l} = \sum_{j=1}^{n} \sum_{k=1}^{p} a_{i,j} b_{j,k} c_{k,l} \, , \tag{51}$$

matr-assocr3

woraus leicht erkennbar ist, dass diese beiden Ausdrücke gleich sind, da sie sich lediglich durch eine Umordnung der Summationsreihenfolge (die Summen sind ja endlich!) unterscheiden. Genaugenommen benutzt man auch noch die Assoziativität (nicht einmal die Kommutativität) der Multiplikation der Skalare in $\mathbb{K}$ [37].

Noch besser ist es, das Ganze nicht nur mit Koordinaten, sondern in Form von Linear-Kombinationen der Spaltenvektoren $(\vec{\mathbf{a}}_j)$ zu schreiben, und mit entsprechender Notation für die anderen Matrix-Matrix-Produkte und den involvierten Spaltenvektoren. Dann haben wir für $1 \leq k \leq p$ und , $1 \leq l \leq q$:

$$\vec{\mathbf{g}}_k = \sum_{j=1}^{n} b_{j,k} \vec{\mathbf{a}}_j, \quad \vec{\mathbf{u}}_l = \sum_{k=1}^{p} c_{k,l} \vec{\mathbf{g}}_k$$

woraus sich durch Einsetzen/Substitution ergibt:

$$\vec{\mathbf{u}}_l = \sum_{j=1}^{n} \left(\sum_{k=1}^{p} b_{j,k} c_{k,l} \right) \vec{\mathbf{a}}_j = \sum_{j=1}^{n} h_{j,l} \vec{\mathbf{a}}_j .$$

□

Bei dieser Art der Darstellung spielt es nun keine Rolle mehr, dass die Vektoren $(\vec{\mathbf{a}}_j)_{1 \leq j \leq n}$ eine Kollektion von Spaltenvektoren im $\mathbb{R}^m$ sind. Es könnten dies genauso andere (abstrakte oder konkrete Vektoren) sein.

[37] Es wird stark empfohlen, sich diese Beweisführung gut einzuprägen, weil sie viel über das Aufschreiben von Matrix-Multiplikationen in variablem Kontext, mit verschiedenen Dimensionen und verschiedenen Indizierungen aussagt!, als ein gutes Trainingsfeld darstellt.

Durch Uminterpretation des Beweises kommen wir also zu folgender Interpretation: *Gegeben seien n Vektoren $(\mathbf{v}_j)_{1\leq j\leq n}$ aus denen p verschieden Linearkombinationen gewonnen werden, deren Koeffizienten wir der Einfachheit halber in der Form einer $n \times p$ Matrix $\boldsymbol{B}$ beschreiben. Jede Spalte enthält dann die benötigten n Koeffizienten, und die Nummer der Spalte gibt an die wievielte Linear-Kombination gebildet wird. Bildet man daraus nun unter Zuhilfenahme einer $p \times q$ weitere q Linear Kombinationen (von Linearkombinationen), so besagt der oben ausgeführte Rechnung, dass diese sich wieder umschreiben lässt als eine Linear-Kombination der ursprünglichen n Vektoren. Weil wir als Endprodukt gerade q Vektoren gebildet haben, benötigen wir zur Beschreibung eine $n \times q$ Matrix, und unsere Rechnung oben besagt: die Matrix $\boldsymbol{H} = \boldsymbol{B} * \mathbf{C}$ ist genau die benötigte Matrix! Wir haben also bewiesen (unter Verzicht auf Details):*

linlincomb1

Lemma 10. *Die Menge der endlichen Linear-Kombinationen einer Menge von Vektoren ist selbst bezüglich gegenüber der Bildung weiterer Linear-Kombinationen abgeschlossen. Schreibt man die verwendeten Koeffizienten in Matrix-Form an, so entspricht die iterierte Bildung von Linear-Kombinationen genau der bereits definierten Matrix-Multiplikation. Dabei muss man sich die Matrizen (mit entsprechendem Format nach dem Domino-Prinzip von Rechts auf die Spalten-Vektoren von $\boldsymbol{A}$ bzw. die abstrakten Vektoren wirkend denken).*

Als ein Spezial-Fall dieser Überlegung können wir folgende Konsequenz angeben, welche den Fall $n = p = q$ betreffen. Nehmen wir an, dass wir zwei $n \times n$-Matrizen $\boldsymbol{B}$ und $\mathbf{C}$ haben, welche invers zueinander sind, d.h. welche $\boldsymbol{B} * \mathbf{C} = \boldsymbol{N}_n$ erfüllt.

Bildet man mit Hilfe der $n \times n$-Matrix $\boldsymbol{B}$ n Linear-Kombinationen $(\vec{\mathbf{g}}_k)$ von n Vektoren $(\vec{\mathbf{a}}_j)$, so kann man die ursprünglichen Vektoren daraus durch Linear-Kombination mit den Koeffizienten aus $\mathbf{C}$ wieder zurückgewinnen!

Eine wichtige Folgerung aus dieser Beobachtung ist der folgende Satz:

erzteilraum1

Theorem 4. *Hat man eine beliebige Menge M von "Vektoren" in einem Vektorraum*[38] $\boldsymbol{V}$*. Dann gibt es einen kleinsten Teilraum (manchmal als "linear span", oder lineare Hülle von M erzeugter Teilraum in $\boldsymbol{V}$ bezeichnet) $\boldsymbol{V}_M$, der M enthält. Es ist dies genau die Menge der endlichen Linearkombinationen von Elementen aus M. Wir werden gelegentlich auch schreiben (2014)* $\mathbf{LH}(\mathbf{M})$.

Proof. Klar ist, dass die Menge der endl. Linearkombinationen M enthält, und (soeben bewiesen), dass dies ein Vektorraum ist (d.h. ein Teilraum von $\boldsymbol{V}$).

Umgekehrt enthält jeder Vektorraum $\boldsymbol{W}$, welcher ein Teilraum von $\boldsymbol{V}$ ist und M enthält, notwendigerweise alle Summen, alle Vielfachen von Vektoren aus M und daher (!Induktionsbeweis) beliebige endliche Linear-Kombinationen [diese Aussage wäre ein eigenes Lemma wert!], daher die Menge aller endlichen Linear-Kombinationen, also genau das oben beschriebene *lineare Erzeugnis* der Menge $M \subset \boldsymbol{V}$. □

[38] An dieser Stelle wird NICHT vorausgesetzt, dass diese Ausgangsmenge endlich ist, lediglich die Linear-Kombinationen werden nur aus endlich vielen aus dieser vorgegebenen Menge von Vektoren gebildet. Beachte auch, dass die Reihenfolge der Vektoren beim Bilden von Linear-Kombinationen keine Rolle spielt, d.h. dass es sinnvoll ist, von einer Menge zu sprechen.

Terminology: Wenn $\boldsymbol{W} \subset \boldsymbol{V}$ das lineare Erzeugnis von M ist, so nennt man M ein **Erzeugendensystem** für (den Teilraum) $\boldsymbol{W}$. Wir sagen auch: "$\boldsymbol{W}$ stimmt mit der linearen Hülle (der Menge aller Lin.Komb. von Elementen aus m überein).

Ein Vektorraum mit *endlichem Erzeugendensystem* wird ein **endlich-dimensionaler Vektorraum** genannt. Wir werden später noch sehen (sobald der Begriff der Basis geklärt ist) dass es ein Raum mit einer *endlichen Basis* ist (die durch "Weglassen" von Elementen aus dem endlichen Erzeugersystem entsteht), der somit eine (!endliche) und wohldefinierte *Dimension* hat[39].

Bemerkung: Wenn $M = \mathbf{v}$ ein einziger Vektor ist, so ist der davon erzeugte (eindimensionale) Teilraum einfach von der Form $\boldsymbol{W} = \{\lambda\mathbf{v}, \lambda \in \mathbb{K}\}$. Man spricht oft auch von einer "Geraden" in $\mathbb{K}^n$.

Wir werden uns bald mit *minimalen Erzeugendensystemen* bzw. *Basen* befassen. Beispielsweise ist die Menge der Monome $\{1, t, t^2, \ldots, t^n\}$ ein Erzeugendensystem für den Teilraum der Polynomfunktionen vom Grad $\leq n$ (über $\mathbb{R}$ oder $\mathbb{C}$), den wir mit $\mathscr{P}_n(\mathbb{R})$ bzw. $\mathscr{P}_n(\mathbb{C})$ bezeichen werden. Wir werden noch sehen, unter welchen Bedingungen eine andere Kollektion von $n+1$ Polynomfunktionen in $\mathscr{P}_n(\mathbb{R})$ ebenfalls eine Basis (de facto also ein minimales Erzeugendensystem ist) ist, z.B. weil es linear unabhängig ist.

monomindep

Lemma 11. *Für jedes $n \in \mathbb{N}$ ist die Menge der $n+1$ Monome $\{1, t, t^2, \ldots, t^n\}$ eine Basis für $\mathscr{P}_n(\mathbb{C})$ bzw. $\mathscr{P}_n(\mathbb{R})$.*

Proof. Wir schreiben o.B.d.A. den Beweis für $\mathscr{P}_n(\mathbb{R})$ auf. Es ist klar die Monome ein Erzeugendensystem von $\mathscr{P}_n(\mathbb{R})$ sind, so ist der Raum ja innerhalb des Vektorraums aller reellen Funktionen mit punktweiser Addition definiert.

Für den Beweis der linearen Unabhängigkeit (dann ist klar, dass die Monome ein Basis sind!) erscheint es als vorteilhaft, die Monome als eine *geordnete Basis* mit absteigenden Potenzen anzuordnen, d.h. $\mathcal{M}^n := \{t^n, t^{n-1}, \cdots, t^2, t, 1\}$, so wie es in MATLAB/OCTAVE gehandhabt wird. Einem Vektor $\vec{\mathbf{a}} \in \mathbb{R}^{n+1}$ entspricht dann das Polynom

$$p(t) = \sum_{k=1}^{n+1} a_k t^{n-k}, \tag{52}$$

evalpolML

welches auch vom MATLAB Befehl `polyval` so verstanden wird. Der Befehl

```
>> vals = polyval(a,x)
```

etwa für $\vec{\mathbf{x}} = [x_1, x_2, \cdots, x_m]$ ergibt gerade die Werte des Polynoms $p(t)$ in (52) an den angegebenen Stellen[40].

Sei nun $p(t) = 0$ für eine nichtverschwindende Koeffizientenfolge $\vec{\mathbf{a}} \in \mathbb{K}^{n+1}$. Dann gibt es einen führenden (den ersten in dieser Reihenfolge, der zur höchsten Potenz, die wirklich vorkommt gehört) Koeffizienten. Da mit $p(t) = 0$ auch $\lambda p(t) = 0, \forall t \in \mathbb{R}$ gilt, können wir o.B.d.A. annehmen, dass $a_{k_0} > 0$ gilt.

[39] Aber zur reinen Definition ist es nicht notwendig, den Begriff der Basis zu verwenden, und schon gar nicht den der Dimension!

[40] Mit dem Befehl `plot(x,polyval(a,x));` kann man auch Polynomfunktionen plotten!

Für $t > 0$ ist aber dann auch $r(t) := p(t) \cdot t^{-k_0} \equiv 0$, aber in Potenzen ausgedrückt ist die (rationale) Funktion $r(t)$ einfach gleich

$$r(t) = a_{k_0} + \sum_{j=1}^{n+1-k_0} a_{k_0+j} t^{-j}.$$

(? richtige Exponenten). Nimmt man den Limes für $t \to \infty$ erhält man (das ist jetzt eine Analysis Argument)

$$0 = a_{k_0} + \lim_{t\to\infty} \sum_{j=1}^{n+1-k_0} a_{k_0+j} t^{-j} = a_{k_0} + 0,$$

im Widerspruch zur Annahme $a_{k_0} > 0$. □

Später werden wir noch mit Hilfe der Lagrange Interpolationspolynome sehen, dass auch $\mathcal{P}_n(X)$, die Menge der auf eine endliche Menge $X \subset \mathbb{C}$ eingeschränkten Polynomfunktionen linear unabhängig ist genau dann wenn X mindestens $n+1$ verschiedene Punkte enthält. Insbesondere sind Mengen wie ein Intervall $I = [a, b]$ oder die Menge der Einheitswurzeln $\mathbf{U}_N \subset \mathbb{C}$ (Ecken es regelmässigen N-Eckes am Einheitskreis in der komplexen Ebene, d.h. innerhalb $\mathbf{U} := \{z \in \mathbb{C} \,|\, |z| = 1\}$, solche Mengen, wenn nur $N > n+1$ gilt.

Besteht die Menge nur aus einem Element/Vektor $\mathbf{m} \in M$ dann ist das lineare Erzeugnis einfach die Menge der Vielfachen von $\mathbf{m}$, d.h.

$$\boldsymbol{V}_M = \{\lambda \mathbf{m}\,,\, \lambda \in \mathbb{K}\}.$$

Erstes Beispiel (zum Nachdenken): Die Menge $M = \{(t-1)^k\,,\, 0 \le k \le n\}$ ist ebenfalls eine Erzeugendensystem für $\mathscr{P}_n(\mathbb{R})$, oder gleich noch allgemeiner: Für jedes $\alpha \in \mathbb{R}$ ist die Menge $M = \{(t-\alpha)^k\,,\, 0 \le k \le n\}$ ein Erzeugendensystem für $\mathscr{P}_n(\mathbb{R})$ bzw. $\mathscr{P}_n(\mathbb{C})$.

Wenn die Menge M die Menge der Zeilenvektoren einer Matrix ist, so ist das Erzeugnis genau der **Zeilenraum** von $\boldsymbol{A}$. Wir werden wohl das Symbol $\mathbf{Z}(\boldsymbol{A})$ verwenden. Entsprechend gibt es das linear Erzeugnis der Menge der Spaltenvektoren von $\boldsymbol{A}$, den wir dann **Spaltenraum** von $\boldsymbol{A}$ nennen wollen, und mit $\mathbf{Sp}(\boldsymbol{A})$ bezeichnen wollen[41].

Mit der Aussage über den Zusammenhang zwischen iterierter Linearkombinationsbildung und Matrix-Multiplikation lässt sich leicht zeigen: Hat man n Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in einem Vektorraum, und zwei zueinander inverse Matrizen $\boldsymbol{A}$ und $\boldsymbol{B}$, mit $\boldsymbol{A} * \boldsymbol{B} = \mathbf{Id}_\mathbf{n} = \boldsymbol{B} * \boldsymbol{A}$, und bildet man mittels $\boldsymbol{A}$ eine neue Kollektion von ebenfalls n Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_n$, dann haben beide Kollektionen (die alte wie die neue) das gleiche! Erzeugnis!

Proof. Sei $\mathbf{x}$ im linearen Erzeugnis der $\mathbf{v}$-Vektoren. Dann gibt es einen Zeilenvektor λ der Länge n (eine $1 \times n$-Matrix) mit den Koeffizienten. Wir können aber auch

$$\lambda = \lambda * \mathbf{Id}_n = (\lambda * \boldsymbol{B}) * \boldsymbol{A}$$

[41] Eine der grundlegenden Aussagen der Linearen Algebra ist die Aussage, dass diese beide Räume die gleiche *Dimension* haben, die sich aus der Zahl der Unbekannten minus der Zahl der relevanten Gleichungen, d.h. der nicht-trivialen Gleichungen im Gleichungssystem nach Anwendung der Gausselimination, d.h. minus der Zahl der führenden Einsen bzw. Pivot-Elemente.

schreiben. Die Wirkung von $\boldsymbol{A}$ auf den $\mathbf{v}$-Vektoren macht aber gerade aus den $\mathbf{v}$-Vektoren die $\mathbf{w}$-Vektoren, also kann man mit den Koeffizienten $\mu = \lambda * \boldsymbol{B}$ denselben Vektor $\mathbf{x}$ als Linear-Kombination der $\mathbf{w}$-Vektoren schreiben. Also ist $\mathbf{x}$ auch im linearen Erzeugnis der $\mathbf{w}$-Vektoren.

Mit vertauschten Rollen kann man auch zeigen, dass die Umkehrung gilt, somit haben die beiden Kollektionen das gleiche! lineare Erzeugnis.

Wir haben dazu nur(!?) die Assoziativität der Matrix-Multiplikation benutzt! Weiters ist es nicht einmal notwendig vorauszusetzen, dass $\boldsymbol{B}$ eine quadratische, invertierbare Matrix ist. Es hätte genügt vorauszusetzen, dass $\boldsymbol{B}$ eine rechtsinverse Matrix $\mathbf{C}$ hat (die dann notwendigerweise dasselbe Format wie $\boldsymbol{B}^t$ haben muss), mit $\boldsymbol{B} * \mathbf{C} = \boldsymbol{I}_n$.

□

10.1 Euler's Formel revisited

Als ein konkretes Beispiel nehmen wir uns die beiden folgenden 2×2-Matrizen vor:
Die beiden Matrizen

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \tag{53}$$ Eulerb2

und

$$\mathbf{C} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tag{54}$$ Eulerc2

Angewendet auf die Funktionen $\cos(x)$ und $\sin(x)$ können wir unter Verwendung der sog. *Eulerschen Formel* die (komplexe) Exponential-Funktion als Linear-Kombination dieser beiden Funktionen betrachten. Zunächst einmal

$$e^{ix} = \cos(x) + i\sin(x),\ x \in \mathbb{R}. \tag{55}$$ Euler2a

Daraus ergibt sich weiterhin

$$e^{-ix} = \cos(-x) + i\sin(-x) = \cos(x) - i\sin(x). \tag{56}$$ Euler1a

Schreibt man die verwendeten Koeffizienten in Matrix Format auf so bekommt man genau die oben angegebene Matrix $\boldsymbol{B}$.
Umgekehrt kommt man durch Addieren der beiden Formeln (53) und (54) leicht zu

$$2\cos(x) = e^{ix} + e^{-ix}. \tag{57}$$ Euler3b

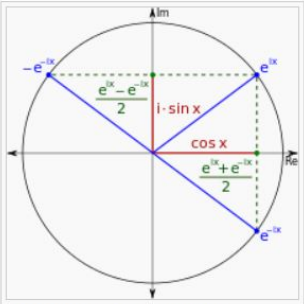
und durch Nehmen der Differenz

$$2i\sin(x) = e^{ix} - e^{-ix}, \tag{58}$$ Euler3c

woraus sich die bekannten Formeln folgen:

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}\ , \quad \sin(x) = \frac{-i(e^{ix} - e^{-ix})}{2} \tag{59}$$ Euler3d

folgen (siehe auch Beginn des Skriptums!).



Beziehung zwischen Sinus, Kosinus und Exponentialfunktion

Abbildung 3: Euler-Demo in WIKI

Ende des neuen Einschubs, 2014 Brauchbare Information dazu gibt es auch auf `http://de.wikipedia.org/wiki/Eulersche_Formel`

EMPFEHLUNG: es ist sehr empfehlenswert, diese Skizze SELBST in GeoGebra nachzubauen!

Es gibt auch noch zwei weitere Sichtweisen auf das Problem, abhängig davon ob man die (komplexe) Exponentialfunktion als *komplex* Linearkombination der (reellen) Funktionen $\sin(x)$ und $\cos(x)$ ansieht (mit Koeffizienten 1 bzw. $\pm i$), oder als reele Linearkombination der (komplexwertigen) Funktionen $\cos(x)$ und $i\sin(x)$. Im ersten Fall ist die Situation durch ein paar zueinander inversere komplexer 2×2-Matrizen beschreibbar, im zweiten Fall geht es lediglich um reelle, invertierbare 2×2-Matrizen!

Reihenfolge ist noch zu adaptieren!

Eine geometrische Interpretation der Matrix $EUL = [1, 1; 1, -1]$ und daraus eine Erklärung, warum diese die "magische Eigenschaft

$$EUL * EUL = 2Id_2$$

haben kann, erfolgt in der Vorlesung. Im Prinzip ist diese eine Drehstreckung ($45°$ im math. positiven Sinn), gefolgt von einer Spiegelung an der Hauptachse, oder in Matrizenschreibweise

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{60}$$

ist gerade das $\sqrt{2}$-Fache einer solchen Drehung, und dann die Spiegelungsmatrix "draufmultipliziert" (bitte selber verifizieren!). Man überlege sich rechnerisch wie auch geometrisch dass das nochmalige Anwenden dieser Operation (bis auf die zweimal eintretende Streckung um den Faktor $\sqrt{2}$) gerade die identische Abbildung ist.

Natürlich muss man nur die inverse Matrix nach der üblichen Formel Berechnen (offenbar ist die Determinante $ad - b = 2$!) eben genau die Hälfte der ursprünglichen Matrix.

Eine der wichtigsten Beobachtung betreffend die Gauss-Elimination ist die Aussage, dass diese Prozedur den Zeilenraum des Systems nicht verändert. Als Begründung kann man argumentieren, dass nur Linearkombinationen der alten Zeilen zu neuen Linear-Kombinationen verbunden werden. Weil jeder der drei Typen von Operationen im Rahmen der G-Elim. durch einen gleichartigen Schritt (z.B. Subtraktion eines Vielfachen einer Zeile von einer anderen als Inversion der Addition...) wieder rückgängig gemacht werden kann, gilt die Relation auch umgekehrt, der von den alten Zeilen aufgespannte Raum ist in dem von den neuen Zeilen aufgespannten Zeilenraum enthalten (die alten sind ja auch als die "ganz neuen", aus dem neuen System durch Rückgängig-Machung erhaltene zu sehen!).

Eine interessante (und weiterführende) Betrachtung ist nun, die einzelnen Schritte der Gauss-Elimination als Matrix-Multiplikation mit Elementar-Matrizen zu interpretieren.

Wir brauchen auch noch den Begriff einer inversen Matrix (der hat natürlich nur für quadratische Matrizen einen Sinn):

Definition 9. Es sei $\boldsymbol{A}$ eine $n \times n$-Matrix. Eine Matrix[42] $\boldsymbol{B}$ heißt **inverse Matrix** zu $\boldsymbol{A}$ (oder zu $\boldsymbol{A}$ inverse Matrix) wenn gilt

$$\boldsymbol{A} * \boldsymbol{B} = \mathbf{Id_n} = \boldsymbol{B} * \boldsymbol{A} \tag{61}$$

invmatr

- Wir werden noch sehen, dass die invertierbaren Matrizen mit der üblichen Matrix-Multiplikation eine (nicht-kommutative) Gruppe bilden ($GL(n, \mathbb{R})$), und dass die inverse Matrix das Gruppeninverse zum Gruppenelement $\boldsymbol{A}$ ist.
- zur Erinnerung (aus der Gruppentheorie): die Matrix $\boldsymbol{B}$ muss also Links- und Rechts-Inverse sein. Wir werden aber später noch sehen, dass es genügt, eine der beiden Gleichheiten in (61) zu haben, weil die andere dann (aufgrund der Struktur der Matrix-Multiplikation, und nicht aus abstrakten, *algebraischen* Gründen) automatisch folgt.
- Im Fall rechteckiger Matrizen ist einfach zu zeigen, dass man "einseitige" inverse Matrizen haben kann, die aber nicht zweiseitig inverse sein können. Die Einbettung der $x - y$-Ebene $\mathbb{R}^2$ in den euklidischen Raum $\mathbb{R}^3$ wird ein einfaches Beispiel sein, dass erst später zu diskutieren sein wird.

10.2 Diverse Interpretation der Matrix Multiplikation

Die erste geometrische Interpretation der üblichen Matrix-Vektor Multiplikation, d.h. der Anwendung einer $m \times n$-Matrix $\boldsymbol{A}$ auf einen Spaltenvektor $\vec{\mathbf{x}}x \in \mathbb{R}^n$ ist genau das **Bilden von Linearkombinationen**. Es gilt

matrlincomb

Theorem 5.

$$\boldsymbol{A} * \mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{a}_k, \tag{62}$$

[42] Diese Matrix muss dann notwendigerweise auch eine $n \times n$-Matrix sein!

d.h. der Vektor $\mathbf{y} := \boldsymbol{A} * \mathbf{x}$ *ist die Linearkombination der Spaltenvektoren* $[\mathbf{a}_1, \cdots, \mathbf{a}_n]$ *mit den Koeffizienten* $[x_1, \ldots, x_n]$ *zu einem neuen Vektor* $\mathbf{y} \in \mathbb{R}^m$.

Insbesondere ergibt die Anwendung der Matrix-Multiplikation auf einen der Einheitsvektoren, d.h. die Wahl $\mathbf{x} = \mathbf{e}_k$ *für ein* $k, 1 \leq k \leq n$, *dass* $\boldsymbol{A} * \mathbf{e}_k = \mathbf{a}_k$, *das ist die k-te Spalte von* $\boldsymbol{A}$.

Eine mögliche weitere Erklärung dafür wird noch im Zusammenhang mit der Diskussion von *linearen Abbildungen* gegeben.

Eine unmittelbare Folgerung aus dieser Beobachtung, ein sogennantes Korollar (auf Engl.: Corollary) ist die folgende Aussage:

Corollary 2. *Der Bildraum der Abbildung* $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ *, d.h. die Menge*

$$\{\mathbf{y} \,|\, \mathbf{y} = \boldsymbol{A} * \mathbf{x} \quad \textit{für ein } \mathbf{x} \in \mathbb{R}^n\}$$

stimmt genau mit dem Spaltenraum $\mathbf{Sp}(\boldsymbol{A})$ *überein.*

Ein lineares Gleichungssystem der Form $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ *ist also genau dann lösbar wenn gilt:* $\mathbf{b} \in \mathbf{Sp}(\boldsymbol{A})$[43].

Wir zielen auf die wichtige Aussage ab: Die allgemeine Lösung eines inhomogenen Gleichungssystems besteht aus einer (beliebigen) *spezielle Lösung des inhomogenene Systems* plus alle Lösungen des entsprechenden *homogenene linearen Gleichungssystems* (d.h. das Gleichungssystem von der Form $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$).

Eine gleichwertige (äquivalente) Formulierung dieser Aussage ist die folgende:

inhomspec **Lemma 12.** *Je zwei Lösungen* $\mathbf{y}_1$ *und* $\mathbf{y}_2$ *eines inhomogenen linearen Gleichungssystems* $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ *unterscheiden sich nur durch eine Lösung des entsprechenden homogenene linearen Gleichungssystems.*

Proof. Hat man zwei solche Lösungen, so gilt für die Differenz, nennen wir sie $\mathbf{z} = \mathbf{y}_1 - \mathbf{y}_2$

$$\boldsymbol{A} * \mathbf{z} = \boldsymbol{A} * (\mathbf{y}_1 - \mathbf{y}_2) =!!! \quad \boldsymbol{A} * \mathbf{y}_1 - \boldsymbol{A} * \mathbf{y}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

D.h. $\mathbf{z}$ ist eine Lösung von $\boldsymbol{A} * \mathbf{z} = \mathbf{0}$. Umgekehrt kann man mit dem gleichen Argument feststellen, wenn man eine bel. derartige Lösung $\mathbf{z}$ zu einer beliebigen Lösung $\mathbf{y}$ von $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ hinzufügt, dass man wieder nur eine Lösung von $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ bekommt. □

Weiteres Material

Die Charakterisierung der Abbildungen der Form $\mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ durch die Eigenschaft der Linearität, d.h. das Respektieren von Linear-Kombinationen (Stoff der Woche vom 22.-25. Nov.), mit den Teilaufgaben: Feststellen der Invertierbarkeit einer linearen Abbildung anhand einer Matrix und deren Inverser (d.h. mittels Gauss).

Injektivität, Surjektivität und Bijektiviät und die entsprechenden Begriffe der LA (*lineare Unabhängigkeit, Erzeugendensystem, bzw. Basis*).

Formale Definitionen:

[43] Fragen der Eindeutigkeit der Lösung etc. werden in Kürze diskutiert

Erzeugsys **Definition 10.** Eine Menge $M \subset \boldsymbol{V}$ heißt **Erzeugendensystem** von $\boldsymbol{V}$, wenn $\boldsymbol{V}_M = \boldsymbol{V}$, d.h. wenn jedes $\mathbf{v} \in \boldsymbol{V}$ sich als (endliche) Linearkombination von Elementen aus M schreiben läßt.

linunabh **Definition 11.** Eine Menge $M \subset \boldsymbol{V}$ wird als **linear abhängig** bezeichnet, wenn es eine endliche Teilfoge in M gibt, die in einer nichttrivialen Relation stehen, d.h. durch eine vom Nullvektor verschiedene Koeffizientefolge $\mathbf{c}$ zum Nullvektor kombiniert werden können, d.h. wenn die Relation $\sum_{k=1}^{s} c_k m_k = \mathbf{0}$ in $\boldsymbol{V}$ gilt, obwohl $\vec{\mathbf{c}} \neq \mathbf{0} \in \mathbb{R}^s$ gilt.

In *Negation* dieser Eigenschaft nennt man eine Menge **linear unabhängig**, d.h. es gilt dann für jede beliebige Teilfolge: Es gibt keinerlei nicht-triviale Relation, oder logisch gleichwertig ausgedrückt: Wenn ein Linearkombination von Elementen den Nullvektor in $\boldsymbol{V}$ darstellt, d.h. wenn $\sum_{k=1}^{s} c_k m_k = \mathbf{0}$ in $\boldsymbol{V}$ gilt, dann muß notwenderweise die Koeffizentenfolge $\vec{\mathbf{c}} = \mathbf{0} \in \mathbb{R}^s$ sein.

basisdef0 **Definition 12.** Eine **Basis** ist ein linear unabhängiges Erzeugendensystem $\mathscr{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\} \subset \boldsymbol{V}$, oder auch ein Koordinatensystem, d.h. *jeder Vektor* $\mathbf{v} \in \boldsymbol{V}$ ist *eindeutig* als Linearkombination der Elemente von $\mathscr{B}$ darstellbar, d.h.

$$\mathbf{v} = \sum_{k=1}^{n} c_k \mathbf{b}_k.$$

Diese eindeutig bestimmte Koeffizientenfolge wird oft auch "in ihrer Funktion" beschrieben, und um in Symbolen auszudrücken, dass $\vec{\mathbf{c}}$ die Folge ist, die dadurch gekennzeichnet ist, dass sie eben die Koeffizienten von $\mathbf{v}$ bilden, wird sinnvollerweise ein eigenes Symbol verwendet. Wir schreiben für den **Koeffizientvektor**

$$[\mathbf{v}]_{\mathscr{B}} := \vec{\mathbf{c}} \quad \forall \mathbf{v} \in \boldsymbol{V}. \tag{63}$$ coeffbas0

Psychlogischer/linguistischer Kommentar:
Der Vorteil diese neuen Schreibweise ist vor allem dann gegeben, wenn man denselben Vektor in *verschiedenen* Basen darstellt, oder mehrere Vektoren braucht, oder mehrer Basen in verschiedenen Räumen in Verwendung hat. Da gehen bei der üblichen Konvention, wo man für Matrizen die Symbole $\boldsymbol{A}, \boldsymbol{B}, \mathbf{C}$ etc. verwendet, rasch die Buchstaben aus. Auch die Verwendung von Laufnummern ist manchmal im Konflikt mit den Konventionen zur Beschreibung von Eintragungen der Matrix. Wir verwenden bisher das Symbol $\vec{\mathbf{a}}_j$ oder $\mathbf{a}_j$ für die j-te Spalte von $\boldsymbol{A}$, bzw. das schreiben $a_{3,5} = 7$, um auszudrücken, dass die Eintragung in der Matrix $\boldsymbol{A}$ in der dritten Zeile und fünften Spalte eben $= 7$ ist. Das wird schwierig, wenn man mehrere Matrizen $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$ hat. Wo soll man da die *Indices* hinschreiben. Ein Ausweg ist gerade noch (wird beim Beweis der Kompositionsregeln genutzt) die involvierten Matrizen mit $\boldsymbol{A}^1, \boldsymbol{A}^2, \boldsymbol{A}^3$ zu verwenden, aber *eigentlich* ist das wieder gefährlich, denn sollte $\boldsymbol{A}^2$ nicht für $\boldsymbol{A} * \boldsymbol{A}$ stehen. Also schreiben wir besser $\boldsymbol{A}^{(2)}$, mit Eintragungen der Form $a^{(2)}_{j,k}$?

Weiters wollen wir zur Sicherung des Begriffes nochmals festhalten, dass im Falle einer Basis $\mathscr{B}$ in $\mathbb{C}^n$, die durch eine (notwendigerweise vom Format) $n \times n$-Matrix $\boldsymbol{B}$ gegeben ist, d.h. $\mathscr{B} = \{\mathbf{b}_1, \cdot, \mathbf{b}_n\}$ die gemachten Konventionen darauf hinauslaufen, dass wir

haben:[44]

$$\mathbf{v} = \boldsymbol{B} * [\mathbf{v}]_{\mathscr{B}}, \quad \text{also} \quad [\mathbf{v}]_{\mathscr{B}} = \boldsymbol{B}^{-1} * \mathbf{v}, \ \forall \mathbf{v} \in \boldsymbol{V}. \tag{64}$$

coeffrep4

Übungsmaterial u.a.: Aufstellen von Matrizen zu bestimmten geometrischen Aufgaben... (später z.B.: Projektion auf eine Ebene oder auf eine Gerade).

SS14: siehe http://www.univie.ac.at/NuHAG/FEICOURS/ss14/LINALGIDEAS.pdf

11 Inverse Abbildungen und Inverse Matrizen

Diese Themenkreis ist zentral für die ganze Vorlesung.

Es geht um die folgenden Fragen, die sich im Zusammenhang mit einer gegebenen linearen Abbildung (oft muss die erst in der Fragestellung entdeckt! werden) zwischen zwei Vektorräumen stellen. Wir haben also die Situation dass $T : \boldsymbol{V} \to \boldsymbol{W}$ eine lineare Abbildung sind. In der allgemeinen Diskussion werden wir meistens voraussetzen, dass $\boldsymbol{V}$ n-dimensional ist und der Zielraum $\boldsymbol{W}$ sei m-dimensional[45].

1. Ist für gegebenes $\mathbf{w}_0 \in \boldsymbol{W}$ die Gleichung $T(\mathbf{v}) = \mathbf{w}_0$ lösbar;
2. Wenn die Gleichung lösbar ist, ist diese Lösung dann eindeutig bestimmt?
3. Ist die Gleichung für *jede mögliche* rechte Seite (r.S.) $\mathbf{w}_0$ lösbar?
4. Ist die Gleichung womöglich für jede rechte Seite eindeutig lösbar (Idealfall!).

Das Skriptum wird sich zunächst mit dem Fall $\boldsymbol{V} = \mathbb{R}^n$ und $\boldsymbol{W} = \mathbb{R}^m$ befassen. Dann ist bekanntlich jede lineare Abbildung T von der Form $\vec{\mathbf{x}} \mapsto \boldsymbol{A} * \vec{\mathbf{x}}$ für eine eindeutig bestimmte $m \times n$-Matrix $\boldsymbol{A}$ (über $\mathbb{K}$)[46].

ACHTUNG: am 25.Mai wurde hier Umgruppierung durchgeführt, leider nicht die letzte bis zum Abschluss des Semesters, weil nich einiges an Text hinzukommen wird! HGfei

11.1 Empfohlene Realisierung der Gauss-Elimination

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \begin{matrix} \\ \mathbf{II} - \mathbf{2} \cdot \mathbf{I} \\ \mathbf{III} - \mathbf{3} \cdot \mathbf{I} \end{matrix} \Rightarrow \begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \begin{matrix} \\ \\ \mathbf{III} - \mathbf{2} \cdot \mathbf{II} \end{matrix} \Rightarrow \begin{pmatrix} 1 & 4 & 7 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Ein anderes gute Übungsbeispiel sind Matrizen der Form ($k \in \mathbb{N}$)

[44] Dieser Zusammenhang wird in vielen praktischen Anwendungen immer wieder zur Anwendung kommen. Daher sollten sich die Leser mit den Begriffen und der verwendeten Terminologie gut vertraut machen!

[45] Ich nenne das manchmal **Standards-Situation**! Wenn diese Buchstaben in einer konkreten Situation noch nicht anderwärtig vergeben sind, spricht ja auch nichts dagegen, die Bezeichungen so zu wählen, um in "gewohnter Sprache" über die Siuation reden zu können. Wenn nicht anders vorgegeben werden wir die/eine Basis in $\boldsymbol{V}$ mit $\mathscr{B}$ bezeichnen und eine andere Basis für $\boldsymbol{W}$ mit $\mathscr{C}$. Wir haben also $\#\mathscr{B} = n$ und $\#\mathscr{C} = m$.

[46] Der allgemeine Fall wird später noch auf genau diesen Fall zurückgeführt, teilweise verbrämt mit weiterer (kompliziert wirkender) Terminologie, von der man sich aber nicht verwirren lassen sollte.

11.2 Berechnung der inversen Matrix via Gauss-Elimination

Wir wollen in diesem Abschnitt kurz zeigen, dass die Bestimmung der inversen Matrix $\boldsymbol{A}^{-1}$ im Prinzip nichts anderes is, als die Lösung folgender Familie von n linearen Gleichungs-Systemen (jeweils mit derselben System-Matrix $\boldsymbol{A}$), nämlich

$$\boldsymbol{A} * \mathbf{x} = \mathbf{b}, \quad \text{für} \quad \mathbf{b} = \mathbf{e}_k, \quad 1 \leq k \leq n.$$

Naheliegenderweise schreibt man diese diversen rechten Seiten einfach neben der Matrix $\boldsymbol{A}$, bildet also eine rechteckige $n \times 2n$-Matrix, und wendet darauf die Schritte der Gauss-(Jordan)-Elimination an, solange bis der linke Teil, in dem am Anfang die Matrix $\boldsymbol{A}$ gestanden ist, zur Einheitsmatrix $\boldsymbol{N}_n$ vom Format $n \times n$ geworden ist.

Die Behauptung ist, dass dann in der rechnten Hälfte der Matrix, also links von der neuen Einheitsmatrix, genau die inverse Matrix steht!! Das ist einmal schon ein gutes/brauchbares Rezept zur Bestimmung der inversen Matrix, dass man als so ein Rezept hinnehmen und verwenden kann, wir wollen aber natürlich verstehen, warum diese Behauptung stimmt (immer und sozusagen mit Garantie, oder nur unter gewissen Voraussetzung!??). Wir werden sehen, dass die einzige Voraussetzung ist, dass $\boldsymbol{A}$ eine invertierbare Matrix ist, dann kann auch schon gesichert werden, dass man mit Hilfe der Gauss'schen Eliminationsmethode zum Ziel kommt (die genau Strategie ist nicht festgelegt, jede Wahl von Schritten, die zum gewünschen Endergebnis führt, ist willkommen und führt in der Tat zum richtigen/gleichen Ergebnis!) .

11.3 Matrizen und ihre Transponierten

Die Transposition einer Matrix ist durch *Stürzen* der Matrix zu erzielen, d.h. $\boldsymbol{B} := \boldsymbol{A}^t$ bedeutet, dass aus einer $m \times n$ Matrix $\boldsymbol{A}$ eine $n \times m$-Matrix $\boldsymbol{B}$ gemacht wird, mit

$$b_{j,k} = a_{k,j}, \quad 1 \leq k \leq m\,,\, 1 \leq j \leq n.$$

In Worten: Zeilenindex wird zu Spaltenindex und umgekehrt. Beispiel:

$$H = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad \Leftrightarrow \quad H^t = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \tag{65}$$

Dementsprechend macht die Transposition aus einem Zeilenvektor einen Spaltenvektor und umgekehrt. Es gilt klarerweise $(\boldsymbol{A}^t)^t = \boldsymbol{A}$.

Eine Matrix $\boldsymbol{A}$ heißt *symmetrisch* wenn gilt $\boldsymbol{A} = \boldsymbol{A}^t$. Sie heißt *orthogonal* wenn gilt $\boldsymbol{A}^t = \boldsymbol{A}^{-1}$.

In MATLAB wird das Transponieren mit dem Befehl $\boldsymbol{A} \rightarrow \boldsymbol{A}.'$ realisiert. Das ist ein Kommando, das mit einer englischen/US Tastatur leicht realisierbar ist (wegen des Punktes) ein wenig komplizierter als die noch wichtigere Operation $\boldsymbol{A} \rightarrow \boldsymbol{A}'$, welche für komplexen Matrizen den Übergang zur konjugiert transponierten Matrix ergibt, d.h. wir haben die folgende Identität: $\boldsymbol{A}' == conj(\boldsymbol{A}.')$, oder $\boldsymbol{A}.' = conj(\boldsymbol{A}')$.

$$B = \begin{pmatrix} 1 & 3i \\ 2 & 4 \end{pmatrix} \quad \Leftrightarrow \quad B' = \begin{pmatrix} 1 & 2 \\ -3i & 4 \end{pmatrix} \tag{66}$$

So ziemlich die einzigen wichtigen Formel für transponierte Matrizen ist

$$(\boldsymbol{A} * \boldsymbol{B})^t = \boldsymbol{B}^t * \boldsymbol{A}^t \ , \ (\boldsymbol{A}' * \boldsymbol{B}') = (\boldsymbol{B} * \boldsymbol{A})'. \tag{67}$$

matrprodtrans

Der Beweis dieser Behauptungen ist eine einfach Jonglier-Übungen mit den entsprechenden Indizierungen und wird daher den LeserInnen überlassen[47].

EINSCHUB April 2014:

Transpositions Regel

$$(\boldsymbol{A} * \boldsymbol{B})' = \boldsymbol{B}' * \boldsymbol{A}'.$$

Wir setzen $\mathbf{C} := \boldsymbol{A} * \boldsymbol{B}$ und verwenden die Buchstaben G, H für die Matrizen $\boldsymbol{A}'$ bzw. $\boldsymbol{B}'$ und $\mathbf{V} = \mathbf{C}'$.

Nach Definition der Matrix Multiplikation gilt (Domino Prinzip): $c_{j,l} = \sum_{k=1}^{n} a_{j,k} b_{k,l}$ und daher $v_{l,j} = c_{j,l} = \sum_{k=1}^{n} a_{j,k} b_{k,l} = \sum_{k=1}^{n} h_{l,k} g_{k,j}$, was genau der Formel für das Matrix Produkt $\mathbf{G} * \mathbf{H}$ entspricht, was ja zu beweisen war.

Statt mit Koordinaten zu arbeiten (und Summen der Länge n) könnte man auch **Skalarprodukte** (!!) betrachten:

$$c_{j,l} = \langle \vec{\mathbf{z}}_j, \vec{\mathbf{b}}_l \rangle,$$

das Skalarprodukt von j-ter Zeile von $\boldsymbol{A}$ mit der l-ten Spalte von $\boldsymbol{B}$. Die transponierte Matrix $\mathbf{C}^t$ hat also Eintragungen der Form $c_{l,j}, 1 \leq l \leq p, 1 \leq j \leq m$. Nun hat man aber dasselbe Skalarprodukt, denn die l-te Zeile von $\boldsymbol{B}^t$ ist gerade die (alte) j-te Spalte von $\boldsymbol{B}$, und die j-te Spalte von $\boldsymbol{A}^t$ ist dasselbe (als n-Tupel in $\mathbb{K}^n$) wie die j-te Zeile von $\boldsymbol{A}$. Aber dieses Skalarprodukt von j-ter Zeile von $\boldsymbol{A}^t$ mit k-ter Spalte von $\boldsymbol{B}^t$ ist ja gerade die Eintragung der Produktmatrix $\boldsymbol{A}^t * \boldsymbol{B}^t$, in Zeile j und Spalte l.

Ende Einschub 2014

defsymorth

Definition 13. Eine Matrix heißt *symmetrisch* wenn gilt $\boldsymbol{A} = \boldsymbol{A}^t$.

Dieser Begriff ist vor allem auf reelle Matrizen anzuwenden. Im Falle von komplexen Matrizen spricht man von einer *selbst-adjungierten* Matrix, wenn $\boldsymbol{A} = \boldsymbol{A}' := conj(\boldsymbol{A}^t)$ gilt[48] (oder komplex symmetrisch).

Man könnte den Beweis auch so beginnen. Gegeben seien zwei rechteckige Matrizen $\boldsymbol{A}$ und $\boldsymbol{B}$, die miteinander multipliziert werden können (Domino-Prinzip). Bezeichnen wir ihr Product mit $\mathbf{C} := \boldsymbol{A} * \boldsymbol{B}$. Andererseits verwenden wir die Buchstaben $\mathbf{E}$ bzw. $\mathbf{F}$ für $\boldsymbol{A}'$ resp. $\boldsymbol{B}'$, und $\mathbf{G} := \mathbf{F} * \mathbf{E}$. Man zeige dass $\mathbf{C}' = \mathbf{G}$ gilt. Dann schreibe man alles mit entsprechender Indizierung auf, um festzustellen, dass nach einer länglichen, aber elementaren Rechnung alles in Ordnung ist.

Trotz der Einfachheit dieser Formel ist sie von großem praktischen Nutzen, weil sie viele Konventionen von Spaltenvektoren auf Zeilenvektoren (und natürlich umgekehrt) übersetzen hilft.

[47] Vermutlich ist ein kleines selbstgerechnetes Beispiel (mit 2×2 "Buchstaben-Matrizen") überzeugend!

[48] Eine komplexe, selbstadjungierte Matrix hat auf der Diagonale stets nur reelle Werte stehen!

Wissen wir derzeit schon, dass die Wirkung einer Matrix die von *links* auf einen *Spaltenvektor* losgelassen wird, darin besteht, dass der Vektor $\mathbf{x} = [x_1; \ldots; x_n]$ eine Linearkombination der n Spalten der $m \times n$-Matrix $\boldsymbol{A}$ bildet (mit der Koeffizentenfolge $\mathbf{x}$), so ist nun klar dass die Wirkung einer $n \times m$-Matrix $\boldsymbol{B}$ von rechts auf einen Zeilenvektor $\mathbf{y} = [y_1, \ldots, y_n]$ dazu führt, dass wir haben:

$$\mathbf{u} = \mathbf{z} * \boldsymbol{B} = \sum_{k=1}^{n} y_k \mathbf{z}_k \tag{68}$$

veczmat

wobei $\mathbf{z}_1, \ldots, \mathbf{z}_n$ die Zeilenvektoren (alle in $\mathbb{K}^m$) von $\boldsymbol{B}$ sind, d.h. auch die Linearkombination $\mathbf{u}$ dieser Zeilenvektoren liegt in $\mathbb{K}^m$.

Wir haben schon gesehen, dass die Bilder der Einheitsvektoren unter der Matrix-Multiplikation zu den Spalten der agierenden Matrix $\boldsymbol{A}$ werden, d.h. die Wirkung von $\boldsymbol{A}$ gibt dann eine Linearkombination der Spalten. Genaugenommen ist die *Bildmenge*

$$\{\mathbf{y} \,|\, \mathbf{y} = \boldsymbol{A} * \mathbf{x} \quad \text{für [irgend]ein } \mathbf{x} \in \mathbb{R}^n\}$$

genau die Menge der Linearkombinationen von Spalten, d.h. der Spaltenraum $\mathbf{Sp}(\boldsymbol{A})$.

Da die Matrixmultiplikation $\boldsymbol{B} \mapsto \boldsymbol{A} * \boldsymbol{B}$ durch spaltenweise Wirkung von $\boldsymbol{A}$ auf die Spalten von $\boldsymbol{B}$ entsteht, haben wir folgende Beobachtung, gleich kombiniert mit einer entsprechenden Aussage für Zeilenräume (die man leicht mittels Transposition aus der ersten Aussage gewinnt):

paltmatrprod

Lemma 13. *1a) Die Spalten von $\boldsymbol{A} * \boldsymbol{B}$ liegen im Spaltenraum von $\boldsymbol{A}$;*
1b) Der Spaltenraum von $\boldsymbol{A} * \boldsymbol{B}$ ist enthalten im Spaltenraum von $\boldsymbol{A}$;
2a) Die Zeilen von $\boldsymbol{A} * \boldsymbol{B}$ liegen im Zeilenraum von $\boldsymbol{B}$;
2b) Der Zeilenraum von $\boldsymbol{A} * \boldsymbol{B}$ ist im Zeilenraum $\boldsymbol{B}$;

Abgesehen von der Richtigkeit der Aussage ist es auch interessant, welche Sichtweise auf die Matrix-Multiplikation erlaubt. Im ersten Fall entscheiden die Spalten der zweiten Matrix (hier also $\boldsymbol{B}$) welche Linearkombination der Spalten von $\boldsymbol{A}$ in die Produktmatrix $\mathbf{C} := \boldsymbol{A} * \boldsymbol{B}$ an entsprechender Spalte habe. Das Domino-Prinzip (nur Matrizen mit passendem Format können miteinander multipliziert werden!) bedeutet einfach, dass die Zahl der Skalaren in einer Spalte von $\boldsymbol{B}$ der Anzahl der Spalten von $\boldsymbol{A}$ entsprechen muss, damit klar ist, welche Linearkombination der Spalten von $\boldsymbol{A}$ zu bilden ist. Andererseits ist klar, dass $\mathbf{C}$ das Format von einer $m \times n$-Matrix $\boldsymbol{A}$ und einer $n \times k$-Matrix $\boldsymbol{B}$ erbt: es wird eine $m \times k$ Matrix, denn wir bilden mithilfe der Spalten von $\boldsymbol{B}$ genau k Linearkombinationen von Spalten von $\boldsymbol{A}$, und die liegen in $\mathbb{K}^m$, also ist $\mathbf{C}$ eine $m \times k$-Matrix.

Entsprechende Aussagen gelten für die Zeilen. Hier wird klar, dass man $\boldsymbol{A}$ als eine Kollektion von m Zeilenvektoren im $\mathbb{R}^n$ betrachten kann, die durch Rechtsmultiplikation von $\boldsymbol{B}$ Linearkombinationen von Zeilen von $\boldsymbol{B}$ produzieren, und zwar für jede Zeile von $\boldsymbol{A}$ (das sind genau m) wird eine Zeilenvektor im $\mathbb{K}^k$ produziert. Auch das passt zu den allgemeinen Formatvorstellungen!

11.4 Elementar-Matrizen und Gauss Elimination

vgl. Gilbert Strang Buch, Abschnitt 2.3.

elemmatrix

Definition 14. Eine $m \times m$ Matrix $\boldsymbol{E}$ heißt **Elementarmatrix**, wenn sie aus einer $m \times m$ Einheitsmatrix $\mathbf{1}_m$ durch Anwenden einer der drei Typen von elementaren Zeilenoperationen, wie sie in der Gauss-Elimination erlaubt sind, entsteht.

Dementsprechend gibt es drei Typen von Elementarmatrizen, die typischerweise (Beispiele im 3×3-Format) so aussehen:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{69}$$

oder im Fall, dass man vier mal die zweite Zeile zur ersten addiert ($III + 4 \cdot II$):

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix} \tag{70}$$

Es gilt folgende wichtige Aussage:

Proposition 1. *Die Anwendung eines elementaren Schrittes des Gauss-Elimination auf eine $m \times n$-Matrix $\boldsymbol{A}$ ist gleichwertig mit der Links-Matrix-Multiplikation durch die entsprechende Elementarmatrix.*

Obwohl das auch leicht durch Index-Spielerein zu verifizieren ist, ist es wohl instruktiver das Anhand kleiner konkreter Beispiele (oder eines MATLAB Experimentes) zu verifzieren.

Da Elementarmatrizen eine einfache Form haben, ist es auch leicht, die Behauptung durch die Wirkungsweise der Matrix-Multiplikation zu erklären, wie auch immer.

Gilbert Strang verwendet spezielle Symbole für die diversen Elementarmatrizen, z.B. $P_{1,2}$ für das Vertauschen von Zeile 1 mit Zeile 2, oder $E_{3,2}$ wenn die *dritte* Zeile durch Addieren eines Vielfaches von Zeile *zwei* modifizert wird.

Nun brauchen wir noch das Assoziativ-Gesetz der Matrix-Multiplikation (Details folgen hier noch, mit genauer Index-Buchhaltung! Elegantere, aber vielleicht nicht so leicht nachvollziehbare abstrakte Argumente dafür folgen noch später an anderer Stelle!).

$$\boldsymbol{A}, \boldsymbol{B}, \mathbf{C}, \boldsymbol{D}, \boldsymbol{E}, \mathbf{F}, \mathbf{G}$$

Proof. Wir haben zu zeigen, dass für drei Matrizen $\boldsymbol{A}, \boldsymbol{B}, \mathbf{C}$ welche aufgrund ihres Formates zusammenpassen (also eine $m \times n, n \times k$ bzw. eine $k \times l$ Matrix), die Assoziativitätsbedingung

$$(\boldsymbol{A} * \boldsymbol{B}) * \mathbf{C} = \boldsymbol{A} * (\boldsymbol{B} * \mathbf{C}) \tag{71}$$

matr-assoc1

erfüllen. Um dem Beweis eine klarere Struktur zu geben ist es vorteilhaft, für die verschiedenen auftretenden Matrizen eigene Namen zu vergeben. Wir bezeichnen die linke Seite der Gleichung mit $\mathbf{U}$, und die rechte Seite mit $\mathbf{V}$, und setzen weiters $\boldsymbol{E} = \boldsymbol{A} * \boldsymbol{B}$ und $\mathbf{F} = \boldsymbol{B} * \mathbf{C}$, d.h. wir haben zu zeigen, dass $\mathbf{U} = \boldsymbol{E} * \mathbf{C} == \boldsymbol{A} * \mathbf{F} = \mathbf{V}$ gilt. Vom Format her (Domino-Prinzip!) ist alles in Ordnung, denn $\boldsymbol{E}$ ist eine $m \times k$ Matrix, die

mit der $k \times l$ Matrix multipliziert werden kann, ebenso $\boldsymbol{A} * \mathbf{F}$, sodass sowohl $\mathbf{U}$ als auch $\mathbf{V}$ Matrizen vom Format $m \times l$ sind.

Die Hauptproblem für den formalen (allgemeinen) Beweis ist eine vernünftige Wahl der Indizierung. Es ist naheliegend (damit auch die Summierungen über passende Index-Symbole erfolgen kann) die Eintragung der Matrizen mit $(a_{p,q}), (b_{q,r})$ bzw. $(c_{r,s})$ zu bezeichnen, woraus sich in natürlicher Weise die Index-Bezeichnungen $(e_{p,r}), (f_{q,s})$ bzw. $(u_{p,s})$ und $(v_{p,s})$ ergeben.

Einfach der *Definition* des Matrix-Produktes folgend, haben wir

$$e_{p,r} = \sum_{q=1}^{n} a_{p,q} b_{q,r} \quad , \quad f_{q,s} = \sum_{r=1}^{k} b_{q,r} c_{r,s}$$

bzw.

$$u_{p,s} = \sum_{r=1}^{k} e_{p,r} c_{r,s} \quad , \quad v_{p,s} = \sum_{q=1}^{n} a_{p,q} f_{q,s}$$

bzw. nach Einsetzen der ersten Gleichungen

$$u_{p,s} = \sum_{r=1}^{k} \sum_{q=1}^{n} a_{p,q} b_{q,r} c_{r,s} \qquad v_{p,s} = \sum_{q=1}^{n} a_{p,q} \sum_{r=1}^{k} b_{q,r} c_{r,s} = \sum_{q=1}^{n} \sum_{r=1}^{k} a_{p,q} b_{q,r} c_{r,s} \ , \qquad (72)$$ matr-assoc3

woraus leicht erkennbar ist, dass diese beiden Ausdrücke gleich sind, da sie sich lediglich durch eine Umordnung der Summationsreihenfolge (die Summen sind ja endlich!) unterscheiden. Genaugenommen benutzt man auch noch Assoziativität der Multiplikation der Skalare in $\mathbb{K}$. □

11.5 Das Gauss-Verfahren zur Inversion von Matrizen

Nun sind wir bereit, eine **Rezept** für die Inversion von (quadratischen und invertierbaren) $n \times n$-Matrizen $\boldsymbol{A}$ anzugeben.

Man bilde eine erweiterte Matrix der Form $[\boldsymbol{A}, \mathbf{1}_n]$, wende darauf die Gauss-Elimination an, so lange bis man an der Stelle von $\boldsymbol{A}$ die matrix $\mathbf{1}_n$ vorfindet, also eine Matrix der Form $[\mathbf{1}_n, \boldsymbol{B}]$. Die Behauptung ist: Dann ist $\boldsymbol{B}$ die zu $\boldsymbol{A}$ inverse Matrix, d.h.

$$\boldsymbol{A} * \boldsymbol{B} = \mathbf{1}_n = \boldsymbol{B} * \boldsymbol{A}.$$

12 Invertierbare Matrizen

Es ist schnell klar, dass nur quadratische Matrizen invertierbar sind[49] Ein wichtiger Satz ist der folgende:

[49] d.h. es wird schon sinnvollerweise als Teil der Definition vorausgesetzt, obwohl man zeigen kann, dass eine von einer $m \times n$-Matrix induzierte Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ nur dann als Abbildung bijektiv sein kann, wenn $m = n$ ist, und natürlich noch zusätzliche Eigenschaften erfüllt sind.

groupmatr

Theorem 6. *Die Menge aller invertierbaren Matrizen in $\mathcal{M}_{n,n}(\mathbb{R})$ bildet eine Gruppe bezüglich der Matrix-Multiplikation. Sie wird mit dem Symbol $GL(n,\mathbb{K})$ bezeichnet und heißt die* **allgemeine lineare Gruppe über** $\mathbb{K}$[50].

Siehe beispielsweise durchaus auch:

`http://de.wikipedia.org/wiki/Allgemeine_lineare_Gruppe`

An dieser Stelle wäre es durchaus angebracht, ein paar Beispiele von inversen Matrizen zu geben, insbesondere für eine invertierbare Matrix der Form $\boldsymbol{A} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ die inverse Matrix in symbolischer Form (beispielsweise mit unbestimmtem Ansatz, siehe Übungen) zu finden und die Bedingung(en) an die Koeffizienten, welche die Invertierbarkeit garantieren, festzustellen [51]

Wir kommen nur zur Erklärung, warum und in welcher Weise das "Gauss-Rezept", also ein konkreters Verfahren (ein *Algorithmus*) zur Bestimmung der inversen Matrix führt. Rein logisch beschränken wir uns zuerst darauf, zu zeigen, dass eine erfolgreiche Realisierung des Verfahrens (d.h. wenn der Endzustand die Einheitsmatrix $\mathbf{1}_n$ ist) garantiert, dass die Matrix invertierbar ist und dass auf der rechten Seite die (aus gruppentheoretischen Überlegungenen stets *eindeutig bestimmte*) inverse Matrix steht.

Die umgekehrte Richtung, d.h. die Aussage, dass für jede invertierbare Matrix einen *gangbaren Weg* (d.h. eine endliche Folge von Gauss Schritten gibt) gibt, der zu diesem Ziel führt, ist noch später zu diskutieren[52] Wichtiger ist vielleicht [nach dem logischen Prinzip $(A) \Rightarrow (B)$ ist gleichwertig zu $(non-B) \Rightarrow (non-A)$], dass man sicher sein kann (wie gesagt, nach Diskussion dieser Umkehrung) dass eine Matrix *nicht invertierbar ist* wenn eine Durchführung (und wie man zeigen kann auch jeder andere gangbare Weg) zu einer Anzahl $r < n$ von Pivot-Elementen führt.

Ohne die einzelnen Schritte durchzuführen, schauen wir uns ein Beispiel an. Dazu verwenden wir die Matrix $\boldsymbol{A}$ (man beachte dass $\boldsymbol{A}(9) = 0$, im Gegensatz zum sonst oft verwendeten Beispiel, das hat zur Folge - wie wir gleich sehen werden - dass $\boldsymbol{A}$ invertierbar ist!):

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 0 \end{pmatrix} \tag{73}$$

[50] AUF ENGLISCH: General Linear Group over $\mathbb{K}$. Daher die Abkürzung! Für $n > 1$ ist diese Gruppe nicht kommutativ! Die 1×1 Matrizen sind natürlich nichts anderes als die Elemente von $\mathbb{K}$ selber, sodass $GL(1,\mathbb{K})$ mit der multiplikativen Gruppe $\mathbb{K}^*$ aller von Null verschiedenen Elemente von $\mathbb{K}$ identifiziert werden kann.

[51] bekannlich: $ad - bc = \det(\boldsymbol{A}) \neq 0$!

[52] Das ist für die praktische Anwendung nicht so wichtig, wie oft verwendet man ein Elektro-Gerät noch bevor man den Garantieschein ausgefüllt hat. Ausserdem muss gesagt werden, dass diese Überlegung nicht automatisch eine Handlungsanweisung gibt, welcher Schritt wann anzuwenden ist. Andererseits ist bereits jetzt aus der Praxis klar, dass man typischerweise durch Vertauschen von Zeile versuchen wird, Pivotelemente - wenn nötig - in die führende Position zu bringen. In vielen Fällen ist das gar nicht nötig, z.B. im Falle von *diagonal dominanten* Matrizen, wie sie in der Numerischen Mathematik diskutiert werden. Danach erzeugt man durch Multiplizieren mit einer Zahl ungleich Null eine 1, und kann danach in der Spalten der Pivot-Elemente ober- und unterhalb des pivotierenden 1-Elements Nullen erzeugen. Es kommt also im Prinzip nur darauf an zu zeigen, dass man genau n Pivot-Elemente finden kann. Das kommt bald!

Die erweiterte Matrix ist dann von der Form

$$\begin{pmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{74}$$

Unter Anwendung des MATLAB-Befehls `rref` (oder von Hand etc.) kommt man zur reduzierten Zeilenstufenform, d.h.

```
>> AEG = rref([ A, eye(3)]);
```

die folgendes Aussehen hat:

$$AEG = rref([A, eye(3)]) = \begin{pmatrix} 1.0000 & 0.0000 & 0.0000 & -1.7778 & 1.5556 & -0.1111 \\ 0.0000 & 1.0000 & 0.0000 & 0.8889 & -0.7778 & 0.2222 \\ 0.0000 & 0.0000 & 1.0000 & -0.1111 & 0.2222 & -0.1111 \end{pmatrix}$$

Ein Vergleich mit dem eingebauten Befehl zur inversen Matrix mach sicher:

```
>> AI =  AEG(:,4:6)
AI =
   -1.7778    1.5556   -0.1111
    0.8889   -0.7778    0.2222
   -0.1111    0.2222   -0.1111
>> norm( AI - inv(A))
ans =   3.1830e-016  % d.h.\ numerische Uebereinstimmung.
```

Kommen wir nun zum eigentlichen Beweis. Wir erinnern nochmals an folgende Tatsachen betreffend die einzelnen Schritte der Gauss-Elimination:

- jede der Operationen ist invertierbar, durch eine gleichartige Operation;
- jeder Schritt kann als Matrix-Multiplikation von links interpretiert werden, daher auch der Übergang von der Ausgangsmatrix zu irgendeiner Zwischenstufe des Verfahrens, oder auch der Zeilenstufenform;
- jeder dieser Schritte bewahrt den Zeilenraum von $\boldsymbol{A}$, d.h. in *jedem* Schritt des Gauss-Verfahrens ist der Zeilenraum der k-ten Modifikation $\boldsymbol{A}_k$ (beginnend mit $\boldsymbol{A}_0 := \boldsymbol{A}$) gleich dem Zeilenraum des Vorgängers, also ist auch der Zeilenraum von $\boldsymbol{A}$ dem Zeilenraum am Ende der Gauss-Jordan Multiplikation;
- konkret kann man auch beobachten: der Schritt ($III > III + 4 \cdot II$) bedeutet: Addiere das 4-fache der festzuhaltenden zweiten Zeile zur dritten Zeile, als steht in der entsprechenden Elementarmatrix in der *dritten* Zeile die Zahlenfolge bestehend aus lauter Nullen, eine 1 an der dritten Stelle und an der zweiten Stelle eine 4.

Bezeichnen wir die modifizierte Matrix (nach k Schritten der Gauss-Elimination) mit $\boldsymbol{A}_k$, so müssen wir uns zunächst wieder nur um einen einzelnen Schritt kümmern. Wir wenden die k-te Elementarmatrix von links an. Wir bezeichnen den rechten Teil der erweiterten Matrix einfach mit $\boldsymbol{B}_k$, d.h. beginnend mit $\boldsymbol{B}_0 = \mathbf{1}_n$ wollen wir zum Endzustand kommen (l = last step) mit $\boldsymbol{A}_l = \mathbf{1}_n$ und dementsprechend $\boldsymbol{B}_l = \boldsymbol{A}^{-1}$. Das ist die Aussage die wir zu beweisen haben! Für jedes einzelne k haben wir:

$$[\boldsymbol{A}_{k+1}, \boldsymbol{B}_{k+1}] = \boldsymbol{E}_k * [\boldsymbol{A}_k, \boldsymbol{B}_k] \;\; !! \; = [\boldsymbol{E}_k * \boldsymbol{A}_k, \boldsymbol{E}_k * \boldsymbol{B}_k] \tag{75}$$

GaussELINV1

Der entscheidende technische Schritt (mit "!!" markiert) ist die einfache Tatsache, dass die Matrix-Multiplikation spaltenweise auf der rechts stehenden wirkt, und daher in Teilblöcke beliebig zerlegt bzw. zusammengefasst werden kann. Nach drei Schritten steht im rechten Block die Matrix $\boldsymbol{E}_3 * \boldsymbol{E}_2 * \boldsymbol{E}_1$, bzw. allgemein, nach k Schritten die Matrix $\boldsymbol{B}_k = \sqcap_{k=1}^{l} \boldsymbol{E}_k$.

Das Verfahren wird solange fortgeführt, bis wir im vorderen $n \times n$-Block die Einheitsmatrix $\mathbf{1}_n$ stehen haben. Wir bezeichnen den Index der letzten verwendeten Elementar-Matrix mit l. Also haben wir, beginnend mit $k = 0$ und enden mit $k = l - 1$ $\boldsymbol{A}_{k+1} = \boldsymbol{E}_k * \boldsymbol{A}_k$ und $\boldsymbol{A}_l = \mathbf{1}_n$[53]. Also haben wir insgesamt

$$\mathbf{1}_n = \boldsymbol{E}_l * (\boldsymbol{E}_{l-1} * (\cdots * \boldsymbol{E}_1)) * \boldsymbol{A} \tag{76}$$

Ainvprod1

NUN KÖNNEN wir aber das ASSOZIATIVGESETZ anwenden, sowie die (einfache aber wichtige) Tatsache dass ein Produkt von invertierbaren Matrizen auch wieder invertierbar ist, denn es gilt ja

$$(\boldsymbol{A} * \boldsymbol{B})^{-1} = \boldsymbol{B}^{-1} * \boldsymbol{A}^{-1}, \tag{77}$$

matrprodinv

um festzustellen, dass

$$(\boldsymbol{E}_l * \boldsymbol{E}_{l-1} * \cdots * \boldsymbol{E}_1) * \boldsymbol{A} = \mathbf{1}_n$$

gilt, dass also $\sqcap_{k=1}^{l} \boldsymbol{E}_k$ ein Linksinverses zu $\boldsymbol{A}$ ist. Aber genau diese Matrix steht nun auf der rechten Seite der erweiterten Matrix, d.h. $\boldsymbol{B}_l = \sqcap_{k=1}^{l} \boldsymbol{E}_k$ ist eine quadratische Links-Inverse $n \times n$-Matrix zu $\boldsymbol{A}$[54].

Wir lernen aus dem Verfahren folgende Konsequenzen:

1. Wenn das Verfahren bis zu Ende geführt werden kann, d.h. *wenn es gelingt, in der linken Hälfte die Einheitsmatrix herzustellen*, dann *ist* die Matrix $\boldsymbol{A}$ invertierbar. Wir haben ja eine einseitige inverse Matrix $\boldsymbol{B}$ gefunden, die in $GL(n, \mathbb{K})$ liegt, nämlich die Matrix $\boldsymbol{B}$, die ein Produkt von Elementarmatrizen ist. Die sind all invertierbar, also ist auch ihr Produkt, d.i. $\boldsymbol{B}$, invertierbar. Da $\boldsymbol{A}$ eine Rechtsinverse zu $\boldsymbol{B}$ ist, muss $\boldsymbol{A}$ das inverse Element zu $\boldsymbol{B}$ sein, und somit sind $\boldsymbol{A}$ und $\boldsymbol{B}$ beide invertierbar und zueinander invers.

2. Es fehlt noch die Argumentation, warum für jede invertierbare Matrix das Verfahren auch (auf die eine oder andere Art) *auf die vorgegebene Art und Weise* zu einem Erfolg, d.h. zur Bestimmung der inversen Matrix $\boldsymbol{B}$ führt. Das wird daraus

[53] Dann sind wir fertig, das Verfahren war erfolgreich, die letzte Umformung der Ausgangsmatrix $\boldsymbol{A}$ ist die Einheitsmatrix $\mathbf{1}_n$.

[54] Das Symbol $\sqcap$ steht für ein Produkt mit mehreren Faktoren.

folgen, dass der Zeilenzraum von $\boldsymbol{A}$ der ganze $\mathbb{K}^n$ ist (+ Lemma xx). In der Tat, da $\boldsymbol{B} * \boldsymbol{A} = \mathbf{1}_n$ gilt, müssen alle Einheitsvektoren (die Zeilen von $\mathbf{1}_n$) im Bildraum von $\boldsymbol{B} * \boldsymbol{A}$ liegen, der aber liegt bekanntlich im Zeilenraum von $\boldsymbol{A}$ (ist ja in dieser Gleichung der "rechte Faktor").

3. Wir kommen also zum Schluss, dass eine Matrix genau dann invertierbar ist, wenn das Gauss-Eliminationsverfahren (bei passender Wahl der Schritte) zur Einheitsmatrix führt, und die rechte Hälfte der erweiterten Matrix ist dann die (eindeutig bestimmte) inverse Matrix!

4. Vom praktischen Standpunkt aus sei noch ein Hinweis gegeben, wie man vorgehen kann/sollte: Versuchen wir zuerst den ersten Einheitsvektor herzustellen, so müssen wir zunächst *eine* Zeile finden, für die in der ersten Koordinate nicht Null steht. Schreitet man so voran, so kommt man zur Zeilenstufenform (das geht immer). Wir haben aber nun zwei Fälle. Entweder können wir eine Situation herbeiführen, in der wir lauter Einser auf der Hauptdiagonale haben, dann sind wir fertig (subtrahiere passende Vielfache, von unten beginnend, um auf die Form $\mathbf{1}_n$ zu kommen!), *oder* es gibt eine Leerzeile, dann gibt es aber eine Erzeugendensystem des Zeilenraums das aus weniger als n Elementen besteht.[55]

   Wir werden später noch sehen, dass dies garantiert, dass der Zeilenraum von $\boldsymbol{A}$ echt enthalten ist in $\mathbb{K}^n$, und daraus folgt wieder, dass $\boldsymbol{A}$ *nicht invertierbar ist!*.

[55] Der Begriff der Dimension kommt ein wenig später. Natürlich ist die Dimension von $\mathbb{K}^n$ genau n. Wie wir noch sehen werden, ist die Dimension die Minimalzahl von Vektoren eines Erzeugendensystems. Somit kann es nicht sein, dass man den Zeilenraum $\mathbb{K}^n$ mit weniger als n Vektoren aufspannt.

Schauen wir uns ein weiteres, konkretes Beispiel an:

```
>> H = [1,1;1,-1]
H =
     1     1
     1    -1
>> inv(H)
ans =
    0.5000    0.5000
    0.5000   -0.5000
>> HH = [H,eye(2)]
HH =
     1     1     1     0
     1    -1     0     1
HH(2,:) = HH(2,:) - HH(1,:)
HH =
     1     1     1     0
     0    -2    -1     1
>> HH(2,:) = HH(2,:)/-2
HH =
    1.0000    1.0000    1.0000         0
         0    1.0000    0.5000   -0.5000
>> HH(1,:) = HH(1,:) - HH(2,:)
HH =
    1.0000         0    0.5000    0.5000
         0    1.0000    0.5000   -0.5000
```

wobei der rechte Block eben genau die inverse Matrix ist. Wir können uns aber auch die Elementarmatrizen ansehen, die im Laufe der Prozedur verwendet worden sind:

```
>> E = eye(2); E1 = E; E1(2) = - 1
E1 =
     1     0
    -1     1    % subtrahiere I von II
>> E2 = E; E2(4) = 1/-2
E2 =
    1.0000         0
         0   -0.5000  % multipliziere II mit -1/2
>> E3 = E; E3(3) = -1
E3 =
     1    -1      %  I - II
     0     1
>> E3*E2*E1   % das Produkt der Elementar-Matrizen
ans =
    0.5000    0.5000
    0.5000   -0.5000
```

Man beachte, dass dies genau die Koeffizienten sind, die man beim Berechnen von $cos(t)$ bzw. $i \cdot sin(t)$ aus e^{it} und e^{-it} bekommt. Die Koeffizienten $[1, 1; 1, -1]$ waren in der ursprünglichen Eulerschen Formel enthalten!

Daraus folgt aber schon dass $\boldsymbol{B}_l * \mathbf{b}$ eine Lösung des gewüschten Problems ist, d.h. $\mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ ist eine surjektive Abbildung (siehe unten). Weitere Details (die auf besonderen Eigenschaften der Matrizenrechnung beruhen) folgen in Kürze!

Eine weitere Beobachtung kann hilfreich sein, um die Frage zu beantworten, ob man die inverse Matrix bestimmen soll, um die Lösung des (in)homogenen Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ zu lösen, indem man $\mathbf{x} = \boldsymbol{A}^{-1} * \mathbf{b}$ berechnet, oder indem man das Gauss-Eliminations Verfahren einfach auf die erweiterte Systemmatrix $\widetilde{\boldsymbol{A}} = [\boldsymbol{A}, \mathbf{b}]$ anwenden soll. Im Grunde ist das oben beschriebene Konzept gleichwertig zur (systematischen und effizienten, nämlich simultanen) Lösung des inhomogenen Systems $\boldsymbol{A} * \mathbf{x} = \mathbf{e}_k$, für $k = 1, \ldots, n$. [56] Die Antwort liegt auf der Hand. Wenn man oft mit derselben Systemmatrix zu arbeiten hat, oder wenn diese sozusagen "offline" vor dem Eintreffen der Daten (rechte Seite) bestimmt werden kann, dann ist es Vorteilhaft. Hingegen ist es unnötiger Aufwand die ganze inverse Matrix zu bestimmen, wenn man nur *wenige* Gleichungen derselben Bauart zu lösen hat. Natürlich ändert sich die Sichtweise ein wenig, wenn man z.B. MATLAB die Arbeit tun läßt, dann ist die Anwendung der inversen Matrix konzeptuell einfacher.

Aber ist das auch wirklich ein Lösung? Und ist diese eindeutig bestimmt? Ist da noch etwas zu tun? De facto sind beide Fragen (nochmals unter Benutzung der Assoziativität der Matrix-Multiplikation!) einfach zu beantworten[57]

1. Ist $\boldsymbol{A}$ invertierbar, so ist der Vektor $\mathbf{x} := \boldsymbol{A}^{-1} * \mathbf{b}$ wohldefiniert, und es gilt

$$\boldsymbol{A} * \mathbf{x} = \boldsymbol{A} * (\boldsymbol{A}^{-1} * \mathbf{b}) \quad =!! \quad (\boldsymbol{A} *^{-1} \boldsymbol{A}) * \mathbf{b} = \mathbf{b}. \tag{78}$$ Axb-invmat

2. Umgekehrt: Sei $\mathbf{y}$ eine Lösung des Gleichungssytems $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$, dann gilt:

$$\boldsymbol{A} * \mathbf{y} = \mathbf{b} \Rightarrow \boldsymbol{A}^{-1} * (\boldsymbol{A} * \mathbf{y}) = \boldsymbol{A}^{-1} * \mathbf{b} \Rightarrow \mathbf{y} = \boldsymbol{A}^{-1} * \mathbf{b}. \tag{79}$$ Axb-uniq

   Eine nützliche Formel beschreibt das Vertauschen von Transponieren und Konjugieren:

$$(\boldsymbol{A}^{-1})^t = (\boldsymbol{A}^t)^{-1}. \tag{80}$$ mattranspinv

   Der Beweis sei eine Übungsaufgabe!!! (bitte selber probieren!), zuerst anhand kleiner Beispiele, später in allgemeiner Form, d.h. mit Index-Jonglieren, unter Verwendung von (67). matrprodtransp

   Aus $\boldsymbol{B}_l * \boldsymbol{A} = \mathbf{1}_n$ folgt also $\boldsymbol{A}^t * \boldsymbol{B}_l^t = \mathbf{1}_n$, somit könnte man argumentieren, dass der Zeilenraum von $\boldsymbol{B}_l$ der ganze $\mathbb{K}^n$ ist (noch zu sehen, was das bringt...).

[56] Das sollte sich jede/r LeserIn individuell klarmachen, sollte aber sofort klar sein, wenn man sich ansieht, dass Matrix-Multiplikation auf der rechten Seite spaltenweise definiert ist!

[57] ACHTUNG: Diese kurzen! Beweise sind absolut Prüfungsstoff!

12.1 Invertierbarkeitskriterien

Hier geht es weiterhin um eine Behandlung der Frage, wann eine vorgegebene $n \times n$-Matrix $\boldsymbol{A}$ invertierbar ist. Man spricht auch von *nicht-singulären* Matrizen. Die *singulären* Matrizen sind also die *"schlechten"* Matrizen, die einen gewissen Defekt haben. Dieser wird sogar gemessen in Form der Dimension des Nullraums Null($\boldsymbol{A}$).

$$\text{Defekt}(\boldsymbol{A}) := \dim(\text{null}(\boldsymbol{A})). \tag{81}$$

defectdef1

Der folgende Satz gehört zu den wichtigsten Resultaten der Linearen Algebra. Er zeigt auf, dass für die Invertierbarkeit von Matrizen ausreicht, die Surjektivität (Eigenschaft der Spaltenvektoren: sie bilden ein Erzeugendensystem) oder die lineare Unabhängigkeit (die Abbildung $\vec{\mathbf{x}} \mapsto \boldsymbol{A} * \vec{\mathbf{x}}$ ist injektiv) zu überprüfen, und die andere ergibt sich automatisch

Warnung: Diese Äquivalenz hat mit der Bedingung $n = m$ zu tun und ist keineswegs wahr, wenn $n \neq m$ gilt! Andererseits ist die Frage der Invertierbarkeit einer linearen Matrix nicht relevant, wenn $n \neq m$ gilt (weil Räume verschiedener Dimension nie isomorph sein können).

invcrit007

Theorem 7. *Es sei $\boldsymbol{A}$ eine $n \times n$-Matrix über $\mathbb{K}$. Dann sind die folgenden Eigenschaften zueinander äquivalent, d.h. es gilt eine (oder eben nicht), genau dann wenn alle anderen Bedingungen gelten (oder eben alle nicht gelten):*

(i) Die Spalten von $\boldsymbol{A}$, d.h. das System $basA = \{\vec{\mathbf{a}}_1, \cdots, \vec{\mathbf{a}}_n\}$ bilden eine Basis für $\mathbb{K}^n$;

(ii) Die Spalten von $\boldsymbol{A}$ bilden eine Erzeugendensystem für $\mathbb{K}^n$;

(iii) Die Spalten von $\boldsymbol{A}$ sind linear unabhängig;

(iv) Die Gauss-Elimination erlaubt es, durch sukzessives Anwenden von Elementarmatrizen (Matrixmultiplikation von links) auf die Einheitsmatrix umzuformen;

(v) Die Matrix $\boldsymbol{A}$ ist invertierbar;

Remark 1. Obwohl einigen der Querverbindungen auch relativ leicht direkt verifizert werden kann, machen wir einen zyklischen Beweis.

Der Beweis wurde in der Vorlesung vom 21. Mai vorgetragen!

Beispielsweise ist es nicht schwer, die Äquivalenz von (i) und (v) zu verifizieren (Übungsaufgabe). Andererseits ist es nur eine Sache der Definitionen zu verifizieren, dass die gleichzeitige Gültigkeit von (ii) und (iii) äquivalent zur Gültigkeit von (i) ist. Es ist also "nur erstaunlich" (und eine Konsequenz der Gauss-Elimanation, jedenfalls in unserer Beschreibung des Sachverhaltes), dass Surjektivität von $T : \vec{\mathbf{x}} \mapsto \boldsymbol{A} * \vec{\mathbf{x}}$ zur Injektivität dieser Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^n$ äquivalent ist.

Wir haben eine Reihe von unmittelbaren Konsequenzen auf diesem Resultat:

unitaryobs1

Corollary 3. *Eine Matrix $\mathbf{U}$ erfüllt $\mathbf{U}' * \mathbf{U} = \mathbf{Id}_n$ (sie ist orthogonal) genau dann wenn $\mathbf{U} * \mathbf{U}' = \mathbf{Id}_n$ gilt, d.h. wenn für jeder $\vec{\mathbf{x}} \in \mathbb{K}^n$ gilt:*

$$\vec{\mathbf{x}} = \sum_{k=1}^{n} \langle \vec{\mathbf{x}}, \vec{\mathbf{u}}_k \rangle \vec{\mathbf{u}}_k.$$

13 Lineare Abbildungen und Matrizen

Der Begriff der *linearen Abbildung* ist einer der zentralen Begriffe im Rahmen der *Linearen Algebra*.

def-linT **Definition 15.** Es seien $\boldsymbol{V}$ und $\boldsymbol{W}$ Vektorräume über einen festen Körper. Eine Abbildung T von $\boldsymbol{V}$ nach $\boldsymbol{W}$ heißt *linear*, wenn T die Vektorraum-Struktur in folgender Weise respektiert:

1. $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) \quad \forall\ \mathbf{v}_1, \mathbf{v}_2 \in \boldsymbol{V}$;
2. $T(\lambda \cdot \mathbf{v}) = \lambda \cdot T(\mathbf{v}) \qquad \forall \lambda \in \mathbb{K}, \mathbf{v} \in \ \boldsymbol{V}$;

Man beachte, dass hier unterschiedliche Additionen (einmal in $\boldsymbol{V}$ und auf der rechten Seite in $\boldsymbol{W}$) und skalare Multiplikationen genommen werden.

Da wir schon gesehen haben, dass man durch Anwenden der beiden Prinzipien: Multipliziere mit einer Zahl (d.h. sklarare Multiplikation mit einem $\lambda \in \mathbb{K}$) bzw. durch Addition sukzessive beliebige Linear-Kombinationen aufbauen kann, ist folgendes harmlose (aber praktische sehr nützliche) Lemma sehr hilfreich:

lincomblemma **Lemma 14.** *Eine Abbildung T von $\boldsymbol{V}$ nach $\boldsymbol{W}$ ist* **genau dann linear**, *wenn sie Linearkombinationen respektiert, d.h. wenn gilt für beliebige Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_l$ und entsprechende Skalare $\lambda_k, 1 \leq k \leq l$:*

$$T\left(\sum_{k=1}^{l} \lambda_k \mathbf{v}_k\right) = \sum_{k=1}^{l} \lambda_k T(\mathbf{v}_k) \tag{82}$$ lincombresp

Proof. Zuerst einmal ist klar, dass die Gültigkeit von (82) die Linearität impliziert, denn klarerweise sind gewöhnlich Summen auch Linear-Kombinationen, ebenso ist $\lambda\mathbf{v}$ eine spezielle Linear Kombination mit genau einem Summanden. In anderen Worten, die Linearität ist eine notwendige Folge des Respektierens von Linear-Kombinationen.

Es geht also um die Umkehrung. Der Beweis erfolgt durch Induktion, nach der Länge der Linear-Kombination. Haben wir nur einen Summanden, so ist die Eigenschaft (82) eine Folge von Forderung ii) in der Definition der Linearität.

Setzt man die Gültigkeit für Linear-Kombinationen bis zur Länge l_0 voraus (Induktionsvoraussetzung), so kann man leicht mit Hilfe von Eigenschaft (i) in der Definition der Linearität auf Linear-Kombinationen der Länge $l_0 + 1$ schliessen, denn es gilt:

$$T\left(\sum_{k=1}^{l_0+1} \lambda_k \mathbf{v}_k\right) = T\left(\sum_{k=1}^{l_0} \lambda_k \mathbf{v}_k\right) + \lambda_{l_0+1} \cdot T(\mathbf{v}_{l_0+1}).$$

Aufgrund der Verwendung der Induktionsvoraussetzung kann man den Term mit $\sum_{k=1}^{l_0}$ durch eine Linearkombination von Bild-Elementen unter T, also der Vektoren $T(\mathbf{v}_k)$, mit $1 \leq k \leq n_0$ darstellen, woraus sich wie gewuenscht für die rechte Seite ergibt:

$$T\left(\sum_{k=1}^{l_0} \lambda_k \mathbf{v}_k\right) + \lambda_{l_0+1} \cdot T(\mathbf{v}_{l_0+1}) = \sum_{k=1}^{l_0} \lambda_k T(\mathbf{v}_k) + \lambda_{l_0+1} \cdot T(\mathbf{v}_{l_0+1}) = \sum_{k=1}^{l_0+1} \lambda_k T(\mathbf{v}_k).$$

□

Einer der wichtigsten Sätze (wenn auch unscheinbar wirkend), der die Rolle der Matrizenrechnung innerhalb der linearen Algebra beschreibt, ist der folgende Satz:

LinAbbRnRm **Theorem 8.** *Jede Abbildung von der Form* $\mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ *von* $\mathbb{R}^n$ *nach* $\mathbb{R}^m$ *ist linear, d.h. respektiert Linear-Kombinationen.*

Umgekehrt, ist jede lineare Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ *durch eine eindeutig bestimmte Matrix* $\boldsymbol{A}$ *auf die oben angegebene Weise (durch Matrixmultiplikation von links*[58]*) möglich.*

Die zu einer Abbildung T *gehörige Matrix* $\boldsymbol{A}$ *hat als Spalten* $(\mathbf{a}_k)$ *einfach die Bilder der Einheitsvektoren* $(\mathbf{e}_k)_{k=1}^n$ *in* $\mathbb{R}^m$*.*

Wir werden später noch einiges zur Umkehrung hören, d.h. wann und wie es möglich ist, durch eine Matrix eine lineare Abbildung zu definieren (siehe Abschnitt Dualräume).

Proof. Der Beweis hat zwei Richtungen. Erstens, ist zu zeigen (einfache Übungsaufgabe, Rechnen mit Matrizen), dass die Matrix-Multiplikation die entsprechenden Eigenschaften hat.

Umgekehrt, sei T eine (abstrakte) lineare Abbildung, das heißt, eine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ welche die Eigenschaft (82) hat, d.h. Linearkombinationen respektiert. Da wir genau n Einheitsvektoren in $\mathbb{R}^n$ vorfinden, deren Bilder alle in $\mathbb{R}^m$ liegen, kann man leicht eine Matrix $\boldsymbol{A}$ definieren, indem man einfach für jedes $k, 1 \leq k \leq n$ das Bild $T(\mathbf{e}_k)$ in die $k-$te Spalte schreibt, also $\mathbf{a}_k := T(\mathbf{e}_k)$. Das ergibt die gewünschte Matrix, den klarerweise stimmen die beiden Abbildungen, d.h. die gegebene T und die durch $\boldsymbol{A}$ definierte Abbildung $S : \mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ auf den Einheitsvektoren überein, und damit überall, denn

lincombresp

$$T(\mathbf{x}) = T\left(\sum_{k=1}^{n} x_k \mathbf{e}_k\right) = \sum_{k=1}^{n} x_k T(\mathbf{e}_k) = \sum_{k=1}^{n} x_k \mathbf{a}_k = \boldsymbol{A} * \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n.$$

□

Das zuletzt verwendete Argument kann leicht benutzt werden, um folgende allgemeinere Aussage zu beweisen:

identlinabb **Lemma 15.** *Es seien* T_1 *und* T_2 *zwei lineare Abbildungen von* $\boldsymbol{V}$ *nach* $\boldsymbol{W}$*, welche auf einem Erzeugendensystem* M *übereinstimmen, d.h. es gelt (nur)* $T_1(\mathbf{m}) = T_2(\mathbf{m}), \mathbf{m} \in \mathbf{M}$*. Dann stimmen sie für alle* $\mathbf{v} \in \boldsymbol{V}$ *überein, d.h.* $T_1 = T_2$ *(im Sinne von Abbildungen, also* $T_1(\mathbf{v}) = T_2(\mathbf{v}), \forall \mathbf{v} \in \boldsymbol{V}$*.).*[59]*.*

[58] Es gäbe im Prinzip einen gleichwertigen Satz, mit Matrix-Multiplikation von rechts, wenn man die Vektoren $\mathbf{x}$, die ja nur geordnete n-Tupel sind, nicht als Spaltenvektoren, sondern als Zeilenvektoren auffassen würde. Der Übergang von dem einen Setting in das andere ist natürlich ganz einfach durch Transposition möglich! Ein Vorteil von dieser Sichtweise wäre, dass dann die entsprechende Matrix vom Format $n \times m$, d.h. der erste Parameter wäre die Dimension des Def. Bereiches und der zweite $(= m)$ wäre dann die Dimension des Zielbereiches.

[59] Erst später werden wir sehen, dass es im Falle einer linear unabhängigen Menge möglich ist, eine Abbildung einfach auf einer Basis vorzugeben

In Wirklichkeit gilt sogar noch mehr: Man hat nicht nur eine (nicht-konstruktive) Eindeutigkeitsaussage, sondern kann das Ergebnis direkt verifizieren. Es muss (aufgrund der Linearität) das Bild von $\mathbf{v}$ als Linearkombination der Bilder $T(\mathbf{m}), m \in M$ schreibbar sein (> gute Übung zum selber Überlegen!).

Es ist eine gute Übungsaufgabe (!Achtung, ebenfalls naheliegender Prüfungsstoff) folgende Aussagen zu verifizieren:

linabbfacts **Lemma 16.** *1. Die Komposition von linearen Abbildungen ist wieder linear;*

2. Ist eine lineare Abbildung bijektiv, dann ist auch die Umkehrabbildung linear;

3. Die Linearkombination von linearen Abbildungen ist linear, d.h. sind T_1 und T_2 lineare Abbildungen, dann ist auch ihre Linearkombination $\lambda_1 T_1 + \lambda_2 T_2$ ebenfalls eine lineare Abbildung

Proof. Die Linearität der Umkehrabbildung folgt recht einfach[60] aus folgender Überlegung. Wir führen nur die Verträglichkeit mit der Addition vor:

Gegeben seien also $\mathbf{w}_1, \mathbf{w}_2 \in \boldsymbol{W}$. Da T bijektiv ist, gibt es genau entsprechende Elemente $\mathbf{v}_i = T^{-1}(\mathbf{w}_i)$, also mit $T(\mathbf{v}_i) = \mathbf{w}_i, i = 1, 2$. Die Linearität von T (direkte Richtung) impliziert natürlich, dass

$$\mathbf{w}_1 + \mathbf{w}_2 = T(\mathbf{v}_1) + T(\mathbf{v}_2) = T(\mathbf{v}_1 + \mathbf{v}_2)$$

gilt. Wendet man auf beide Seiten diese Gleichung die Umkehr-Abbildung T^{-1} an, so bekommt man:

$$T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = T^{-1}(T(\mathbf{v}_1 + \mathbf{v}_2)) = \mathbf{v}_1 + \mathbf{v}_2 = T^{-1}\mathbf{w}_1 + T^{-1}\mathbf{w}_2.$$

Analog geht man mit der skalaren Multiplikation um.[61] MORE TO BE DONE HERE! □

HIER ist eine genauere/ausführliche Version der Aussage:

13.1 Inverse Abbildung und inverse Matrix

invlinmap1 **Lemma 17.** *Es sei T eine lineare Abbildung von einem Vektorraum $\boldsymbol{V}$ in sich selbst, welche invertierbar ist, mit inverser Abbildung T^{-1}. Dann ist auch T^{-1} linear!*

Proof. Gegeben zwei Elemente $\mathbf{w}_i \in \boldsymbol{V}$, welche als Bilder von anderen Vektoren $\mathbf{v}_i$ unter T sind, d.h. $\mathbf{w}_i = T(\mathbf{v}_i), i = 1, 2$. Weiters seien $\lambda_1, \lambda_2 \in \mathbb{K}$ Skalare. Wir interessieren uns für das Bild der allgemeinen Linearkombination $\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2$ unter T^{-1}.

Die folgende Kette von Gleichungen benützt nur die Tatsache dass $T^{-1} \circ T = Id$ ist und dass T eine lineare Abbildung ist:

$$T^{-1}(\lambda_1\mathbf{w}_1 + \lambda_2\mathbf{w}_2) = T^{-1}(\lambda_1 T\mathbf{v}_1 + \lambda_2 T\mathbf{v}_2) = T^{-1}(T(\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2)) = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2,$$

[60] Trotzdem, oder gerade deshalb ist es sehr empfehlenswert, sich den Beweis gut einzuprägen, und zu analysieren, wie die Voraussetzungen und Begriffe zusammenspielen.

[61] Im Prinzip kommt es darauf an, dass $T^{-1}\mathbf{w}_1 + T^{-1}\mathbf{w}_2$ ein Element ist, dessen Bild aufgrund der Linearität gleich $\mathbf{w}_1 + \mathbf{w}_2$ ist, sodass es aufgrund der Eindeutigkeit gleich sein muss $T^{-1}(\mathbf{w}_1 + \mathbf{w}_2)$, dessen Bild unter T trivialerweise gerade $\mathbf{w}_1 + \mathbf{w}_2$ ergibt!

aber der letzte Ausdruck ist ja gerade $\lambda_1 T^{-1}(\mathbf{w}_1) + \lambda_2 T^{-1}(\mathbf{w}_2)$, w.z.z.w..

Daraus folgt insbesondere, dass T^{-1} mit Summen ($\lambda_1 = 1 = \lambda_2$) bzw. Skalarprodukten ($\lambda_1 = \lambda, \lambda_2 = 0$) verträglich.

In kompakterer Form: Betrachte Linear-Kombinationen der Elementen $\mathbf{w}_k, 1 \leq k \leq n$, also Ausdrücke der Form $\sum_{k=1}^{n} \lambda_k \mathbf{w}_k$, mit $T(\mathbf{v}_k) = \mathbf{w}_k, 1 \leq k \leq n$. Dann gilt:

$$T^{-1}[\sum_{k=1}^{n} \lambda_k \mathbf{w}_k] = T^{-1}[\sum_{k=1}^{n} \lambda_k T(\mathbf{v}_k)] = T^{-1}[T(\sum_{k=1}^{n} \lambda_k \mathbf{v}_k)] = \sum_{k=1}^{n} \lambda_k \mathbf{v}_k = \sum_{k=1}^{n} \lambda_k T^{-1} \mathbf{w}_k.$$

□

13.2 Matrix Multiplikation und Zusammensetzung

Nun zeigen wir, dass die Matrix-Multiplikation gerade so gemacht ist, dass sie mit der Komposition von linearen Abbildungen verträglich ist!

composlin1

Theorem 9. *Es seien zwei lineare Abbildungen $T_2 : \mathbb{R}^p \to \mathbb{R}^n$ und $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ gegeben. Die zugehörigen Matrizen seien $\boldsymbol{B}$ (Format $n \times p$) bzw. $\boldsymbol{A}$ (Format $m \times n$). Dann ist die Matrix $\mathbf{C}$ vom Format $p \times m$, welche die zusammengesetzte lineare (!) Abbildung $T := T_1 \circ T_2$ von $\mathbb{R}^p$ nach $\mathbb{R}^m$ repräsentiert also mit $T(\mathbf{v}) = T_1(T_2(\mathbf{v}))$ durch das Matrix-Produkt $\mathbf{C} = \boldsymbol{A} * \boldsymbol{B}$ gegeben.*

Proof. Wir müssen die Matrix $\mathbf{C}$ (Format ist OK) bestimmen, indem wir die Bilder der Einheitsvektoren, also $T(\mathbf{e}_k)$ in die k-te Spalte von $\mathbf{C}$ schreiben. Weil T_2 zuerst zur Anwendung kommt (obwohl die Matrix $\boldsymbol{A}$ links steht!) betrachten wir zunächst das Bild von $\mathbf{e}_k$ unter T_2 ($1 \leq k \leq p$), welches (so ist ja $\boldsymbol{B}$ bestimmt worden) eben genau $\vec{\mathbf{b}}_k$ ist, die k-te Spalte von $\boldsymbol{B}$. Somit müssen wir nur noch $T_1(\vec{\mathbf{b}}_k)$ bestimmen, also $\boldsymbol{A} * \vec{\mathbf{b}}_k$. Aber der Definition von $\boldsymbol{A} * \boldsymbol{B}$ ist das genau die k-te Spalte dieser Produkt-Matrix. Kurz:

$$T(\mathbf{e}_k) = T_1\left(T_2(\mathbf{e}_k)\right) = T_1(\boldsymbol{B} * \mathbf{e}_k) = T_1(\mathbf{b}_k) = \boldsymbol{A} * \mathbf{b}_k = \mathbf{c}_k$$

für die Produktmatrix $\mathbf{C} = \boldsymbol{A} * \boldsymbol{B}$, definiert wie üblich durch die Formel $\mathbf{c}_k := \boldsymbol{A} * \mathbf{b}_k$. □

Als Folgerung haben wir: Die Matrix, die der inversen linearen Abbildung entspricht es *genau* die inverse Matrix, da die Matrix für $\mathbf{Id}_n$ die Einheitsmatrix ist.

Als Begründung kann man angeben: Da T^{-1} eine lineare Abbildung ist (wie wir gerade gesehen haben), gibt es für T^{-1} eine Matrix. Die Komposition von T mit T^{-1} ist aber (in beiden Reihenfolgen!) die identische Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$, also muss dass Matrix Produkt, das ja die beiden zusammengesetzen Abbildungen darstellt, jeweils die Matrix der identischen Abbildung sein, d.h. die Einheitsmatrix. Wir haben also

$$\boldsymbol{A} * \boldsymbol{B} = \mathbf{I}_n = \boldsymbol{B} * \boldsymbol{A},$$

aber das ist genau die definierende Eigenschaft der inversen Matrix! *Somit ist es legitim und sinnvoll, anstatt des Symbols $\boldsymbol{B}$ ab sofort das Symbol $\boldsymbol{A}^{-1}$ zu verwenden!*

Wir können daraus auch nochmals eine (weitere) Rechtfertigung für das Berechnen der inversen Matrix mittels Gauss-Elimination ziehen. Wenn die inverse Matrix $\boldsymbol{B}$ existiert,

dann sind die Spalten der Bilder von $\mathbf{e}_k$ unter $\boldsymbol{B}$, aber $\boldsymbol{A}^{-1}(\mathbf{e}_k) = \mathbf{b}_k$ ist offenbar gleichbedeuten (Matrix Multiplikation mit $\boldsymbol{A}$ ist eine bijektive Abbildung) mit der Aussage dass $\boldsymbol{A} * \mathbf{b}_k = \mathbf{e}_k$. Mit anderen Worten, die k-te Spalte von $\boldsymbol{A}^{-1} = \boldsymbol{B}$ ist nichts anderes als die eindeutige bestimmte Lösung des linearen Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{e}_k$. Wenn man also dieses Gleichungssystem *"auf einmal für alle Einheitsvektoren"* lösen will, muss man die Kollektion der Einheitsvektoren, d.h. die Matrix $\mathbf{I}_n$ an die rechte Seite schreiben und für jedes $k, 1 \leq k \leq n$ nach $\mathbf{x}$ auflösen. Aus Gründen der Ökonomie macht man es aber natürlich auf einmal, weil die notwendigen Schritte in der Gauss-Elimination ja nur von der Matrix $\boldsymbol{A}$ abhängen, und nicht von der rechten Seite. Damit haben wir also das bekannte Rezept zur *Berechnung der inversen Matrix* mit Hilfe der Gauss-Jordan Elimination nochmals gerechtfertigt.

Als kleinen Zusatznutzen der obigen Überlegungen kann man anführen, dass es auf diese Weise auch einzelen SPALTEN der inversen Matrix berechnet werden können!

Weitere Anmerkung: Man kann natürlich auch einzelne Zeilen der inversen Matrix[62] bestimmen, indem man die Spalten der inversen Matrix von $\boldsymbol{A}^t$ berechnet.

ENDE DES EINSCHUBES, APRIL 2014

Der letzte Teil der Aussage kann leicht so zusammengefasst werden.

linabbspac **Lemma 18.** *Für je zwei Vektorräume $\boldsymbol{V}$ und $\boldsymbol{W}$ ist die Menge der linearen Abbildungen (normalerweise mit $\mathcal{L}(\boldsymbol{V}, \boldsymbol{W})$ bezeichnet) selbst wieder ein linearer Raum bzgl. der üblichen Linearkombinationsbildung (elementweise).*

linisom **Definition 16.** Eine invertierbare lineare Abbildung T von $\boldsymbol{V}$ nach $\boldsymbol{W}$ definiert einen **Isomorphismus** zwischen $\boldsymbol{V}$ und $\boldsymbol{W}$. Zwei Vektorräume (über demselben Körper) heissen **isomorph**, wenn es einen solchen Isomorphismus gibt.

Da die Abbildung T^{-1} ebenfalls linear ist, ist diese Definition gleichwertig mit der Existenz einer linearen Abbildung T sowie einer zu T inversen linaren Abbildung S von $\boldsymbol{W}$ nach $\boldsymbol{V}$ gibt, also zur Existenz von zwei linearen Abbildungen T, S welche folgende Identitäten erfüllen:

$$T \circ S = \mathbf{1}_{\boldsymbol{W}} \quad \text{and} \quad S \circ T = \mathbf{1}_{\boldsymbol{V}}. \tag{83}$$ compiso1

Im Falle eines Isomorphismus von $\boldsymbol{V}$ zu sich selbst spricht dann von **Automorphismen von $\boldsymbol{V}$** ist das folgende Lemma:

GLnRgroup **Lemma 19.** *Die Menge der Automorphismen eines Vektorraums $\boldsymbol{V}$ in sich, kurz geschrieben $\mathcal{A}ut(\boldsymbol{V})$, ist eine Gruppe bezüglich Komposition. Das zu T inverse Element in dieser Gruppe ist die in Def. 16 auftauchende lineare Abbildung S. Das neutrale Element ist die identische Abbildung: $T_0 = Id_{\boldsymbol{V}} : \mathbf{v} \mapsto \mathbf{v}$, $\mathbf{v} \in \boldsymbol{V}$.*

linisom

Proof. Wir haben nicht viel zu beweisen, ausser dass z.B. die Komposition zweier Automorphismen wieder ein Automorphismus ist. In der Tat ist die Komposition linearer Abbildung selber wieder linear, denn

$$T_2(T_1(\mathbf{v}_1 + \mathbf{v}_2)) = T_2(T_1(\mathbf{v}_1) + T_1(\mathbf{v}_2)) = T_2(T_1(\mathbf{v}_1)) + T_2(T_1(\mathbf{v}_2))$$

[62] Diese bilden insgesamt ein Biorthogonal-System zum System der Spalten von $\boldsymbol{A}$, aber das wird erst später bzw. anderswo genauer diskutiert!

etc., und klarerweise ist $S_2 \circ S_1$ die zugehörige inverse lineare Abbildung. Da diese (beide) linear und bijektiv sind, ist klar, dass diese inversen Elemente selbst wieder Automorphismen bilden, etc.. □

Beispiele von konkreten linearen Abbildungen folgen im nächsten Abschnitt. Dann erweist sich die oben gemachte Aussage als gleichwertig zur Aussage: Das Produkt von zwei invertierbaren $n \times n$-Matrizen ist wieder eine invertierbare $n \times n$-Matrix (und das ist ja eine bekannt/einfache Aussage, siehe Formel (77)). matrprodinv

Ein weitere und ebensfalls grundlegende Eigenschaft, die man als Begründung für die Form der Definition der Matrix-Multiplikation sehen kann, ist durch den folgenden Satz gegeben.

matrix-comp1 **Theorem 10.** *Seien T_1 und T_2 zwei lineare Abbbildung, die eine von $\mathbb{R}^n$ nach $\mathbb{R}^m$, die andere von $\mathbb{R}^m$ nach $\mathbb{R}^k$, und seien $\boldsymbol{A}_1$ und $\boldsymbol{A}_2$ die entsprechenden Matrizen, die also vom Format $m \times n$ bzw. $k \times m$ sind. Dann ist die zu $T_2 \circ T_1$ gehörige Abbbildung durch die $k \times h$-Matrix $\boldsymbol{A}_2 * \boldsymbol{A}_1$ gegeben*[63].

Proof. Es ist ganz einfach, wenn man bedenkt, dass die Matrix $\mathbf{C}$ zu $T_2 \circ T_1$ als Spalten die Bilder der Einheitsvektoren unter dieser zusammengesetzten Abbildung enthält. Dazu muss man die Bilder der $\mathbf{e}_k$ unter T_1, das *sind aber genau die Spalten von* $\boldsymbol{A}_1$! unter T_2 Abbilden, d.h. aber genau, dass man die zugehörige Matrix auf jeden dieser Vektoren abbilden muss. Das ist aber gerade die übliche Definition der Multiplikation von Matrizen. Fasse den rechten Faktor (hier $\boldsymbol{A}_1$) als eine Kollektion von Spaltenvektoren auf, und wende darauf einzeln die Matrix $\boldsymbol{A}_2$ an. Kurz, $\mathbf{C} = \boldsymbol{A}_2 \circ \boldsymbol{A}_1$. □

DASSELBE IN ALTERNATIVER SCHREIBWEISE, ANGEPASST AN DEN BEWEIS DES ASSOZIATIVGESETZES: MARCH 2014:

composlin1 **Theorem 11.** *Es seien zwei lineare Abbildungen $T_2 : \mathbb{R}^p \to \mathbb{R}^n$ und $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ gegeben. Die zugehörigen Matrizen seien $\boldsymbol{B}$ (Format $n \times p$) bzw. $\boldsymbol{A}$ (Format $m \times n$).*

Dann ist die Matrix $\mathbf{C}$ vom Format $p \times m$, welche die zusammengesetzte lineare (!) Abbildung $T := T_1 \circ T_2$ von $\mathbb{R}^p$ nach $\mathbb{R}^m$ repräsentiert also mit $T(\mathbf{v}) = T_1(T_2(\mathbf{v}))$ durch das Matrix-Produkt $\mathbf{C} = \boldsymbol{A} * \boldsymbol{B}$ gegeben.

Proof. Wir müssen die Matrix $\mathbf{C}$ (Format ist OK) bestimmen, indem wir die Bilder der Einheitsvektoren, also $T(\mathbf{e}_k)$ in die k-te Spalte von $\mathbf{C}$ schreiben. Dazu betrachten wir zunächst das Bild von $\mathbf{e}_k$ unter T_1 ($1 \leq k \leq p$), welches (so ist ja $\boldsymbol{B}$ bestimmt worden) eben genau $\vec{\mathbf{b}}_k$ ist, die k-te Spalte von $\boldsymbol{B}$. Somit müssen wir nur noch $T_1(\vec{\mathbf{b}}_k)$ bestimmen, also $\boldsymbol{A} * \vec{\mathbf{b}}_k$. Aber der Definition von $\boldsymbol{A} * \boldsymbol{B}$ ist das genau die k-te Spalte dieser Produkt-Matrix.

Kurzversion

$$T(\mathbf{e}_k) = T_1(T_2(\mathbf{e}_k)) = T_1(\boldsymbol{B} * \mathbf{e}_k) = T_1(\mathbf{b}_k) = \boldsymbol{A} * \mathbf{b}_k = \mathbf{c}_k$$

[63] Man beachte, dass die Wirkung durch Links-Hinzuschreiben kommt, daher sind notwendigerweise die Matrizen in der umgekehrten Reihenfolge hinzuschreiben als die Abbildungen angewendet werden, nämlich T_1 zuerst, und dann T_2.

für die Produktmatrix $\mathbf{C} = \boldsymbol{A} * \boldsymbol{B}$, also gilt nach dem allgemeine Darstellungsprinzip

$$T(\mathbf{x}) = \mathbf{C} * \mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^p.$$

□

Für die Bestimmung einer Basis für den Spaltenraum ist es nützlich, folgendes allgemeine Prinzip zu verifizieren:

isoprops1 **Lemma 20.** *Es sei T ein Isomorphismus zwischen zwei Vektorräumen $\boldsymbol{V}$ und $\boldsymbol{W}$, mit inversem Isomorphismus $S = T^{-1}$. Dann gelten folgende Aussagen:*

1. *die Menge M ist linear unabhängig (bzw. abhängig) $\Leftrightarrow$ $T(M)$ ist linear unabhängig (bzw. abhängig);*
2. *eine Menge $M \subset \boldsymbol{V}$ ist ein Erzeugendensystem von $\boldsymbol{V}$ $\Leftrightarrow$ $T(M)$ ist ein Erzeugendensystem von $\boldsymbol{W}$;*
3. *M ist eine Basis für $\boldsymbol{V}$ $\Leftrightarrow$ $T(M)$ ist eine Basis für $\boldsymbol{W}$.*

Als Korollar hat man also sofort: Hat $\boldsymbol{V}$ eine Basis mit n Elementen, dann muss auch $\boldsymbol{W}$ eine Basis mit n Elementen haben. Würde man (durchaus zulässigerweise) die *Dimension* von $\boldsymbol{V}$ bzw. $\boldsymbol{W}$ als die Minimalzahl benötigter Elemente für ein Erzeugendensystem definiert haben, könnte man schon an dieser Stelle schließen, dass zwei Ismorphe Räume gleiche Dimension haben. Es gilt aber auch die Umkehrung (siehe später).

Ebenso ist folgendes Lemma ganz einfach zu beweisen (Übung!)

idLinAbb **Lemma 21.** *Es seien T_1 und T_2 zwei lineare Abbildungen, und M ein Erzeugendensystem. Wenn $T_1(m) = T_2(m)$ für alle $m \in M$ gilt, dann sind die beiden Abbildungen identisch, d.h. $T_1(\mathbf{v}) = T_2(\mathbf{v})$ für alle $\mathbf{v} \in \boldsymbol{V}$.*

Dieses unscheinbare Lemma kann dazu dienen, den wichtigen Darstellungssatz für lineare Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ zu beweisen.

LinDarst0 **Theorem 12.** *Zu jeder linearen Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ gibt es eine eindeutig bestimmte Matrix $\boldsymbol{A}$, sodass $T(\mathbf{x}) = \boldsymbol{A} * \mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$ gilt. Die Spalten $\mathbf{a}_1, \ldots, \mathbf{a}_n$ aus $\mathbb{R}^m$ dieser Matrix $\boldsymbol{A}$ sind genau die Bilder der Einheitsvektoren $\mathbf{e}_1, \ldots, \mathbf{e}_n$ unter T.*

Proof. Wir wissen bereits, dass für jede $m \times n$-Matrix $\boldsymbol{A}$ die Abbildung $T_2 : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ von $\mathbb{R}^n$ nach $\mathbb{R}^m$ linear ist. Wir wissen auch, dass $\boldsymbol{A} * \mathbf{e}_k = \mathbf{a}_k$, der k-te Spaltenvektor von $\boldsymbol{A}$ ist.

Ist also eine beliebige linear Abbildung $T : \mathbb{R}^n \to \mathbb{R}^m$ gegeben, so ist es naheliegend, die Matrix $\boldsymbol{A}$ (die dann T *darstellen* soll) so zu wählen, dass $\mathbf{a}_k := T(\mathbf{e}_k)$ gilt (diese Wahl ist durch T eindeutig bestimmt, und in der Tat auch eindeutig (d.h. verschiedene Abbildungen T müssen wegen Lemma (21) idLinAbb oben auch auf mindestens einem der Einheitsvektoren (die ja eine Basis für $\mathbb{R}^n$ bilden) verschieden sein.

Die Abbildung $T_1 : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ für die soeben definierte Matrix $\boldsymbol{A}$ ist dann klarerweise mit T auf den Einheitsvektoren $\mathbf{e}_k, 1 \leq k \leq n$ gleich, also nach Lemma (21) idLinAbb als Abbildungen identisch, was zu zeigen war. □

Bemerkung: Man könnte die Aussage des vorhergehenden Satzes noch ein wenig ausschmücken und davon sprechen, dass die Abbildung $T \mapsto \boldsymbol{A}$ (der Abbildung T wird in eindeutig und umkehrbarer Weise eine Matrix $\boldsymbol{A}$ zugeordnet) selbst wieder ein Isomorphimus ist, zwischen dem Vektorraum all dieser linearen Abbildungen (z.B. mit Summe $(T+S)(\mathbf{x}) = T(\mathbf{x}) + S(\mathbf{x})$, etc.) und dem Vektorraum aller $m \times n$-Matrizen, mit der üblichen "koordinatenweisen" Addition und skalaren Multiplikation.

Eine andere interessante Konsequenz ist folgendes Lemma:

isodim1 **Lemma 22.** *Es gebe einen Isomorphismus von $\mathbb{R}^n$ nach $\mathbb{R}^m$, dann gilt $m = n$.*

Proof. Indirekte Annahme: Es sei $n \neq m$. O.B.d.A. gelte $n > m$ (im anderen Falle betrachtet man den inversen Isomorphismus). Die zugehörige Matrix hat dann aber echt mehr Zeilen als Spalten, somit kann es aber nur $m < n$ Pivot-Elemente geben, es gibt also mindestens eine freie Variable, also haben wir ein Problem mit der Injektivität diese Isomorphismus! Dieser Widerspruch erzwingt[64] $m = n$. □

compisom01 **Lemma 23.** *Die Komposition von Isomorphismen (also natürlich auch von Automorphismen) ist wieder ein Isomorphismus (und die Inversen haben in der umgekehrten Reihenfolge angewendet zu werden).*

Der Beweis kann den LeserInnen überlassen werden. Die Gruppe der *Automorphismen des* $\mathbb{R}^n$ wurde schon mit der $GL(n, \mathbb{R})$ identifiziert (es ist einfach die Gruppe der invertierbaren, reellen $n \times n$ Matrizen).

Eine weitere Konsequenz, und in der Tat ein wichtiges *praktisches Kriterium* zur Überprüfung der linearen (Un-) Abhängigkeit ist:

linabhsp **Lemma 24.** *Es sei $J \subset \{1, \ldots, n\}$ eine beliebige Kollektion von Spalten-Indices. Die Menge der entsprechenden Spalten von $\boldsymbol{A}$, d.h. $(\mathbf{a}_j)_{j \in J}$ ist linear unabhängig genau dann wenn die entsprechende Kollektion von Spaltenvektoren $(g_j)_{j \in j}$ der entsprechenden Gauss'schen Zeilenstufenform linear unabhängig sind.*

Andererseits ist eine Vektor $\mathbf{b}$ von einer Familie von Spaltenvektoren genau dann linear abhängig, wenn die Anwendung der Gauss-Elimination auf die erweiterte Systemmatrix zeigt, dass die letzte Spalte linear abhängig ist, also kein neues Pivot-Element erzeugt.

Daraus leitet man die Basis-Identifizierung ab (eine Basis ist ein maximal linear unabhängiges System in einem [endlichdim.] Vektorraum).

Proof. Es ist nur mehr wichtig festzustellen, dass die Kollektion der Spalten, in denen die Pivotelemente stehen, diese Eigenschaft hat. Dass sie eine linear unabh. Kollektion von Vektoren im $\mathbb{R}^m$ bilden, ist einfach zu sehen (genauso wie man es beim Zeilenraum gemacht hat).

Bleibt also nur mehr die Überprüfung der *Maximalität*, d.h. man muss nur zeigen, dass die Vektoren, *die sich hinter den Pivot-Spalten verstecken*, d.h. die rechts davon stehen aber nicht selber Pivot-Spalten sind, eben Linear-Kombinationen der Pivot-Spalten davor!! sind. Das ist aber unmittelbar einsichtig. □

[64] Natürlich ist umgekehrt im Falle $m = n$ jede invertierbar Matrix ein Isomorphismus, genaugenommen sind das dann auch alle derartigen Isomorphismen.

In Worten: Eine Nicht-Pivot-Spalte "versteckt" sich hinter der davorliegenden Pivotspalte, und in jeder Zeile (bis zum Zeilen-Index der davorliegenden Stufe = Pivotzeile) steht ein Einser. Somit ist klar (man rechne ein einfaches Beispiel bzw. male sich die Situation in schematischer Form, das heißt mit $*-$Symbolen) dass diese Nicht-Pivotzeile sich als Linearkombination der davorliegenden Spalten erweist.

BEMERKUNG: Die soeben durchgeführte Diskussion zeigt, dass die im Prinzip plausible, aber aufwändige Diskussion der Erstellung einer Basis des Spaltenraumes durch fortlaufendes Hinzufügen von neuen Spaltenvektoren, die zu den bisher ausgewählten linear unabh. sind (somit neue "Dimensionen" eröffnen), durch die Neu-Interpretation des Gauss-Algorithmus (und hier ist es z.B. wichtig, dass man sich auf Matrix-Multiplikation von links, mit invertierbaren Elementar-Matrizen beschränkt!) auf eine algorithmische Basis gestellt wird. Dass diese konkrete Auswahl möglicherweise numerisch nicht besonders stabil ist (schleifende Schnitte!) ist auch klar, wenn man sich die Lage geometrisch vorstellt. Es ist keineswegs gesagt, dass man auf diese Weise eine "gute Basis" bekommt. Letztendlich ist ja auch die Reihenfolge der Vektoren festgelegt, sobald man diese als Spalten in eine Matrix $\boldsymbol{A}$ einpflanzt.

Typischer Fall: Ergibt die Gauss-Elimination, dass die Spalten mit den Nummern $1, 2, 4, 5, 7$ einer 8×8-Matrix Pivot-Elemente enthalten, kann man sofort sagen, daß die Spalte mit Nr.3 sich als Linearkombination der 1. und 2.-ten Spalte darstellen läßt, ebenso die Spalte mit Nr. 6 als Lin. Komb. von Nr. 1,2,4,5, etc. Die Endform der Gauss-Elimination liefert dann 5 (weil es 5 Pivot-Elemente gibt) Zeilenvektoren mit führenden Einsen, und diese bilden *eine Basis* für den Zeilenraum von $\boldsymbol{A}$.

Alternative Problemstellung: Gegeben eine Kollektion von Vektoren im $\mathbb{R}^m$, d.h. eine (endliche) Menge $M \subset \mathbb{R}^m$. Welche von diesen Vektoren kann man entfernen, um eine Basis für ihr *lineares Erzeugnis*, d.h. den Raum $\boldsymbol{V}_M$ zu bekommen. Antwort: Man schreibe diese Vektoren als Spalten in eine $m \times n$-Matrix $\boldsymbol{A}$ und wende dann das soeben beschriebene Verfahren an.

Später werden wir noch sehen, wie man für einen (endlich erzeugten) Teilraum des $\mathbb{R}^n$ sogar eine Orthonormalbasis finden kann. Diese ist besonders schön, weil sie erlaubt, die orthogonale Projektion auf den Teilraum direkt (numerisch) zu realisieren!

Zum Abschluss noch eine andere Standardaufgabe: Man hat einen n-dimensionalen Vektorraum $\boldsymbol{V}$ (der Begriff der Dimension muss noch genauer nach den Weihnachtsferien diskutiert werden: die Minimalzahl von Elementen eins Erzeugersystems) und darin ein System von n Vektoren. Man stelle fest, ob diese Vektoren eine Basis für $\boldsymbol{V}$ bilden: Antwort: Man muss nur feststellen, ob sie linear unabhängig sind! (und das kann oft durch Studium der Matrix geschehen, die man benutzt, um das System aus einer (*Standard*) beliebigen bekannten Basis zu erhalten. Wenn diese Matrix invertierbar ist (und nur dann), hat man eine (neue) Basis (aus einer alten gemacht).
Gehört wohl anderswohin, ist Wdhlg. von Lemma 22. isodim1

RnRmlemma **Lemma 25.** *Angenommen es gibt einen linearen Isomorphismus zwischen $\mathbb{R}^n$ und $\mathbb{R}^m$, dann gilt $n = m$.*

Proof. Bezeichnen wir den angenommenen Isomorphismus mit T, und seinen inversen Isomorphismus von $\mathbb{R}^m$ nach $\mathbb{R}^m$ mit S. Wir führen einen indirekten Beweis. Angenommen $m \neq n$. Wir diskutieren den Fall $n > m$ in Detail unter der Nutzung von T

(andernfalls ist $m > n$ und wir könnten den Widerspruch mittels S erzielen!). Nach Theorem 12 gibt es für T eine Matrix-Darstellung, die also mehr Spalten als Zeilen hat. Daraus folgt aber unter Anwendung der Gauss-Elimination, dass es mindestens eine freie Variable gibt, oder anders ausgesprochen, das homogene Gleichungssystem $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$ hat mehr als eine (sogar unendliche viele) Lösungen, was aber im Widerspruch zur Injektivität der Abbildung T steht. Dieser Widerspruch impliziert (zweimal angewendet), dass $m = n$ gelten muss. □

Der Beweis enthält auch schon in Argument für folgende Aussage:

indepvects

Lemma 26. *Im $\mathbb{R}^n$ sind mehr als n Vektoren stets linear abhängig.*

Genauso gilt: weniger als n Vektoren können nie den ganzen $\mathbb{R}^n$ aufspannen, die Begründung folgt später.[65]

Hinweis: Wir unterscheiden in der Folge nicht zwischen den Symbolen $\mathbf{c}, \vec{c}$ oder $\vec{\mathbf{c}}$.

basisomorph

Theorem 13. *Jede Basis $\mathscr{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ in einem (endlich erzeugten) Vektorraum $\boldsymbol{V}$ erzeugt einen Isomorphismus[66] $\Phi_{\mathscr{B}}$ von $\boldsymbol{V}$ nach $\mathbb{R}^n$, indem jedem Element $\mathbf{v} \in \boldsymbol{V}$ mit der eindeutig bestimmten Darstellung $\mathbf{v} = \sum_{k=1}^n c_k \mathbf{b}_k$ die Koeffizientenfolge $\vec{\mathbf{c}} := (c_k)_{k=1}^n$ zugeordnet wird. Die inverse Abbildung ist [natürlich] die Syntheseabbildung $S_{\mathscr{B}} : \vec{\mathbf{c}} \mapsto \sum_{k=1}^n c_k \mathbf{b}_k$. Man nennt diesen eindeutig bestimmten Vektor auch den Koordinatenvektor von $\mathbf{v}$ relativ zur (oder bezüglich der) Basis $\mathscr{B}$ und verwendet dafür das Symbol $\Psi_{\mathscr{B}}$ (wird später noch genauer ausgeführt)*

$$\Psi_{\mathscr{B}} : \mathbf{v} \mapsto [\mathbf{v}]_{\mathscr{B}} := \vec{\mathbf{c}} \quad \textit{wenn gilt:} \quad \mathbf{v} = \sum_{k=1}^{n} c_k \mathbf{b}_k. \tag{84}$$

cordvec1

oder äquivalent beschrieben:

$$\Phi_{\mathscr{B}}(\mathbf{v}) = [\mathbf{v}]_{\mathscr{B}}, \quad \Phi^{-1}(\vec{\mathbf{c}}) = \sum_{k=1}^{n} c_k \mathbf{b}_k. \tag{85}$$

cordmapdef

Man nennt daher die Abbildung $\Psi_{\mathscr{B}} = \Phi_{\mathscr{B}}^{-1}$ auch Syntheseabbildung. Man kombiniert die Information von $\mathbf{c}$ (Koeffizientenvektor) und $\mathscr{B}$ (Basis) einfach indem man sie zur Bildung einer Linearkombination heranzieht.

Es gilt auch die Umkehrung von dieser Aussage!

Es ist vielleicht von Vorteil, den konkreten Fall, den wir am Anfang der Vorlesung betrachtet haben, nochmals unter dieser allgemeinen Formulierung zu rekapitulieren.

Betrachten wir den Fall $\boldsymbol{V} = \mathbb{K}^n$ (und denken wir an $\mathbb{R}^n$ oder $\mathbb{C}^n$). Dann besteht eine Basis $\mathscr{B}$ aus einer Folge $(\vec{\mathbf{b}}_j)_{j=1}^n$ von Spaltenvektoren, die wir wie üblich zu einer Matrix

[65] Kurz gesagt: Wenn das möglich wäre könnte man - notfalls durch weiteres Weglassen - eine Basis mit weniger als n Elementen herstellen, aber dann müssten die n Einheitsvektoren linear abhängig sein, was sie aber offensichtlich nicht sind!

[66] Es sei mit $\Phi_{\mathscr{B}}$ die von $\mathscr{B}$ induzierte *Koeffizienten-Abbildung* bezeichnet, in der früheren Version des Skriptums wurde das Symbol $K_{\boldsymbol{V}}$ verwendet.

$\boldsymbol{B}$ zusammenfassen können. Die Basis Eigenschaft von $\mathscr{B}$ ist dann äquivalent zur Invertierbarkeit der Matrix $\boldsymbol{B}$. Die Synthese-Abbildung bzw. die Koeffizienten-Abbildung sind dann wie folgt gegeben:[67]

$$\Phi_{\mathscr{B}}^{-1}(\vec{\mathbf{c}}) = \boldsymbol{B} * \vec{\mathbf{c}}, \quad \text{bzw.} \quad \Phi_{\mathscr{B}}(\mathbf{v}) = \boldsymbol{B} * \mathbf{v}. \tag{86}$$

matrsyncof1

Proof. Eine **Basis** ist per definitionem ein linear unabhängiges Erzeugendensystem. In Worten: Die Eindeutigkeit der Darstellung von beliebigen Vektoren ist gewährleistet. □

Wenn man zwei verschiedene "abstrakte" Basen in einem allgemeinen Vektorraum $\boldsymbol{V}$ hat, sagen wir $\mathscr{C}$ und $\mathscr{B}$, mit entsprechenden Isomorphismen $\Phi_{\mathscr{B}}$ und $\Phi_{\mathscr{C}}$, mit $\Phi_b asB(\mathbf{v}) = [\mathbf{v}]_{\mathscr{B}}$ bzw. $\Phi_b asC(\mathbf{v}) = [\mathbf{v}]_{\mathscr{C}}$ (nach Definition dieser Abbildung), dann ist klar, dass der Übergang von $[\mathbf{v}]_{\mathscr{B}}$ nach $[\mathbf{v}]_{\mathscr{C}}$ durch den (!) Isomorphismus $\Phi_{\mathscr{C}} \circ \Phi_{\mathscr{B}}^{-1}$ gegeben ist. D.h. der Basis-Wechsel-Matrix ist ebenfalls eine $n \times n$-Matrix, die sich als Produkt von zwei Isomorphismen schreiben lässt.

13.3 Kompositions-Regel und Matrizen

VERBALE BESCHREIBUNG DER SITUATION Jeder linearen Abbildung T zwischen zwei Vektorräumen über demselben Körper kann bei jeweils fest gewählter Basis eine Matrix zugeordnet werden. Je nach Dimension der Räume haben diese Matrizen entsprechendes Format. Die Spalten dieser Matrix enthalten die Koordinaten der Bilder der Basis-Vektoren des ersten Basis, beschrieben in den Koordinaten der Basis die im Zielraum verwendet wird.

matrcomp123

Theorem 14. *Schaltet man zwei lineare Abbildungen hintereinander und beschreibt bei jeweils in festen Basen der zugehörigen Räume, dann ist die Matrix, die der zusammengesetzten Abbildung zugeordnet wird, genau die Produktmatrix der zugehörigen Matrizen. In Symbolen:*

$$[T_2 \circ T_1]_{\mathscr{B}_3 \leftarrow \mathscr{B}_1} = [T_2]_{\mathscr{B}_3 \leftarrow \mathscr{B}_2} * [T_1]_{\mathscr{B}_2 \leftarrow \mathscr{B}_1} \tag{87}$$

matrkomp

Wie man sich leicht überzeugen kann, passen die zugehörigen Matrizen jedenfalls vom Format zusammen, d.h. es gibt kein grundsätzliches Problem mit der Matrix-Multiplikation (Achtung auf die Reihenfolge).

Proof. Beginnen wir damit, entsprechende Bezeichnungen zu wählen. Es ist naheliegend zu sagen, es gibt einen ersten, zweiten und dritten Vektorraum, also $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3$ und lineare Abbildungen $T_1 : \boldsymbol{V}_1 \to \boldsymbol{V}_2$ bzw. $T_2 : \boldsymbol{V}_2 \to \boldsymbol{V}_3$. Jeder der Räume sei mit einer (geordneten) Basis $\mathscr{B}^1, \mathscr{B}^2, \mathscr{B}^3$ ausgestattet, sodass wir eine Matrix $\boldsymbol{A}_1$ zur Beschreibung von T_1 und eine Matrix $\boldsymbol{A}_2$ zur Beschreibung von T_2 haben. Nimmt man an, dass die Dimension des Raumes $\boldsymbol{V}_k$ gerade n_k ist, mit $k = 1, 2, 3$, dann ist das Format von $\boldsymbol{A}_1$ gerade $n_2 \times n_1$ und das Format von $\boldsymbol{A}_2$ ist $n_3 \times n_2$, also ist das Matrix Produkt $\boldsymbol{A} := \boldsymbol{A}_2 * \boldsymbol{A}_1$ wohldefiniert und vom Format $n_3 \times n_1$, also ist es sinnvoll, diese Matrix mit der Matrix zu vergleichen, welche T (mit den Basen $\mathscr{B}^1$ bzw. $\mathscr{B}^3$) beschreibt.

[67] Diese wichtigen Relationen sollten im Detail verstanden werden, was jedenfalls erfordert, dass sich die LeserInnen mehrfach mit dem Thema auseinandersetzen!

Zur allgemeinen Idee: Die Matrix $\boldsymbol{A}$ beschreibt den Operator, indem sie die Bilder aller Basisvektoren einer Basis $\mathscr{B}$ $T(\mathbf{b}_j)$ durch seine Koordinaten bzgl. $\mathscr{B}'$ im Zielbereich durch einen Vektor (Spalte Nr. j von $\boldsymbol{A}$) $\vec{\mathbf{a}}_j$ speichern, oder als Formel:

$$T(\mathbf{b}_j) = \sum_{r=1}^{s} a_{r,j}\mathbf{b}'_r.$$

Insbesondere haben wir für T_1, T_2 mit entsprechenden Matrizen $\boldsymbol{A}^1$ bzw. $\boldsymbol{A}^2$.

$$T_1(\mathbf{b}_j^1) = \sum_{r=1}^{n_2} a_{r,j}^1 \mathbf{b}_r^2.$$

$$T_2(\mathbf{b}_r^2) = \sum_{l=1}^{n_3} a_{l,r}^2 \mathbf{b}_l^3.$$

Betrachten wir nun die Wirkung von $T = T_2 \circ T_1$ auf die Basis $\mathscr{B}^1$ bestehend aus den Vektoren $\mathbf{b}_1^1, \cdot, \mathbf{b}_{n_1}^1$, so bekommen wir unter Verwendung der Linearität von T_2:

$$T(\mathbf{b}_j^1) = T_2(T_1(\mathbf{b}_j^1)) = T_2\left(\sum_{r=1}^{n_2} a_{r,j}^1 \mathbf{b}_r^2\right) = \sum_{r=1}^{n_2} a_{r,j}^1 T_2(\mathbf{b}_r^2) =$$

$$= \sum_{r=1}^{n_2} a_{r,j}^1 \sum_{l=1}^{n_3} a_{l,r}^2 \mathbf{b}_l^3 = \sum_{l=1}^{n_3} \left(\sum_{r=1}^{n_2} a_{l,r}^2 a_{r,j}^1\right) \mathbf{b}_l^3$$

Charakteristische Eigenschaft von $\boldsymbol{A}$ (wird deshalb auch als $[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ bezeichnet):

$$\boldsymbol{A} * [\mathbf{v}]_{\mathscr{B}} = [T(\mathbf{v})]_{\mathscr{C}} \quad \Leftrightarrow \quad [T(\mathbf{v})]_{\mathscr{C}} = [T]_{\mathscr{C}\leftarrow\mathscr{B}} * [\mathbf{v}]_{\mathscr{B}}.$$

Das gilt nach Konstruktion! (Definition) für die Basisvektoren, denn wir haben

$$[\mathbf{b}_j]_{\mathscr{B}} = \vec{\mathbf{e}}_j, \quad \text{also} \quad \boldsymbol{A} * \vec{\mathbf{e}}_j = \mathbf{a}_j = [T(\mathbf{v})]_{\mathscr{C}}.$$

Die Abbildung, die jedem Koeffizienten-Vektor $[\mathbf{v}]_{\mathscr{B}} \in \mathbb{K}^n$ den entsprechenden Vektor $[T(\mathbf{v})]_{\mathscr{C}} \in \mathbb{K}^m$ zuordnet (als zusammengesetzte lineare Abbildung der Form $\Phi_{\mathscr{C}} \circ T \Phi_{\mathscr{B}}^{-1}$) und die durch die Matrix-Multiplikation definierte lineare Abbildung stimmen also auf den Einheitsvektoren überein, *SIND ALSO GLEICH!* (für alle $\mathbf{v} \in \boldsymbol{V}$). □

$$[T_2 \circ T_1]_{\mathscr{B}_3\leftarrow\mathscr{B}_1} = [T_2]_{\mathscr{B}_3\leftarrow\mathscr{B}_2} * [T_1]_{\mathscr{B}_2\leftarrow\mathscr{B}_1} \tag{88}$$ matrkomp123

Diese Formel hat viele wichtige Konsequenzen. Trivialerweise gilt

$$[\mathbf{Id}_{\boldsymbol{V}}]^{-1}_{\mathscr{C}\leftarrow\mathscr{B}} == \left([\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{C}\leftarrow\mathscr{B}}\right)^{-1} = [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B}\leftarrow\mathscr{C}},$$

oder in der Cap'schen Notation:

$$\left([\mathbf{Id}_{\boldsymbol{V}}]^{\mathscr{B}}_{\mathscr{C}}\right)^{-1} = [\mathbf{Id}_{\boldsymbol{V}}]^{\mathscr{C}}_{\mathscr{B}}$$

weil offensichtlich wegen $\mathbf{Id}_{\boldsymbol{V}} \circ \mathbf{Id}_{\boldsymbol{V}} = \mathbf{Id}_{\boldsymbol{V}} = \mathbf{Id}_{\boldsymbol{V}} \in$ gilt:

$$\mathbf{Id}_n = [\mathbf{Id}_{\boldsymbol{V}}]^{\mathscr{B}}_{\mathscr{B}} = [\mathbf{Id}_{\boldsymbol{V}}]^{\mathscr{B}}_{\mathscr{C}} * [\mathbf{Id}_{\boldsymbol{V}}]^{\mathscr{C}}_{\mathscr{B}}.$$

13.4 Matrix-Darstellungen von Operatoren: Vergleich

Verschiedene Notationen, die in diversen Büchern verwendet werden, im Vergleich.

13.4.1 Cap Skriptum

`http://www.mat.univie.ac.at/~cap/files/Linalg.pdf`

$$[T]_{\mathscr{C}}^{\mathscr{B}}, \quad [T]_{\mathscr{C}}^{\mathscr{B}} * [\mathbf{v}]_{\mathscr{B}} = [T(\mathbf{v})]_{\mathscr{C}}, \quad [T_2 \circ T_1]_{\mathscr{B}_3}^{\mathscr{B}_1} = [T_2]_{\mathscr{B}_3}^{\mathscr{B}_2} * [T_1]_{\mathscr{B}_2}^{\mathscr{B}_1}.$$

Kommt dort in der Form (Prop.,4.15):

$$[f(v)]_{\mathscr{B}'} = [f(v)]_{\mathscr{B}'}^{\mathscr{B}} [v]_{\mathscr{B}}.$$

Satz 4.15 (Cap, p.62): Für feste Basen $\mathscr{B}$ und $\mathscr{B}'$ in $\boldsymbol{V}$ bzw. $\boldsymbol{W}$ etabliert die Abbildung $f \mapsto [f]_{\mathscr{B}'}^{\mathscr{B}}$ einen linearen Isomorphismus zwischen $\mathcal{L}(\mathbf{V},\mathbf{W})$ und $M_{m,n}(\mathbb{K})$, dem Vektorraum der $m \times n$-Matrizen über $\mathbb{K}$.

A. Cap verwendet für (lineare) Abbildungen den typischen Abbildungterm und nennt sie einfach f. Das erscheint für einige *meiner Beispiele* wenig vorteilhaft, wenn man als Vektorräume Funktionen (wie $f_1, f_2, ...$) hat (vgl. unser Beispiel mit der Eulerschen Formel).

Die Kompositionsregel ist dann in Satz 4.16 (p.57) ausformuliert. Ebenso die Äquivalenz der Invertierbarkeit einer linearen Abbildung mit der Invertierbarkeit der entsprechenden Matrix (das gilt für jede!! Wahl von Basen, und manchmal macht es die Wahl der Basen leichter oder schwerer, die Inverierbarkeit zu erkennen).

13.4.2 Haller: Skriptum

`http://www.mat.univie.ac.at/~stefan/files/LA/LA+LA1+LA2.Skriptum.pdf`

Haller (typisches Beispiel/Symbol). In seinem Skriptum werden die linearen Abbildungen mit φ bezeichnet (weil das das typische Symbol für strukturerhaltende Abbildungen ist), und die Basis-Wechsel-Matrizen werden als "Transformationen von $\mathbb{K}^n$" gesehen, und bekommen daher den Buchstaben T als Symbol zugeordnet.

$$[\varphi]_{\tilde{C}\tilde{B}} = T_{\tilde{C}C}\, [\varphi]_{CB}\, T_{B\tilde{B}}, \quad \text{oder} \quad [v]_{\tilde{B}} = T_{\tilde{B}B} [v]_B$$

OK/Vorteil im Vergleich zu Cap ($[T]_{\mathscr{C}}^{\mathscr{B}}$), dass alles auf einer Ebene ist, und die Reihenfolge so gewählt wurde, dass die Kompositionsregel (auch das Domino-Prinzip, wenn man die Anzahl der jeweiligen Basis-Elemente bedenkt) respektiert, aber man muss sich erst "daran gewöhnen". Der Pfeil (hgfei) im Symbol $[T]_{\mathscr{B}\hookleftarrow\mathscr{B}}$ ist in dem Sinn dafür unterstützend.

Nachteil bei Haller: Er schreibt/denkt wohl an *Morphismen* zwischen Vektorräumen, und verwendet daher den Buchstaben φ für lineare Abbildungen, und andererseits T für Koordinaten-Transformationen.

13.4.3 Howard Anton

Das Buch von Howard Anton ([1]) ist recht "angewandt", behandelt aber den Basiswechsel in der richtigen Allgemeinheit (Kap.8.4). Er verwendet für die Abbildung T (gleiche Konvention wie in diesem Skriptum) das Symbol $[T]_{\mathscr{B},\mathscr{B}'}$ wenn $\mathscr{B}'$ die Basis im Definitionsbereich $\boldsymbol{V}$ ist und $\mathscr{B}$ diejenige im Zielbereich. Das hat die größte Ähnlichkeit mit der von uns gewählten (nächster Abschnitt) $[T]_{\mathscr{B}\hookleftarrow\mathscr{B}'}$. (d.h. es fehlen nur der "Enterhaken"). Kleiner Nachteil seiner Notation ist die Tatsache, dass er die Basen (e.g. Satz 8.4.2.) $\mathscr{B}, \mathscr{B}', \mathscr{B}''$ nennt, was einen potentiellen Konflikt mit der Funktion des MATLAB/OCTAVE Befehls $\boldsymbol{B} \to \boldsymbol{B}'$ ergibt.

13.4.4 Paul A. Fuhrmann: A polynomial approach to Lin.Alg.

In seinem Buch [3] verwendet P.A. Fuhrmann die folgende Konvention, um z.B. den Basiswechsel zu beschreiben (p.47):

$$[v]^{\mathscr{B}_1} = [Id]_{\mathscr{B}}^{\mathscr{B}_1}[v]^{\mathscr{B}}, \quad \mathbf{v} \in \boldsymbol{V}. \tag{89}$$

baschangfu12

Auch hier (wie bei Cap etc.) schlucken Basen die *unten stehen* die Basen die oben stehen, und die neue Basis (hier $\mathscr{B}_1$) bleibt übrig. Es gibt den Vorteil, dass man im Prinzip einzelne Koordinaten gut ansprechen kann, d.h. es ist keine Problem, das Symbol $[v]_3^{\mathscr{B}_1}$ als dritte Koordinate des Vektors $[v]^{\mathscr{B}_1}$ zu interpretieren, was bei dem Symbol $[x]_{\mathscr{B}_1}$ schwieriger auszudrücken ist[68].

Trotzdem ist es ein nicht sehr "praktische" Methode (d.h. Cap's Schreibweise ist da zu bevorzugen). Im Buch [3] wird übrigens für die Standard-Basis der Polynomfunktionen, d.h. z.B. für $\mathscr{P}_n(\mathbb{C})$, d.h. für $\{1, z, z^2, \cdot, z^n\}$ das Symbol $\mathscr{B}_{st}$ verwendet.

13.4.5 hgfei Matrix-Notation

$$[T]_{\mathscr{C}\hookleftarrow\mathscr{B}}, [T]_{\mathscr{B}\hookleftarrow\mathscr{B}}, \text{ oder nur } [T]_{\mathscr{B}}.$$

Allgemein gilt: Der Inversion der linearen Abbildung (klarerweise bildet T^{-1} von $\boldsymbol{W}$ nach $\boldsymbol{V}$ ab) entspricht der Inversion der zugehörigen Matrizen, unter Beibehaltung der Basen in den entsprechenden Räumen:

$$[T^{-1}]_{\mathscr{B}\hookleftarrow\mathscr{C}} = [T]^{-1}_{\mathscr{C}\hookleftarrow\mathscr{B}} = \quad \text{or} \quad ([T]_{\mathscr{C}\hookleftarrow\mathscr{B}})^{-1} \quad \text{und} \quad [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B}\hookleftarrow\mathscr{B}} = \mathbf{Id}_n = [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{C}}.$$

Weitere Kommentare sind noch vorgesehen.

[68] Man könnte nur schreiben α_3, mit $\alpha = [x]_{\mathscr{B}_1}$, oder $([x]_{\mathscr{B}_1})_3$, was natürlich unschön aussieht und eher zu vermeiden ist!

13.5 Die Fritz Mayer Story

Ein Schüler der 3B hatte eine sehr schlechte Mathematik-Schularbeit geschrieben, und deshalb lud *die Lehrerin, welche diese Schularbeit korrigiert hatte*, den *Vater des Schülers, der diese schlechte Schularbeit geschreiben hatte*, zur Vorsprache in die Schule.

Auf dem Weg zur Schule beschloß *der Vater des Schülers*, sich noch kurz im Kaffeehaus zu stärken. Er setzte sich an den ersten Tisch rechts vom Eingang und bestellte einen Kaffee. Kurz darauf brachte *der Mann, bei dem er den Kaffee bestellt hatte*, genau den Kaffee, den er bei ihm bestellt hatte. Nach Bezahlung eben dieses Kaffees *bei dem Mann, bei dem er den Kaffee bestellt hatte*, fuhr der Vater in die Schule.

KURZVERSION

Bevor Fritz Mayer wegen der schlechten Schularbeit seines Sohnes bei der Lehrerin vorsprach, ließ er sich vom Kellner im Kaffeehaus eine Großen Braunen servieren.

- Was sind die sprachlichen Vorteile der zweiten Fassung;
- u.a. wurde *Unwesentliches weggelassen*;
- der Name Fritz Mayer ist nicht "informativ" sondern (was die Geschichte betrifft) willkürlich gewählt, aber verkürzt die Darstellung
- es geht um drei Personen: Lehrerin, Vater, Kellner, die in einer Aktionskette miteinander etwas zu tun haben; der Zusammenhang muss geklärt sein (warum fährt der Vater in die Schule, was hat der Kellner damit zu tun, etc.)
- die zentralen Themen in der Geschichte sind (für uns) nicht der Schüler (könnte Michi Mayer sein),
- Die Terminologie "Lehrerin" bzw. Kellner (auch Vater) ist wohl etabliert und allgemein verständlich. Hingegen: beim Namen "Fritz Mayer" denkt man nicht sofort an den Vater von beispielsweise Michi Mayer.
- Die Kurzversion ist in einem Punkt sogar genauer: der Vater bestellte einen "Grossen Braunen" (nicht irgendeinen Kaffee), auch wenn das inhaltlich für die Story nicht viel bedeutet!
- unnötige Information kann mit Vorteil in der Beschreibung weggelassen werden, etwa wo der Vater gesessen ist und dass er wohl auch den Kaffee bezahlt hat.
- Man könnte in der Kurzversion noch anmerken, dass der Vater das Kaffeehaus aufsucht, bzw. den Kaffee bestellte, um sich vor dem Gespräch zu stärken.
- Wir reden von einer Matrix $\boldsymbol{A}$, die eine lineare Abbildung T von $\boldsymbol{V}$ (mit Basis $\mathscr{B}$) nach $\boldsymbol{W}$ (mit Basis $\mathscr{C}$) darstellt. $\boldsymbol{A}$ ist Fritz Mayer. $[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ entspricht dem *Vater von dem Schüler der die schlechte Schularbeit geschrieben hat.* Der *Kellner* (eine andere Matrix $\boldsymbol{B}$) entspricht *dem Mann, bei dem der Vater .. den Kaffee bestellt hatte...*, etc. (Raum für Ausschmückungen). z.B. das Auto von Fritz Mayer (wohl nicht das Auto von dem Mann dessen Sohn...) entspricht der *dritten Spalte von* $\boldsymbol{A}$, was natürlich die dritte Spalte von $[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ ist, also $[T(\mathbf{b}_3)]_{\mathscr{C}}$.

13.5.1 Noch ein anderer sprachlicher Vergleich

Es geht um die Wahl von Bezeichnungen. Manchmal wird eine Matrix durch ihre Funktionalitaet beschreiben , manchmal durch Grossbuchstaben wie A,B,C.. wobei nicht wichtig ist, welcher Buchstabe verwendet wird....

Auch (die Namen/Buchstaben für) Indices sind nicht wichtig (ich meine j, k, l, etc.) aber sie sollten konsistent verwendet werden!

Ich könnte auch Gesetzestext hernehmen. Dort steht nicht, dass Fritz Mayer für's Schnellfahren Strafe zahlen muss, sondern "der Fahrer eines Autos, welches im Stadtgebiet die dort erlaubte Höchstgeschwindigkeit um mehr als 10% ueberschreitet hat das Problem, dass er Zahlen muss"!

Ein analoger Fall aus der linearen Algebra:

- (die verbale Formulierung) ein Matrix, die mehr Spalten als Zeilen hat, hat das Problem, dass ihre Spalten linear abhaengig sind.

  Finde ich besser als:

- (*) Wenn $\boldsymbol{A}$ eine Matrix vom Format $m \times n$ ist und wenn $n > m$ ist, dann sind die Spalten von $\boldsymbol{A}$ linear abhängig.

Ist natürlich dasselbe. Aber was ist, wenn ich eine Matrix $\boldsymbol{A}$ vom Format 5×3 bekomme, und frage, ob die Zeilen von $\boldsymbol{A}$ linear abhängig sind? Dann muss ich $\boldsymbol{A}^t$ bilden, und das m in $(*)$ ist das n von $\boldsymbol{A}$, also 3 und ... (>> Knopf im Hirn!?)... odr??

13.6 Matrix-Darstellung: verschiedene Perspektiven

Eine mögliche Beschreibung bzw. Begründung für die spezifische Wahl der Matrix-Darstellung bzgl. zweier allgemeiner Basen ergibt sich (begrifflich) einerseits aus der Tatsache, dass in allgemeinen Vektorräumen einfach keine "Einheitsvektoren" vorhanden sind, andererseits das selbst im $\mathbb{R}^k$ nicht klar ist, warum die sozusagen *willkürlich* eingeführten Euklidischen Koordinaten die einzigen sein sollten, die zur Beschreibung von Vektoren im $\mathbb{R}^n$ dienen sollten. Selbst wenn man sich auf Orthonormal-Basen einschränken will, gibt es eine *Vielzahl* von Möglichkeiten, die nun in der richtigen Allgemeinheit diskutiert werden können.

Hier kommt später noch ein Diagramm!, Mai 2014: hgfei

Hat man also eine lineare Abbildung $T \in \mathcal{L}(\mathbf{V},\mathbf{W})$ und die übliche Ausstattung der Räume mit Basen $\mathscr{B}$ bzw. $\mathscr{C}$, mit n bzw. n Elementen. Dann will man die Wirkung von T *durch die jeweiligen Koordinaten* beschreiben, d.h. man will wissen, wie man von den $\mathscr{B}$-Koordinaten von $\mathbf{v} \in \boldsymbol{V}$ zu den $\mathscr{C}$-Koordinaten von $T(\mathbf{v}) \in \boldsymbol{W}$ kommt. Anders ausgedrückt, man sucht eine Abbildung zu finden, welche $[\mathbf{v}]_{\mathscr{B}}$ in $[T(\mathbf{v})]_{\mathscr{C}}$ überführt.

Das Prinzip ist einfach (hier einmal zur Abwechslung verbal beschrieben, aber sicher besser durch ein Diagramm repräsentiert, > 6./7. Mai?): Hat man die $\mathscr{B}$-Koordinaten von $\mathbf{v}$, so muss man die Synthese-Abbildung $\Psi_{\mathscr{B}} = \Phi_{\mathscr{B}}^{-1}$ anwenden, um $\mathbf{v}$ zu bekommen. Dann kann man T anwenden, um daraus $T(\mathbf{v})$ zu gewinnen, und schliesslich wendet man

die Koeffizienten-Abbildung $\Phi_{\mathscr{C}}$ an, um die $\mathscr{C}$-Koordinaten von $T(\mathbf{v})$ zu bekommen. Wir haben also eine *zusammengesetzte lineare (!) Abbildung* von der Form:

$$TBC := \Phi_{\mathscr{C}} \circ T \circ \Phi_{\mathscr{B}}^{-1}$$

welche offenbar von $\mathbb{R}^n$ nach $\mathbb{R}^m$ geht, und somit eine Matrix-Darstellung hat.

Wenn man es genauer betrachtet (und ein paar kleine Üeberlegungen anstellt), so ist die Matrix $[T]_{\mathscr{C}\leftarrow\mathscr{B}}$ genau die Matrix-Darstellung von TBC, wie wir sie schon lange kennen.

Wieder spezialisiert auf den Fall $\boldsymbol{V} = \mathbb{R}^n$ und $\boldsymbol{W} = \mathbb{R}^m$, allerdings nun mit allgemeinen Basen, dann ist $\mathscr{B}$ durch eine $n \times n$-Matrix $\boldsymbol{B}$ und im Zielraum $\mathscr{C}$ durch eine $m \times m$-Matrix $\mathbf{C}$ beschrieben, und wir erhalten durch Kombination der bisherigen Beobachtungen, dass die "darstellende Matrix" für die lineare Abbildung $T : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ geben ist durch:

$$[T]_{\mathscr{C}\leftarrow\mathscr{B}} = \boldsymbol{B}^{-1} * \boldsymbol{A} * \mathbf{C} \tag{90}$$ genmatrep3

Ist insbesondere $\boldsymbol{V} = \boldsymbol{W}$ und werden auf beiden Seiten die gleiche Basis verwendet (entweder $\mathscr{B}$ oder $\mathscr{C}$), dann gelten die folgenden Identitäten:

$$[T]_{\mathscr{B}} := [T]_{\mathscr{B}\leftarrow\mathscr{B}} = \boldsymbol{B}^{-1} * \boldsymbol{A} * \boldsymbol{B} \quad \text{and} \quad [T]_{\mathscr{C}} = \mathbf{C}^{-1} * \boldsymbol{A} * \mathbf{C} \tag{91}$$ genmatrep3

Daraus ist leicht der Zusammenhang zwischen diesen beiden Darstellungen berechenbar, indem man etwa die erste Gleichung zuerst von links mit $\boldsymbol{B}$ und dann noch mit $\mathbf{C}^{-1}$ multipliziert, also insgesamt mit $\mathbf{H} := \mathbf{C}^{-1} * \boldsymbol{B}$ multipliziert, und entsprechend auf der rechten Seite mit $\mathbf{H}^{-1}$. Wir haben also insgesamt die wichtige Formel:

$$[T]_{\mathscr{C}} = \mathbf{H} * [T]_{\mathscr{B}} * \mathbf{H}^{-1} \quad \text{mit} \quad \mathbf{H} := \mathbf{C}^{-1} * \boldsymbol{B}. \tag{92}$$ changmatrep1

Man spricht ganz allgemein beim Übergang von $\boldsymbol{A}$ zu einer Matrix der Form $\boldsymbol{B} * \boldsymbol{A} * \boldsymbol{B}^{-1}$ (oder gleichwertig: $\boldsymbol{A} \mapsto \mathbf{C}^{-1} * \boldsymbol{A} * \mathbf{C}$) von einer *Konjugation* der Matrix $\boldsymbol{A}$ mit der Matrix $\boldsymbol{B}$ bzw. $\mathbf{C}$. In unserem Fall findet die Konjugation mit

$$\mathbf{H} := \mathbf{C}^{-1} * \boldsymbol{B} = [\mathbf{Id}_{\mathbb{R}^n}]_{\mathbf{C}\leftarrow\boldsymbol{B}},$$

der sogenannten *Basis-Wechsel-Matrix* statt.

In unserer Schreibweise kann dies aber auch noch einfacher und allgemeiner ausgedrückt werden, wobei die Vielfalt der Begriffe etwas reduziert wird. Alle auftretenden Matrizen haben in dieser Betrachtungsweise ihre natürliche Rolle und es ist nicht so mysteriös, warum und in welcher Reihenfolge sie auftreten.

Offenbar gilt $T = \mathbf{Id}_{\boldsymbol{V}} \circ T \circ \mathbf{Id}_{\boldsymbol{V}}$, wobei $\mathbf{Id}_{\boldsymbol{V}}$ nur vorgeschaltet wird, um die Wechsel von einer Basis zu einer anderen zu beschreiben. Wir wählen nun für T zunächst (innne) eine Basis $\mathscr{B}$, wollen aber "aussen" mit $\mathscr{C}$-Koordinaten arbeiten. Dann ist nach den Kompositionsregeln klar, dass gilt (Formel für Basis-Wechsel):

$$[T]_{\mathscr{C}\leftarrow\mathscr{C}} = [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{C}\leftarrow\mathscr{B}} * [T]_{\mathscr{C}\leftarrow\mathscr{C}} * [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B}\leftarrow\mathscr{C}}.$$

13.6.1 Konkrete Beispiele von Basiswechsel

Einerseits wird in der Vorlesung das mit der Eulerschen Formel verbunden Beispiel genauer diskutiert werden. Andererseits einfache, konkrete Bespiele...

Wir betrachten den Teilraum von $\boldsymbol{C}_b(\mathbb{R})$, dem Raum aller stetigen (Continuous) und beschränkten (Bounded) komplexwertigen Funktionen auf $\mathbb{R}$. Dies ist ein Vektorraum über dem Körper $\mathbb{C}$ der komplexen Zahlen (d.h. $\mathbb{K} = \mathbb{C}$), welcher von den Exponential-Funktionen e^{ix}, e^{-ix} aufgespannt wird.

Welche Dimension hat dieser Raum. Klar ist, dass er höchstens zwei-dimensional ist (nur zwei Erzeuger), aber auch dass die beiden nicht (auch nicht komplexe) Vielfache voneinander sind, also ist der Raum zwei-dimensional (über $\mathbb{C}$). Wir werden gleich noch sehen, dass man auch zwei Linear-Kombinationen bilden kann, die "offensichtlich" linear unabhängig sind, also muss auch deshalb der Raum zweidimensional sein. Wir haben also: $\boldsymbol{V} = \mathbf{LH}(M)$, mit $M = \{e^{ix}, e^{-ix}\}$.

Eine alternative Basis ist durch Linear-Kombination aus diesen Vektoren berechenbar, unter Verwendung der Eulerschen Formel (und deren Umkehrung). Die Winkelfunktionen $\cos(x)$ und $i\sin(x)$ (ist bequemer zum Rechnen, weil es erlaubt nur mit reellen Linear-Kombinationen zu arbeiten!) spannen diesen Raum ebenfalls auf.

Nun betrachten wir die Gruppe der Translationsoperators $T_z, z \in \mathbb{R}$ auf diesem Teilraum, und zwar einmal in der Exponential-Basis $\mathscr{B} := \text{Expi} := \{e^{ix}, e^{-ix}\}$ und ein andermal in der *CoSi*-basis $\mathscr{C} := \text{CoSi} := \{\cos(x), i\sin(x)\}$.

In der ersten Basis ist die Darstellung leicht, weil aufgrund des Exponentialgesetzes

$$e^{i(x-z)} = e^{-iz}e^{ix}, \quad \text{and} \quad e^{-i(x-z)} = e^{iz}e^{-ix}, \quad x, z \in \mathbb{R}$$

gilt wir also für die Matrix-Darstellung eine Diagonalmatrix bekommen

$$[T_z]_{\text{Expi}} = \begin{pmatrix} e^{-iz} & 0 \\ 0 & e^{iz} \end{pmatrix}. \tag{93}$$

Um nun daraus die Darstellung derselben Operatoren bzgl. der anderen $\mathscr{C} = \text{CoSi}$ zu bekommen, müssen wir zuerst die Basis-Wechsel-Matrix bestimmen!

Die Eulersche Formel für $e^{\pm ix} = \cos(x) \pm i\sin(x)$ liefert sofort, dass

$$[\mathbf{Id}_{\boldsymbol{V}}]_{\text{CoSi}\hookleftarrow\text{Expi}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{94}$$

Die inverse Matrix dazu ist die Matrix selber, bis auf den Faktor $1/2$ (weil $\boldsymbol{A}*\boldsymbol{A} = 2\mathbf{Id}_2$). Wir erhalten also unter Verwendung der allgemeinen Formel

$$[T_z]_{\text{CoSi}} = [\mathbf{Id}_{\boldsymbol{V}}]_{\text{CoSi}\hookleftarrow\text{Expi}} * [T_z]_{\text{Expi}\hookleftarrow\text{Expi}} * [\mathbf{Id}_{\boldsymbol{V}}]_{\text{Expi}\hookleftarrow\text{CoSi}}$$

durch Ausführen der Rechnung

$$\boldsymbol{A}_z := [T_z]_{\text{CoSi}} = \begin{pmatrix} \cos(z) & -i\sin(z) \\ -i\sin(z) & \cos(z) \end{pmatrix} \tag{95}$$ TzCosi1

Wenn also f von der Form $f(x) = c_1 \cos(x) + c_2 i \sin(x)$, wenn also $[f]_{\mathrm{CoSi}} = \vec{\mathbf{c}} \in \mathbb{C}^2$ gilt, so ist auch $T_z(f)$ von derselben Form, mit Koeffizienten der Form $\vec{\mathbf{d}} = \boldsymbol{A}_z * \vec{\mathbf{c}}$, bzw. $[T_z f](x) = d_1 \cos(x) + d_2 i \sin(x)$,
Speziell auf die beiden Einheitsvektoren angewandt, also $\vec{\mathbf{c}} = \vec{\mathbf{e}}_1$, als $f(x) = 1 \cos(x)$ ergibt sich eine (bis auf das Umstellen der Reihenfolge der Terme) die bekannt Formel:

$$[T_z \cos](x) = cos(x - z) = \cos(z)\cos(x) + \sin(z)\sin(x). \tag{96}$$ `cosshift1`

sowie für die (komplexe) Sinus-Funktion:

$$[T_z(i \sin)](x) = i \cdot sin(x - z) = i \cdot [\cos(z)\sin(x) - \sin(z)\cos(x)]. \tag{97}$$ `sinshift1`

oder umgeformt auf die übliche Schreibweise:

$$\sin(x \mp z) = \sin(x)\cos(z) \pm \cos(x)\sin(z). \tag{98}$$ `sinshiftr1`

sowie

$$\cos(x \pm z) = \cos(x)\cos(z) \mp \sin(x)\sin(z). \tag{99}$$ `cosshift2`

Da die Operatoren $(T_z)_{z\in\mathbb{R}}$ auf natürliche Weise bzgl. Komposition eine Gruppe bilden (die de facto isomorph zu $(\mathbb{R}, +)$ ist!) müssen auch die entsprechenden darstellenden Matrizen eine Gruppe bilden. Das ist recht klar bei den "exponentiellen Diagonalmatrizen" aber weniger evident bei den "komplexen Drehmatrizen" (welche übrigens $\det(\boldsymbol{A}_z) = 1$ erfüllen!)

Als Alternative kann man die Situation mit Hilfe der *reell-wertigen Basis* $\{\cos(x), \sin(x)\}$, allerdings nun mit *komplexen* Koeffizienten, d.h. mit Hilfe der komplexen Matrizen vom Kapitel 10.1 (Formel (2),(3)), section 10.1: Euler's Formel. In dieser Basis haben die Translationsoperatoren die folgende Form:

$$\boldsymbol{B}_z := [T_z]_{\mathrm{CoSi}_R} = \begin{pmatrix} \cos(z) & -\sin(z) \\ \sin(z) & \cos(z) \end{pmatrix}. \tag{100}$$ `TzCosi1R`

13.7 Pivotelemente, Biorthogonalität

Den Zusammenhang zwischen inversen Matrizen bzw. Biorthogonal-Systemen werden wir erst später behandeln, möglicherweise erst im Winter- Semester (Dr. Ehler).

13.8 Information in der reduzierten Zeilenstufenform

Bemerkung (28.3.14, vielleicht anderswohin?)

Ein lineare Abbildung $T : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ von $\mathbb{K}^n$ nach $\mathbb{K}^m$ ist injektiv genau dann, wenn in jeder Spalte der RZSF (reduzierten Zeilgenstufenform) ein Pivotelement zu finden ist (daher notwendigerweise $n \leq m$!). Andererseits ist T genau dann surjektiv wenn die RZSF mindestens (de facto genau) m Pivotelemente aufweist. Im Falle $m = n$ gilt also: T ist genau dann bijektiv wenn die Matrix $\boldsymbol{A}$ invertierbar ist, bzw. die Spalten von $\boldsymbol{A}$ sind eine Basis (wobei auch zu bemerken ist, dass diese der Fall ist genau dann wenn $\boldsymbol{A}^t$ bzw. genauso $\boldsymbol{A}'$ invertierbar sind). Wir werden später noch diskutieren, dass die Spalten von $\boldsymbol{B} := \boldsymbol{A}'^{-1} = \boldsymbol{A}^{-1'}$ (geometrisch interpretiert) ein sog. *Biorthogonalsystem* zur Familie der Spalten von $\boldsymbol{A}$ sind (und umgekehrt), weil gilt (Kronecker Delta-Symbol, definiert durch $\delta_{j,j} = 1$, und $\delta_{j,k} = 0 \,\text{für}\, j \neq k$):

$$\langle \vec{\mathbf{b}}_j, \vec{\mathbf{a}}_k \rangle = \delta_{j,k} \tag{101}$$ biorthdef1

Die Bedeutung dieser Interpretation der inversen Matrix (im Prinzip der Zeilen der inversen Matrix) wird durch folgende Identität klar:

$$\vec{\mathbf{x}} = (\boldsymbol{A} * \boldsymbol{A}^{-1}) * \vec{\mathbf{x}} = \boldsymbol{A} * (\boldsymbol{B}' * \vec{\mathbf{x}}) = \sum_{k=1}^{n} \langle \vec{\mathbf{x}}, \vec{\mathbf{b}}_k \rangle \vec{\mathbf{a}}_k. \quad \forall \vec{\mathbf{x}} \in \mathbb{K}^n. \tag{102}$$ biorthrep1

Es gilt auch die Umkehrung davon. Wenn die Gleichung (102) (für alle Vektoren $\vec{\mathbf{x}} \in \mathbb{R}^n$, oder auch nur für die Elemente eines Erzeugendensystems gilt) für eine $n \times n$-Matrix $\boldsymbol{B}$, dann ist $\boldsymbol{B}'$ (d.h. $conj(\boldsymbol{B}^t)$) die inverse Matrix zu $\boldsymbol{A}$. Die Begründung für die Identität $\boldsymbol{A} * \boldsymbol{B}$ (die auch nur im Falle $m = n$ möglich ist, wird noch folgen!). Wesentlicher Teil der Begründung: Wenn $\boldsymbol{B}' * \boldsymbol{A} = \mathbf{I}_n$ gilt, dann müssen die Spalten von $\boldsymbol{B}$ ein Erzeugendensystem sein, da aber der $\mathbb{R}^n$ eine Basis mit n Elementen hat (d.h. n-dimensional ist) sind diese Vektoren auch (automatisch!, wegen ihrer geringen Zahl) linear unabhängig sein, somit ist $\boldsymbol{B}'$ invertierbar, somit aber auch $\boldsymbol{B}$ und weiters $\boldsymbol{A}$ und die beiden Matrizen sind zueinander invers.

Achtung: es ist zu prüfen, ob diese Argumente an dieser Stelle schon verwendbar sind! (oder erst später!?)

dimchar1 **Lemma 27.** *Je zwei Basen haben gleich viele Elemente. Diese gemeinsame Eigenschaft aller möglichen Basis eines endlichdimensionalen Vektorraumes, nämlich die Anzahl der Basiselemente, ist daher eine unter Isomorphie (von Vektorräumen) invariante Größe.*

Proof. Hat man zwei Basen $\mathscr{B}_1$ bzw. $\mathscr{B}_2$ mit m bzw. n Elementen. Dann definiert $\Phi_{\mathscr{B}_1} \circ S_{\mathscr{B}_2}$ als Komposition von zwei Isomorphismen wieder einen Isomorphismus, der nun aber $\mathbb{R}^n$ isomorph auf $\mathbb{R}^m$ abbildet. Nach Lemma (25) gilt daher $m = n$, w.z.z.w.. □ RnRmlemma

dimVR00 **Definition 17.** Die (von der konkreten Basis unabhängige) Anzahl der Elemente einer (beliebigen) Basis $\mathscr{B}$ eines Vektorraumes $\boldsymbol{V}$ wird die **Dimension** von $\boldsymbol{V}$ genannt.

dimchar2 **Lemma 28.** *Die Dimension ist gleich der Minimalzahl von Erzeugenden, oder aber (ebenso) gleich der Maximalzahl von linear unabhängigen Vektoren in $\boldsymbol{V}$.*

Proof. Die Behauptung beruht auf dem Prinzip, dass eine Erzeugendensystem genau dann nicht minimal ist, wenn eines seiner Elemente entfernt werden kann, ohne die lineare Hülle zu verändern. Das ist genau dann der Fall, wenn die Menge lineare abhängig ist. Die nicht mehr weiter reduzierbaren Erzeugersysteme, d.h. die minimalen (die nach Entfernung eines beliebigen Elementes eine echt kleinere Hülle hätten) sind also genau die *linear unabhängigen* Erzeugendensystem, also genau die Basen.

Ebenso ist eine linear unabhängige Menge, welche noch nicht den ganzen Vektorraum erzeugt, welche also nicht maximal ist, immer auch durch Hinzufügen eines weiteren von der Menge linear unabhängigen Elementes vergrößerbar. Mit anderen Worten: Die maximalen, linear unabhängigen Familien in $\boldsymbol{V}$ sind genau diejenigen linear unabh. Mengen, die auch Erzeugendensysteme sind. □

April 2014:
Eine weitere Konsequenz der bisher Betrachtungen ist die folgende:

maplinindep1 **Lemma 29.** *Es $\boldsymbol{V}$ ein n-dimensional Vektorraum, und $\mathbf{v}_1, \cdots \mathbf{v}_s$ sei eine linear unabhängige Familie von Vektoren in $\boldsymbol{V}$. Weiters sei $\boldsymbol{W}$ ein andere Vektorraum, in dem eine gleiche Anzahl von Vektoren $\mathbf{w}_1, \cdots, \mathbf{w}_s$ vorgegeben sind. Dann gibt es eine lineare Abbildung $T : \boldsymbol{V} \to \boldsymbol{W}$ welche $T(\mathbf{v}_j) = \mathbf{w}_j, 1 \leq j \leq s$ erfüllt.*

Proof. Da man jede linear unabh. Menge zu einer Basis ergänzen kann, gibt es (da ja $n \geq s$ gelten muss!) auch eine Basis $(\mathbf{v}_j)_{j=1}^n$ welche die als Teilmenge die gegebenen Vektoren $\mathbf{v}_j, 1 \leq j \leq s$ enthält. Auch die Folge $(\mathbf{w}_j)$ kann (allenfalls durch Verwenden von $\mathbf{w}_j = \mathbf{0} \in \boldsymbol{W}$ für $j > s$!) entsprechend ergänzt werden.

Da jedes $\mathbf{v} \in \boldsymbol{V}$ eine eindeutige Darstellung $\mathbf{v} = \sum_j c_j \mathbf{v}_j$ hat, ist die Abbildung $\mathbf{v} \mapsto (c_j) \in \mathbb{K}^n$ linear, aber ebenso $(c_j) \mapsto \sum_{j=1}^n c_j \mathbf{w}_j$ linear ist, ist auch die (zusammengesetzte) Abbildung $\mathbf{v} \mapsto \sum_{j=1}^n c_j \mathbf{w}_j$ linear, und hat die gewünschten Eigenschaften. □

Und wie sieht es aus, wenn die Familie $(\mathbf{v}_j)_{j=1}^s$ *linear abhängig* ist?? Die Antwort ist relativ einfach, denn man hat:

mapfamdep1 **Theorem 15.** *Gegeben eine endliche Folge $\mathbf{v}_1, \cdots \mathbf{v}_s$ in einem Vektorraum $\boldsymbol{V}$ und eine gleich lange Folge $\mathbf{w}_1, \cdots, \mathbf{w}_s$ in einem anderen Vektorraum $\boldsymbol{W}$, dann gibt es eine lineare Abbildung $T : \boldsymbol{V} \to \boldsymbol{W}$ mit $T(\mathbf{v}_j) = \mathbf{w}_j, 1 \leq j \leq s$ genau dann wenn gilt:*[69]

$$\sum_{j=1}^{k} c_k \mathbf{v}_k = \mathbf{0} \in \boldsymbol{V} \quad \Rightarrow \quad \sum_{j=1}^{k} c_k \mathbf{w}_k = \mathbf{0} \in \boldsymbol{W}. \tag{103}$$ lindepVW1

[69] Wenn den vorhandenen lineare Abhängigkeiten der Folge in $\boldsymbol{V}$ auch entsprechende lineare Abhängigkeiten der vorgesehenen Bildvektoren in $\boldsymbol{W}$ entsprechen, wenn also gilt:

Proof. Der Beweis kann per Reduktion auf den vorigen Fall zurückgeführt werden. WENN die Folge $(\mathbf{v}_j)_{j=1}^k$ nicht linear unabhängig ist, dann kann man eine (maximale) linear unabhängige Teilmenge finden die aus $s < k$ Elementen besteht, und deren Linear-Kombinationen die restlichen Elemente der Folge sind. O.b.d.A. können wir (allenfalls durch Neunummerierung er Elemente!) annehmen, dass die ersten s Elemente linear unabhängig sind, und die Elemente $\mathbf{v}_j, s < j \leq k$ davon linear abhängig sind.

Unter Anwendung des vorigen Satzes kann man die lineare Abbildung dann durch die Wahl $T(\mathbf{v}_j) := \mathbf{w}_j, 1 \leq j \leq s$ festlegen. Da wegen der Gleichung (103) auch die Elemente $\mathbf{w}_j, s < j \leq k$ mit den *gleich Koeffizienten* aus den ersten Elementen (mit $j \leq s$) gewonnen werden können, ist klar, dass aufgrund der Linearität automatisch gilt: $T(\mathbf{v}_j) = \mathbf{w}_j$, auch für $j > s$. Somit ist die Behauptung bewiesen. □

lindepVW1

Remark 2. Im Falle $\boldsymbol{V} = \mathbb{R}^n$ ist die Situation besonders einfach. Wenn die Elemente $(\mathbf{v}_j)_{j=1}^n$ eine Basis bilden, dann geht es um die Fragt: Gibt es eine $n \times n$-Matrix $\boldsymbol{A}$ sodass $\mathbf{W} = \boldsymbol{A} * \mathbf{V}$? Die Antwort ist klar, da ja $\mathbf{V}$ eine invertierbare Matrix ist. Somit muss gelten: $\boldsymbol{A} = \mathbf{W} * \mathbf{V}^{-1}$.

13.9 Lineare und Nichtlineare Abbildungen

Es ist eine gute Übung, einen Begriff dadurch einzuüben, dass man sich Beispiele, aber auch Gegenbeispiele genauer anschaut.

Beispielsweise ist die Abbildung, die durch $T(x) = ax + b$ gegeben ist, offenbar für jedes feste paar von reellen Zahlen $a, b \in \mathbb{R}$ eine wohldefinierte Abbildung, aber nur dann eine lineare Abbildung, wenn $b = 0$ gilt!! (bitte auch das selbst überprüfen, es gibt viele verschiedene Ansätze, das zu tun, es muss ja nur *eine!* Verletzung der definierenden Eigenschaften gefunden werden.

Hier tut sich ein Konflikt mit der in der Mittelschule in Österreich festgelegten Terminologie (sogar für die Zentralmatura) und der international üblichen Bezeichnung in der mathematischen Fachliteratur! Die Funktion $f(x) = kx + d$ (deren *Graph* eben eine Gerade, sozusagen eine "gerade" *Linie* in der Ebene ist) wird in der Schul-Literatur als lineare Abbildung bezeichnet, während sie sonst eben als *affine* Abbildung gilt. Für den Kontext dieser Vorlesung werden wir sie eher als eine Polynomfunktion vom *Grad* 1 ansehen, d.h. eine Polynomfunktion mit (offensichtlich) 2 Koeffizienten, also *Ordnung* 2, sodass klar ist, dass es genügt, diese Funktion an zwei Stellen zu kennen um daraus ihren gesamten Verlauf rekonstruieren zu können. In der Tat, die Polynomfunktion mit Grad n bilden einen n-dimensionalen Vektorraum.

`linglgs`

Definition 18. Ein **lineares Gleichungssystem** in n Variablen ist ein Gleichungssystem, das sich in der Form einer Matrix-Gleichung der Form $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ schreiben lässt. Es heißt **homogen** wenn die rechte Seite null ist, d.h. $\mathbf{b} = \mathbf{0}$ gilt.

Ein anderer, wichtiger Satz lautet:

Theorem 16. *Die allgemeine Lösung eines inhomogenene linearen Gleichungssystems* $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ *besteht aus einer beliebigen (sog. speziellen) Lösung von* $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ *plus der allgemeinen Lösung des homogenen Gleichungs- Systems* $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$.

Diese Lösungsmenge ist selbst ein Teilraum von $\mathbb{R}^n$*, genannt der* **Nullraum** *oder* **Kern** *von* $\boldsymbol{A}$*, wir schreiben auch* $\mathrm{Null}(\boldsymbol{A})$ *(English: 'kernel').*[70]

[70] Der OCTAVE/MATLAB Befehl `null(A)` liefert eine Orthonormalbasis für eben diesen Nulllraum, welcher ein Teilraum des $\mathbb{K}^n$ ist, des Definitionsbereichs der linearen Abbildung $T : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$.

14 Geometrische Deutung Linearer Abbildungen

Viele geometrisch interessante (lineare) Abbildungen sind durch wichtige Matrizen realisierbar:

1. Spiegelung an der $x-y$-Ebene:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{104}$$

   Etwas komplizierter ist die Spiegelung an der Ebene (durch den Ursprung) mit Normalvektor (nur als Beispiel) $\vec{\mathbf{n}} = [1;1;1]$ (Diagonale im Einheistwürfel). Die normierte Version dieses Vektors ist $\frac{1}{\sqrt{3}}[1;1;1]$, und daher ist die Projektion in die Richtung dieses Normalvektors gegeben durch:

$$P_n := \frac{1}{3}(\vec{\mathbf{n}}' * \vec{\mathbf{n}}) = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

   Die Spiegelung an der Ebene ist daher $\boldsymbol{N}_3 - 2P_n$, also ist die Matrix dieser Spiegelung von der Form

$$\frac{1}{3}\begin{pmatrix} 1.0000 & -2.0000 & -2.0000 \\ -2.0000 & 1.0000 & -2.0000 \\ -2.0000 & -2.0000 & 1.0000 \end{pmatrix} \tag{105}$$

   Man überzeugt sich leicht davon, dass $\boldsymbol{A} * \boldsymbol{A} = \boldsymbol{N}_3$ gilt.

2. Projektion auf die $x-y$-Ebene

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{106}$$

3. Einbettung der $x-y$-Ebene (oder $\mathbb{R}^2$) in den $\mathbb{R}^3$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{107}$$

4. Drehung in der Ebene um 45° im math. pos. Sinne:

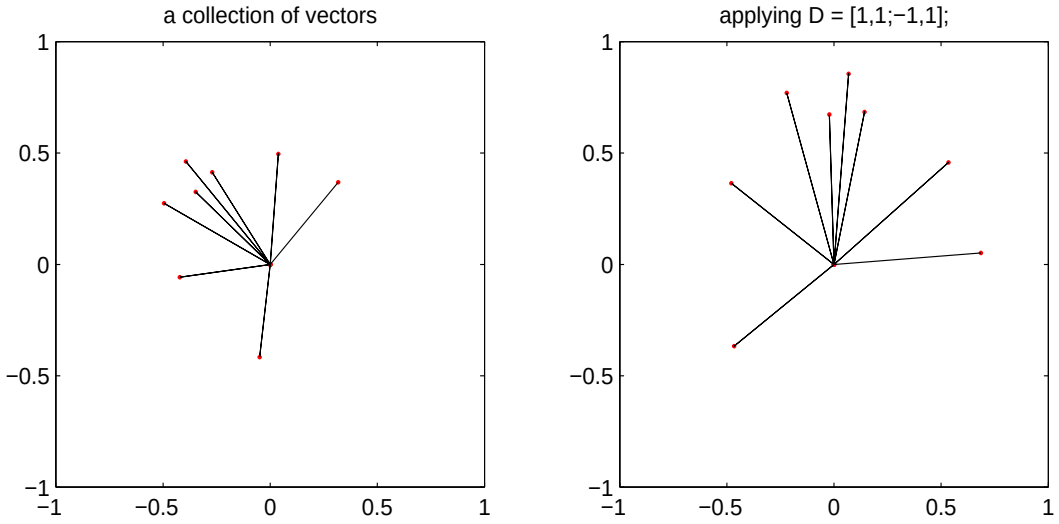
$$\boldsymbol{A} = \begin{pmatrix} 0.7071 & 0.7071 \\ -0.7071 & 0.7071 \end{pmatrix} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1.0000 & 1.0000 \\ -1.0000 & 1.0000 \end{pmatrix} \tag{108}$$

5. $\boldsymbol{A} * \boldsymbol{A}$, d.h. Drehung um 90°:

$$\begin{pmatrix} 0.0000 & 1.0000 \\ -1.0000 & 0.0000 \end{pmatrix} \tag{109}$$

6. Drehstreckungsmatrix: Kombination von Drehung um 45 Grad und Streckung um $\sqrt{2}$.

$$D = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{110}$$



7. Drehstreckungen

$$\begin{pmatrix} r\cos(\varphi) & -r\sin(\varphi) \\ r\sin(\varphi) & r\cos(\varphi) \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{111}$$

for values a, b satisfying $a^2 + b^2 = r^2$, or $r = \sqrt{a^2 + b^2}$.

Test: Diese Matrizen bilden eine kommutative Untergruppe von $GL(n, \mathbb{R})$.

In der Tat, diese Matrizen entsprechen in eindeutiger Weise der Gruppe der komplexen Zahlen ungleich 0, mit gewöhnlicher, punktweiser Multiplikation.

Die Komposition zweier solcher Matrizen ergibt eine gleichartige Matrix, und ist ausserdem *unabhängig von der Reihenfolge* der Matrizen.

$$\begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} * \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} * \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} \tag{112}$$

mit $r = r_1 \cdot r_2$, $\varphi = \varphi_1 + \varphi_2$.

Weitere Erläuterungen zur geometrischen Deutung von Matrizen (als Abbildungen).

Beispiel: Spiegelung an der ersten Mediane, d.h. an der Geraden $y = x$:

Geometrische Beobachtung: Es geht um die Spielung an einer Geraden, die durch Drehung um 45° im math. positiven Sinne aus der $x-$Achse hervorgeht.

Die Spiegelung an der x-Achse ist einfach durch eine Diagonalmatrix gegeben:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{113}$$

Wir haben schon gesehen, dass die Drehstreckungsmatrix

$$\mathbf{U} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \tag{114}$$

eine Kombination von Drehung um 45° im math. pos. Sinne ist, verbunden mit einer Streckung um $\sqrt{2}$.

Die inverse Matrix hiezu ist:

$$\mathbf{U}^{-1} := UI = \begin{pmatrix} 0.5 & 0.5 \\ -0.5 & 0.5 \end{pmatrix} \quad = \quad \frac{1}{2} \cdot \mathbf{U}^t \tag{115}$$

Kombiniert man diese drei Operationen, so erhält man:

```
>> FLX = diag([1,-1])
FLX =
     1     0
     0    -1
>> AA = U * FLX * UI
AA =
     0     1
     1     0

>> AA * [3;5]
ans =
     5
     3
```

Klarerweise ist diese Matrix gerade die Matrix, welche x mit y-Koordinaten vertauscht:

Siehe auch: Multiplikation und Division in der Polarform

`http://de.wikipedia.org/wiki/Komplexe_Zahl#Polarform`

14.1 Permutationsmatrizen

Beispiel 41 (Uebungen, SS14): Lösungshinweise

Bezüglich der Schreibweisen (auch Zyklenschreibweise) für Permutationen siehe auch

`http://de.wikipedia.org/wiki/Permutation`

Eine Permutation π ist eine bijektive Abbildung von $M := \{1, 2, \cdot n\} = \mathbb{Z} \cap [1, n] \subset \mathbb{R}$ in sich. Da die Abbildung bijektiv ist, gibt es eine inverse (ebenfalls bijektive Abbildung) π^{-1}, die dadurch gekennzeichnet ist, dass $\pi^{-1}(l) = k$ genau dann wenn $\pi(k) = l$ ist, weil

$$\pi \circ \pi^{-1} = Id_M = \pi \circ \pi^{-1}$$

gelten muss. Man kann der inversen Permutation auch einen eigenen Namen geben, sagen wir ρ.

Ein Permutation von M induziert in eindeutiger Weise eine Vertauschung auf der Einheitsvektoren, indem man vorschreibt, das $\mathbf{e}_k$ auf $\mathbf{e}_{\pi k}$ abgebildet werden soll. Die Spalten der Matrix, die dieser (!!linearen) Abbildung entsprechen, sind also genau die Bilder der Einheitsvektoren, also genau die Vektoren $(\mathbf{e}_{\pi(k)})_{k=1}^n$.

Klarerweise enthält jede Spalte nur einen 1er, es werden ja nur Einheitsvektoren aufgeschrieben. Weiters wird jeder Einheitsvektor nur genau einmal verwendet, sodass in jeder Zeile (mit Nr. k) nur die Spalte, wo der k-te Einheitsvektor hingestellt wurde (also die Zeile $l = \pi(k)$) den Einser enthält.

Das heißt natürlich, dass die Position der Einser genau sagt, wie die Abbildung π aussieht. Ein Kordinatenpaar (k, l) enthält genau dann eine 1 wenn es von der Form $(k, \pi(k))$ ist. WO ich den Einser mit Nr. k finde sagt mir die Spaltennummer, wie ich den Einheitsvektor in der Matrix finde!

Das ist aber gleichbedeutend damit, dass es von der Form $(\pi^{-1}(l), l) = (\rho(l), l)$ ist, wenn $\rho = \pi^{-1}$ die inverse Permutation ist. Transpontiert man die Matrix, so hat sie also genau dort die Einsen stehen, wo die Index-Paare von der Form $(l, \rho(l)$ sind, das ist aber (wenn man den ersten Teil der Überlegung auf $\rho = \pi^{-1}$ anwendet, genau die Matrix die π^{-1}, also die inverse Permutation (ebenfalls linear) beschreibt, w.z.z.w..

Vorschau auf später: jede Permutation lässt sich als eine Komposition von *Transpositionen* darstellen (ev. auf verschiedene Arten). Eine *Transposition* ist eine Permutation, bei der nur zwei Elemente miteinander Platz tauschen.

Um das einzuüben, folgende Frage: Wie kann man eine zyklische Vertauschen (einmal zum Einüben für $n = 3$ oder $n = 4$) oder die Umkehrung der Reihenfolge auf diese Weise darstellen?

- $\pi_1 : 1 \to 2 \to 3 > 1$; oder $1 \to 2 \to 3 \to 4 \to 1$.
- $\pi_2 : 1 \leftrightarrows 3$ und 2 bleibt fix (alternative Beschreibung: $(1, 3)(2)$).

Man stelle sich vor, dass 3 oder 4 Leute am Tisch sitzen und die Sitzposition nach π_1 oder π_2 verändern wollen, aber so, dass immer nur zwei ihren Platz vertauchen, während alle anderen gleichzeitig sitzen bleiben. Wie kann das gehn?

15 Orthogonale Matrizen, Euklidische Koordinatensysteme

Ein der wichtigsten Unterklassen der invertierbaren Matrizen sind die sogenannten **orthogonalen Matrizen**:

orthmat **Definition 19.** Eine quadratische Matrix $\mathbf{U}$ reeller Zahlen heißt **orthogonal**, wenn (sie invertierbar ist und wenn gilt):

$$\mathbf{U}^{-1} = \mathbf{U}^t,$$

d.h. wenn gilt

$$\mathbf{U} * \mathbf{U}^t = \mathbf{1}_n = \mathbf{U}^t * \mathbf{U}.$$

Im Falle von *komplexen* Matrizen spricht man von **unitären** Matrizen, und verwendet die Matrix $\mathbf{U}'$ (die konjugiert, transponierte Matrix), d.h. man fordert

$$\mathbf{U} * \mathbf{U}' = \mathbf{1}_n = \mathbf{U}' * \mathbf{U}.$$

Beispiele: Spiegelung an einer Ebene, Drehung in der Ebene, etc.

Allgemeine Drehmatrix in der Ebene: $\boldsymbol{A} = [cos(\alpha), -sin(\alpha); sin(\alpha), cos(\alpha)]$ Die Orthogonalität (eigentlich, die Eigenschaft, dass Orthogonalität bei Drehung etc. erhalten bleibt) ist dann geometrisch klar, und folgt (rechnerisch) aus der bekannten Formel $sin^2(\alpha) + cos^2(\alpha) = 1$. (vgl. G. Strang, Beispiel 4.4.1, p.230), mit Skizze, Seite 231.

Die Drehstreckungsmatrizen sind (für passendes $r > 0$) genau von der Form

$$\boldsymbol{A} = r \cdot [cos(\alpha), -sin(\alpha); sin(\alpha), cos(\alpha)]$$

oder alternativ

$$\boldsymbol{A} = r \cdot [a, -b; b, a]$$

sind isomorph zur Gruppe der von Null verschiedenen komplexen Zahlen $\mathbf{z} \in \mathbb{C}$ (mit $r := |z| = \sqrt{z \cdot \overline{z}}$) bezüglich der Multiplikation von komplexen Zahlen.

Die Punkte auf dem Einheitskreis bilden den (ein-dim.) Torus

$$\mathbb{U} = \{z \,|\, |z| = 1\}. \tag{116}$$ torus-def

Funktionen auf dem Torus (man denke sich den Graphen eine solchen, reellwertigen Funktion wie in Plakat auf einer Litfass-Säule aufgeklebt, und spaziere um diese herum, oder lasse sie gleichmässig rotieren) sind $1 - to - 1$ periodische Funktionen.

Wir studieren hier nochmals genauer den Begriff der Orthogonalität:

deforth **Definition 20.** .

- Zwei Vektoren $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^n$ bzw. $\mathbb{C}^n$ heißen **orthogonal zueinander** wenn

$$\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle = \sum_{k=1}^{n} x_k \overline{y_k} = 0 \tag{117}$$ orthskal

gilt. In Anspielung auf die geometrische Deutung der Situation sagt man auch: die beiden Vektoren *stehen aufeinander* (bzw. sind zueinander) *senkrecht* und verwendet die *symbolische Schreibweise* $\mathbf{x} \perp \mathbf{y}$.

- Zwei Mengen $M, N \subset \mathbb{R}^d$ heißen *orthogonal zueinander* wenn je zwei Vektoren $\mathbf{m} \in M$ und $\mathbf{n} \in N$ zueinander senkrecht stehen, symbolisch $M \perp N$.
- Sei $M \subset \mathbb{R}^d$ eine Menge, dann bezeichnet man die Menge aller auf M senkrecht stehenden Vektoren das **orthogonale Komplement**[71] **zu** M:

$$M^\perp := \{\, \mathbf{z} \mid \langle \mathbf{m}, \mathbf{z} \rangle = \mathbf{0} \quad \forall \mathbf{m} \in \mathbf{M} \}. \tag{118}$$ orthcompl

Man beachte, dass die Orthogonalitätsrelation tatsächlich symmetrisch ist, denn es gilt (im Komplexen, daher auch im Reellen),

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \tag{119}$$ skalpsym

und da trivialerweise $\overline{0} = 0$ gilt, ist alles klar.

Wir überlassen es den LeserInnen zu überprüfen, dass für jede Menge M die Menge $M^\perp$ ein linearer Teilraum (ein Unterraum) von $\mathbb{R}^d$ ist, und dass $M \perp N$ impliziert dass $\boldsymbol{V}_M \perp \boldsymbol{V}_N$ gilt. Es wird ähnlich bewiesne wie das folgend das folgende Lemma. Es beruht auf der Linearität des Skalarproduktes in jeder Variablen (im Falle des Körpers $\mathbb{C}$ auf der Sesequilinearität):

orthcomperz **Lemma 30.** *Es sei M eine Menge in $\mathbb{R}^d$. Dann gilt*

$$M^\perp = (\boldsymbol{V}_M)^\perp. \tag{120}$$ ortherz

Insbesondere gilt für zwei Mengen M_1 und M_2 in $\mathbb{R}^d$: Wenn die Mengen denselben Raum erzeugen, dann gilt $M_1^\perp = M_2^\perp$.

Proof. Der Beweis ist wirklich einfach. *Wenn* $M_1 \subset M_2$ gilt, dann ist klar, dass $M_2^\perp M_1^\perp$ (es sind einfach weniger Orthogonalitätrelationen zu erfüllen). Umgekehrt ist klar, dass ein Vektor $\mathbf{n} \in M^\perp$ auch auf alle Elemente des linearen Erzeugnisses von M senkrecht steht, denn per definitionem ist **nn** von der Form $\mathbf{v} = \sum_j c_j m_j$, für passende Koeffizienten, und $m_j \in M$. Dann gilt

$$\langle \mathbf{v}, \mathbf{n} \rangle = \sum_j c_j \langle m_j, \mathbf{n} \rangle = \sum_j c_j \cdot 0 = 0,$$

also liegt $\mathbf{n}$ auch im orthogonalen Komplement von $\boldsymbol{V}_M$. Somit sind die beiden Mengen gleich. □

Ein im Zusammenhang mit linearen Gleichungssystemen wichtiges Korollar (es stellt den Zusammenhang zwischen der Theorie der lin. Gleichungssysteme und den neuen Begriffen her) dazu lautet:

Corollary 4. *Es sei $\boldsymbol{A}$ eine $m \times n$-Matrix über $\mathbb{R}$*[72]*. Dann gilt: Das orthogonale Komplement des Zeilenraumes ist genau die Lösungsmenge des homogenen Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$.*

[71] Die Rechtfertigung für das Vokabel "Komplement" folgt erst später!

[72] Im Falle komplexer Matrizen (und auch sonst, begrifflich gesehen) sollte man lieber von den Spalten von $\boldsymbol{A}'$, also der konjugiert transponierten Matrix sprechen.

Proof. Es gibt nicht viel zu beweisen. Man muss nur die Aussage: "$\mathbf{x}$ gehört der Lösungsmenge an" durch Uminterpretieren in die Aussage "$\mathbf{x}$ steht auf den jeweiligen Zeilenvektor senkrecht". □

Die Begründung, warum man diese Menge das orthogonale *Komplement* nennt, wird noch später erfolgen. Dann werden wir auch zeigen, dass $\boldsymbol{V}_M = (M^\perp)^\perp$ gilt. In den Übungen werden die beiden folgenden Aussagen bearbeitet: a) Man zeige, dass für allgemeine $n \times n$-Matrizen über $\mathbb{C}$ gilt:

$$\langle \boldsymbol{A}' * \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \boldsymbol{A} * \mathbf{y} \rangle \tag{121}$$

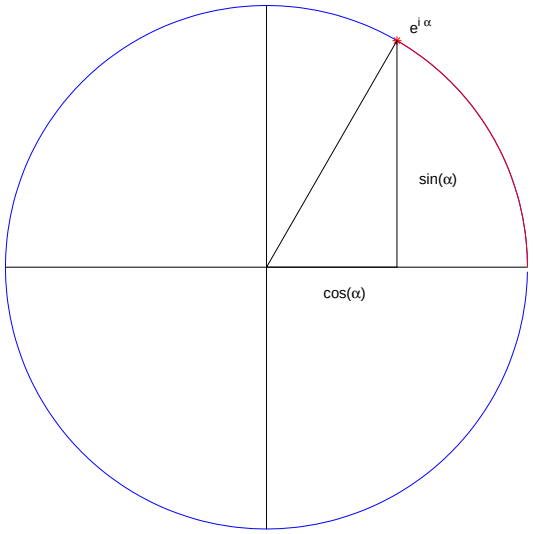
chartranspc

$$\langle \boldsymbol{A} * \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \boldsymbol{A}' * \mathbf{y} \rangle \tag{122}$$

b) Man zeige (unter Verwendung des davorstehenden Beispiels), dass der Nullraum von $\boldsymbol{A}$ mit dem Nullraum von $\boldsymbol{A}' * \boldsymbol{A}$ übereinstimmt.

Man kann daraus auch leicht ableiten, dass normale Matrizen, mit $\boldsymbol{A} * \boldsymbol{A}' = \boldsymbol{A} * \boldsymbol{A}'$ die Eigenschaft haben, dass ihr Zeilenraum gleich dem Spaltenraum ist. Vermutlich brauchen wir dazu noch die Eigenschaften des *orthog. Komplements*.

15.1 Einheitskreis und unitäre/orthogonale Matrizen



Das Bild zeigt einen Teil der *komplexen Ebene*, d.h. genauer der Einheitskreis, die Punkte haben (nach Definition der Winkelfunktionen) die Koordinaten $(\cos(\alpha), \sin(\alpha))$. Bekanntlich gilt ja $\cos^2(\alpha) + \sin^2(\alpha) = 1$.

Andere Interpretation: Wir bilden $t \mapsto e^{2\pi i t}$, dann wickeln wir die Zahlengerade auf dem Einheitskreis auf.

Die Normierung (Umfang des Einheitskreises $= 2\pi$) ist so gewählt, dass jedes Intervall der Länge 1 genau eine volle Umdrehung liefert. Es ist naheliegend, das Standard-Intervall $[0, 1]$ zu nehmen, dann startet man bei $t = 0$, also bei $e^0 = 1$, und wandert

dann innerhalb von *einer Zeiteinheit* (sagen wir eine Minute) einmal rundherum. Wenn man bei $t = 1$ angelangt ist, ist man wieder bei $e^{2\pi} = 1$!.

Der konkret gezeichnete Winkel $\alpha = 60°$ ist also nach 1/6 der Zeiteinheit (das wären bei uns 10 Sekunden) erreicht. Die zurückgelegt Strecke ist das Bogenmaß bis dorthin, also $2\pi/6 = \pi/3$.

Die *neue* Sichtweise auf den Einheitskreis ist nun die Interpretation als *Drehmatrix*. Die Drehmatrizen haben genau dieselben Eigenschaften. Dreht man um den Winkel α, so ergibt ebenfalls eine Drehung um α, dann um β eine Drehung um $\alpha + \beta$, aber ebenfalls *modulo* Vielfache von 360°, etc..

(Unterlagen: Geogebra file `sincosdemo1.ggb` (März 2014))

Andererseits sind das genau die Punkte des Einheitskreises, aufgefasst als Teil der komplexen Ebene. Normalerweise spricht man vom (ein-dimensionalen) Torus $\mathbb{U}$:

$$\mathbb{U} \quad \text{oder} \quad \mathbb{U} := \{z \,|\, |z| = 1\}$$

In diesem Fall hat man die (kommutative Gruppeneigenschaft) durch die gewöhnliche Multiplikation von komplexen Zahlen. Klarerweise gilt auch $1/z = conj(z) = \overline{z}$, denn $|z| = \sqrt{z \cdot \overline{z}} = 1$, wenn $u = (\text{real}(u), \text{imag}(u)) = (a, b)$ auf dem Einheitskreis liegt ($a^2 + b^2 = 1$!). Davon ausgehend kommt man zur Polardarstellung von allgemeinen komplexen Zahlen: $z = r \cdot u$, mit $r = |z|$. Für $z \neq 0$ ist die Darstellung eindeutig!

16 Lineare Unabängigkeit

Einer der natürlichen, ebenso wichtigen Begriffe, der sich im Kontext von allgmeinen Vektorräumen formulieren lässt, ist der Begriff der linearen Unabhängigkeit. Die *traditionelle* Definition lautet wie folgt:

`linunabhdef` **Definition 21.** Eine endliche Folge von Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_l$ in einem Vektorraum $\boldsymbol{V}$ über einem Körper $\mathbb{K}$ heißt **linear unabhängig**, wenn sich aus ihren Linearkombination der Null-Vektor $\mathbf{o} \in \boldsymbol{V}$ nur auf *triviale Weise* darstellen lässt, nämlich durch die Wahl der Koeffizientenfolge $(\lambda_k) = (0)_{1 \leq k \leq l}$ (gemeint ist die Nullfolge der Länge l).

Mehr formal ausgedrückt, ist die Bedingung erfüllt wenn gilt: [73].

$$\sum_{k=1}^{l} \lambda_k \mathbf{v}_k = \mathbf{0} \quad \Rightarrow \quad \lambda_k = 0, \quad \text{for } k = 1, \ldots l \tag{123}$$ `linabhcond`

Eine Menge $M \subset \boldsymbol{V}$ heißt *linear unabhängig*, wenn jede Folge von verschiedenen Elementen aus M eine linear unabh. Folge ist.

`basisdef` **Definition 22.** Ein (endliches) linear unabhängiges Erzeugendensystem in einem Vektorraum $\boldsymbol{V}$ wird eine **Basis** von $\boldsymbol{V}$ genannt.

[73] Da die Linearkombination von Nullen immer den Null-Vektor ergibt, könnte man anstelle des Implikationspfeils auch den Äquivalenzpfeil $\Leftrightarrow$ in dieser Stelle schreiben, ohne den Begriff zu verändern. Da die Überprüfung der Umkehrrichtung aber in jedem Falle erfolgreich sein wird, somit also unnötig in jeder konkreten Situation, ist es vorteilhaft, die Überprüfung in der angegebenen Form zu realisieren

Eine andere Sprechweise ist in dem Wort *Koordinatensystem* enthalten. Die Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ bilden ein Koordinatensystem für $\mathbb{K}^n$ wenn jeder Vektor $\mathbf{x} \in \mathbb{K}^n$ eindeutig einem $n-$Tupel von Koeffizienten zugeordent werden können. Man mache sich klar, warum diese beiden Begriffe tatsächlich logisch gleichwertig sind![74]

endldimVR **Definition 23.** Hat ein Vektorraum $\boldsymbol{V}$ eine endliche Basis, dann wird der Vektorraum *endlichdimensional* genannt[75].

Anmerkung: Die Texte in diesem Abschnitt werden einerseits noch inhaltlich oder in ihrer Reihenfolge neu geordnet. Die Erklaerungen sind eher wortreich, die Argumentationen eher redundant, statt minimalistisch, um auch die unterschiedlichen Argumentationspfade besser auszuleuchten, und diverse Terminologien, die sich nicht in Form von Theorem und Definitionen fassen lassen, einzuueben!

Es gibt ein paar einfache und darüber hinaus noch ein paar recht nützliche Beobachtungen, betreffend linear unabhängige Mengen:

Lemma 31. *Jede Teilmenge einer linear unabh. Menge ist linear unabhängig.*

linabhcrit1 **Lemma 32.** *1. Eine Menge der Form $M = \{\mathbf{v}_0\}$, bestehend aus einem(!) Element, ist linear unabhängig genau dann wenn $\mathbf{v}_0 \neq \mathbf{0}$.*

2. Eine Menge der Form $M = \{\mathbf{v}_1, \mathbf{v}_2\}$ ist genau dann linear (!) abhängig, wenn die beiden Vektoren skalare Vielfache voneinander sind[76]*;*

Proof. Die erste Aussage ist trivial (d.h. nicht, dass man sie nicht *bedenken* sollte). Im zweiten Fall ist ebenfalls klar, dass die Eigenschaft (linabhcond 123) genau dann der Fall ist (durch das "auf die andere Seite bringen eines Terms", wenn die Gleichung $\lambda_1\mathbf{v}_1 = -\lambda_2\mathbf{v}_2$ gilt, für passende Skalare $\lambda_1, \lambda_2 \in \mathbb{K}$. Da die Koeffizientenfolge nicht identisch Null ist, ist einer der Terme (wir können o.B.d.A. annehmen, das dies der erste ist, sonst argumentiert man analog f.d. zweiten), beispielsweise $\lambda_1 \neq 0$. Man kann die Gleichung daher durch λ_1 dividieren, somit bekommt man

$$\mathbf{v}_1 = -\lambda_2/\lambda_1 \cdot \mathbf{v}_2.$$

Die Umkehrung ist klar (Rückformung der Ausdrücke).[77] □

[74] Das Gegenteil von endlichdimensional ist in der normalen Sprachweise *unendlich-dimensional*, obwohl damit nicht gemeint ist, dass es eine effektive, konstruierbare Basis aus unendlich vielen Elementen gibt! Der Begriff "nicht-endlichdim." wäre also sprachlich korrekter!

[75] Wir bewegen uns nun in Richtung auf den Begriff der Dimension eines Vektorraums, d.i. die Anzahl der Basis Elemente. Diese ist nämlich konstant, d.h. für alle beliebigen Basen eines gegebenen Raums gleich. Man nennt einen derartigen Begriff eine *Invariante*.

[76] Es ist ein beliebtes Mißverständnis zu glauben oder zu behaupten, dass auch etwa bei drei Vektoren dieselbe gelte. Kurz gesagt: lieben drei verschiedene Vektoren in einer Ebene, so ist typischerweise keiner eine Vielfaches eines anderen, und trotzdem sind diese Vektoren linear abhängig. D.h. in diesem Falle ist EINER der Vektoren eine Linear-Kombination von ZWEI anderen. Nur wenn man zwei Vektoren hat, dann bedeutet lineare Abhängigkeit, dass ein Vektor eine Linearkombination "der anderen", das ist aber nur EINER, also ein Vielfaches des anderen, ist!

[77] Für mehr als drei Vektoren gibt es keinerlei bestimmte Regeln. Der typische Fall ist, dass jeder der drei Vektoren als Linear-Kombination der anderen darstellbar ist, wenn man z.B. beliebig drei Vektoren in der Ebene nimmt, aber es gibt bel. viele Möglichkeiten, linear Abhängigkeit vorzufinden!

unabhbild **Lemma 33.** *Ist T eine lineare Abbildung von $\boldsymbol{V}$ nach $\boldsymbol{W}$, und ist M eine linear abh. Menge, so ist auch $T(M) = \{T(\mathbf{v}) \,|\, \mathbf{v} \in M\}$ eine linear abhängige Menge.*

Insbesondere gilt für invertierbare, lineare Abbildungen (sog. Isomorphismen von $\boldsymbol{V}$ nach $\boldsymbol{W}$): M *ist* **linear abhängig** $\Leftrightarrow T(M)$ *ist* **linear abhängig.**

Proof. Der Beweis ist ganz einfach. Wenn zwischen gewissen Elementen $(\mathbf{v}_k)_{k=1}^n$ aus M eine nicht-triviale Relation der Form $\sum_{k=1}^n \lambda_k \mathbf{v}_k = \mathbf{0}$ existiert, so gilt natürlich auch wegen der Linearität von T (führt Linear-Kombinationen in Linear-Kombinationen über):

$$\sum_{k=1}^{n} \lambda_k T(\mathbf{v}_k) = T\left(\sum_{k=1}^{n} \lambda_k \mathbf{v}_k\right) = T(\mathbf{0}) = 0, \tag{124}$$

linabtrans

somit kann dieselbe nicht-triviale Lin. Komb. gebildet werden.

Wir sehen also (nur eine leicht veränderte logische Sichtweise) derselben Aussage: Wenn das Bild $T(M)$ linear unabh. ist, dann muß auch M selber linear unabh. sein. Im Falle einer invertierbaren linearen Abbildung ist ja auch die inverse Abbildung T^{-1} wieder linear, und das soeben gebrachte Argument wird symmetrisch. □

Gauss-spalt **Lemma 34.** *Die einzelnen Schritte der Gauss-Elimination bewahren die lineare Unabhängigkeit von Kollektionen von Spaltenvektoren*[78]

Alles was man dazu braucht ist die Beobachtung, dass für jeden beliebigen (aber festgehaltenen) Spalten-Index die neue Spalte (nach Anwendung eines der Elementarschritte) genau das Bild der alten Spalten (an derselben Stelle) ist, unter eben dieser invertierbaren Matrix. Dies garantiert auch, dass lineare Abhängigkeiten ebenso wie lineare Unabhängigkeit erhalten bleiben.

Das folgende Lemma charakterisiert die *Basen* als die *minimalen Erzeugersysteme.*

basminerz **Lemma 35.** *Ein Erzeugendensystem $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ ist eine Basis für einen Vektorraum $\boldsymbol{V}$ genau dann wenn dieses minimal ist, d.h. wenn es die Eigenschaft hat, dass die Entfernung eines beliebigen Elementes aus dieser Menge den Verlust der Eigenschaft, ein Erzeugersystem zu sein, zur Folge hätte.*

Proof. Zuerst zeigen wir, dass man aus einer Basis kein Element entfernen kann, ohne die Erzeuger-Eigenschaft zu verlieren (bzw. zu zerstören). Das ist ganz einfach. Durch Umnumerieren der Elemente können wir unsere Diskussion auf die Entfernung des letzen Elementes einer Basis $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ beschränken. Diese Element hat die nicht-triviale Darstellung $\mathbf{v}_d = 1 \cdot \mathbf{v}_d$. Damit ist aber ausgeschlossen (wegen der allgemeinen Eindeutigkeit der Darstellung von bel. Vektoren in $\boldsymbol{V}$ durch diese Basis), somit ist diese spez. Basis-Element eben nicht mehr darstellbar[79].

[78] Man stelle sich das so vor: Wenn beispielsweise in einer 5×5-Matrix $\boldsymbol{A}$ die dritte und die fünfte Spalte jeweils von den vorhergehenden Spalten linear abh. sind, d.h. sich als Linearkombination derselben darstellen lassen, dann gilt dasselbe auch - im Sinne einer Äquivalenz, für die reduzierte Zeilenstufenform von $\boldsymbol{A}$, an der die lineare Unabhängigkeit von Teilmengen leicht erkennbar ist. Es gibt nämlich eine einfache aber wichtige Identifizierung der größten linear unabh. Teilmenge von Spaltenvektoren, nämlich diejenigen (in der **ursprünglichen Matrix**) welche die Pivot-Elemente enthalten.

[79] Zusammenfassend formuliert: Wir haben also gezeigt, dass eine Basis jedenfalls in minimales Erzeugersystem ist, denn jede mögliche Entfernung eines Elementes führt zu Problemen (eben das entfernte Element kann dann nicht mehr dargestellt werden!).

Sei nun M' ein beliebiges Erzeugersystem für M. Wenn es nicht minimal ist, dann kann man wenigstens ein Element entfernen, ohne die Erzeuger-Eigenschaft zu verletzen. Diese eine Element muß aber dann eine nicht-triviale Linearkombination der übrigen sein, was dasselbe ist, wie das Fehlen der linearen Unabhängigkeit. □

linunabA **Lemma 36.** *Wenn* [80] *eine Folge von n Vektoren* $\mathscr{B} = (\mathbf{b}_k)_{k=1}^n$ *in* $\boldsymbol{V}$ *linear unabhängig ist, und* $\boldsymbol{A}$ *eine invertierbare* $n \times n$*-Matrix ist, dann ist auch die Familie* $\mathscr{C} := \boldsymbol{A}(\mathscr{B})$ *eine Basis, wobei wir darunter die Kollektion* $\mathscr{C}$ *verstehen wollen, die sich aus den Basisvektoren in* $\mathscr{B}$ *durch spaltenweise (die k-te Spalte von* $\boldsymbol{A}$ *enthält die Koeffizienten die* $\mathbf{c}_k$ *bestimmen) Wirkung der Matrix entsteht, dann ist auch* $\mathscr{C}$ *eine Basis für* $\boldsymbol{V}$*. Genauer*

$$\mathbf{c}_k := \sum_{k=1}^{n} a_{j,k}\mathbf{b}_k\,, \quad 1 \le k \le n. \tag{125}$$ lincombB

17 Beispiele von Basen

- Klarerweise ist die Familie der "üblichen Einheitsvektoren" $(\mathbf{e}_k)_{k=1}^n$ eine Basis f.d. $\mathbb{K}^n$, meistens als **Standardbasis** des $\mathbb{K}^n$ bezeichnet (! Beweis > Übung: was ist eigentlich zu beweisen: !jeder Vektor ist auf !eindeutige Weise darstellbar).

  `MATLAB: n=3; e2 = zeros(n,1); e2(2) = 1; gibt: e2 = [0;1;0];`

- Analog kann man im Raum $\mathcal{M}_{m,n}$ aller (quadr. oder rechteckigen) Matrizen (über $\mathbb{K}$) analog sog. Einheitsmatrizen bilden. Ich werde im weiteren die Numerierung der Positionen in der MATLAB Anordnung verwenden.

  `MATLAB: E3 = zeros(2); E3(3) = 1; ergibt:`

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{126}$$

- Für den Vektorraum $\mathscr{P}_n(\mathbb{C})$ (genauso natürlich für $\mathscr{P}_n(\mathbb{R}), \mathscr{P}_2(\mathbb{R}), \mathscr{P}_3(\mathbb{R})$ etc.) sind die *Monome* die natürliche Basis, d.h.

$$(t^k)_{k=0}^n \quad , 0 \le k \le n.$$

  Man beachte, dass eine Basis eine Folge (mit einer festgelegten Reihenfolge ist), denn klarerweise ist es wichtig, welche Koeffizienten zu welchem Basis-Element gehört. Die Eigenschaft, eine Basis zu sein, hängt klarerweise (!?) nicht von der Reihenfolge ab, in der die Vektoren genannt werden, für das konkrete Rechnen mit Basen spielt es sehr wohl eine Rolle.

[80] Es gilt in gewissem Sinne auch die Umkehrung. Wenn eine Matrix nicht invertierbar ist, dann erzeugt sie lineare Abhängigkeiten und somit ist die Kollektion aus Linear-Kombinationen der urspünglichen Basis keine Basis, sogar für keine beliebige Basis. D.h. es gibt nicht nur ein Problem mit der Allgemeinheit der oben gemachten Aussage, sondern es gibt gundsätzlich und immer Probleme. Derzeit mag diese (abstrakte) Argumentationsweise schwierig erscheinen, sie ist aber eine relativ einfache Beobachtung dar.

Deshalb ist es auch wichtig anzumerken, dass beispielsweise MATLAB mit der umgekehrten Reihenfolge der Monome arbeitet:

$$(t^{n-k})_{k=0}^{n}$$

Diese Ordnung entspricht dem, was man mit normalen Dezimalzahlen macht. Die Zahl 123 entspricht hier dem *Wert* des Polynoms $1 \cdot x^2 + 2 \cdot x + 3$ an der Stelle $x = 10$. Der *Stellenwert* einer Ziffer (der Einser repräsentiert den Hunderter) ergibt sich aus der Länge der Ziffernfolge (hier drei) und der Position innerhalb der Ziffernfolge. *Führende* Nullen können weggelassen werden, weiter hinten ist das nicht so. Konkret gilt klarerweise $100 \neq 1$, aber $001 = 1$.

`polyval([1,2,3],10) ergibt in MATLAB die Antwort: ans = 123`

- Konkrete Beispiele: Die Vektoren $\{[1, -1, -0], [1, 1, 0]\}$ bilden eine Basis für die $x - y$-Ebene in $\mathbb{R}^3$. Denn jeder Vektor der Form $[x, y, 0]$ kann als (eindeutige) Linear-Kombination der Vektoren $[1, -1, 0]$ und $[1, 1, 0]$ dargestellt (die ersten beiden Koordinaten bilden eine Basis im $\mathbb{R}^2$) und die dritte ist jeweils Null für alle Beteiligten.

- Die Funktionen $\cos(t)$ und $\sin(t)$ erzeugen einen zwei-dimensionalen Raum (über $\mathbb{R}$ aber auch über $\mathbb{C}$), nämlich die Menge aller ihrer (komplexen) Linear-Kombinationen. In der Tat, sie bilden ein Basis für diesen Vektorraum, denn sie sind offenbar linear unabhängig. Das ist klar, weil sie offenbar nicht vielfache voneinander sind (allgemeines Prinzip), aber auch weil $f(t) = \alpha \cos(t) + \beta \sin(t) = 0,\ \forall t \in \mathbb{R}$ impliziert klarerweise $(\alpha, \beta) = (0, 0) \in \mathbb{C}^2$ (man wähle einfach $t = 0$ und $t = \pi$).

  Derselbe Raum hat aber auch eine andere Basis, nämlich die Funktionen e^{it} und e^{-it}. Diese sind ebenfalls linear unabh. und erzeugen denselben Raum. Dieser Raum ist also zwei-dimensional.

  Dieses Beispiel wird auch noch im Kapitel *"Basiswechsel"* verwendet.

- Der Raum aller trigonometrischen Funktionen vom Maximal-Grad n wird definiert als das lineare Erzeugnis der Funktionen $e^{ikt}, 0 \leq |k| \leq n$.

  Man kann/muss zeigen, dass diese $2n + 1$ Funktionen linear unabhängig sind (z.B. indem man zeigt, dass sie bzgl. eines geeigneten Skalarproduktes ein Orthonormalsystem bilden. Diese Beobachtung kann als Basis für die *Fourier Transformation*, und zwar in den verschiedenen Ausformungen als diskrete FT (DFT), oft als FFT (fast FT) realisiert, oder der entsprechenden Integraltransformation, die vor ca. 200 Jahren entdeckt wurde).

Wenn der Vektorraum $\boldsymbol{V} = \mathbb{K}^m$, aufgefasst als Raum von Spaltenvektoren er Höhe m ist, so ist auch diese Eigenschaft einer Kollektion von Spaltenvektoren durch entsprechende Eigenschaften der (linearen) Abbildung $T : \mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ *charakterisieren*:

linunabhmat

Theorem 17. *Die folgenden Eigenschaften sind äquivalent:*

1. *Die Spalten der Matrix $m \times n$-Matrix $\boldsymbol{A}$ sind linear unabhängig;*
2. *Die Abbildung $\mathbf{x} \to \boldsymbol{A} * \mathbf{x}$ von $\mathbb{K}^n$ nach $\mathbb{K}^m$ ist injektiv;*
3. *Der Nullraum von T ist trivial, d.h.*

$$Null(T) := \{\mathbf{x} \mid T(\mathbf{x}) = \boldsymbol{A} * \mathbf{x}\} = \{\mathbf{0}\}! \tag{127}$$

nulltriv1

Proof. Die Äquivalenz von (1.) und (2.) ist wirklich nichts anderes als eine andere Beschreibung auf der verbalen bzw. symbolischen Ebene, sobald die Äquivalenz von (2.) und (3.) festgestellt ist. Dass die Injektivität der Abbildung T zur Folge hat (da ja jedenfalls $T(\mathbf{0}) = T(0\mathbf{x}) = 0 \cdot T(\mathbf{x}) = \mathbf{0}$ für jedes $\mathbf{x}$ gilt), dass $Null(T) = \{\mathbf{0}\}$ gilt, müssen wir nur die Umkehrung zeigen.

Alles, was wir dazu benötigen, ist dass die Abbildung T mit Addition verträglich ist, d.h. dass gilt $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ gilt. Sei also Bedingung (127) erfüllt, wir wollen zeigen, dass dann T eine injektive Abbildung ist.

Dazu betrachten wir zwei Vektoren $\mathbf{x}_1$ und $\mathbf{x}_2$ in $\mathbb{K}^n$. Wenn ihre Bilder gleich sind, d.h. wenn $T(\mathbf{x}_1) = T(\mathbf{x}_2)$ gilt, dann gilt auch

$$T(\mathbf{x}_1 - \mathbf{x}_2) = T(\mathbf{x}_1) + T(-\mathbf{x}_2) = T(\mathbf{x}_1) - T(\mathbf{x}_2) = 0,$$

also gilt $\mathbf{x}_1 - \mathbf{x}_2 \in \text{Null}(T) \Rightarrow \mathbf{x}_1 - \mathbf{x}_2 = \mathbf{0} \Leftrightarrow \mathbf{x}_1 = \mathbf{x}_2$. Somit ist die Injektivität nachgewiesen. □

Corollary 5. *Die Lösung eines inhomogenen linearen Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ (soferne überhaupt eine existiert) ist genau dann eindeutig bestimmt wenn der Null-Vektor (im Koeffizientenbereich) die einzige Lösung des homogenen linearen Gleichungssystems, d.h. des Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$ darstellt.*

allglsgLGLS

Corollary 6. *Je zwei Lösungen eines linearen, inhomogenen Gleichungssystems unterscheiden sich voneinander (additiv) durch eine Lösung des homogenen Gleichungssystems, d.h. die allgemeine Lösung eines linearen Gleichungssystems besteht aus einer (bel.) speziellen Lösung plus alle Lösungen des zugehörigen homogenene linearen Gleichungssystems.*

Die bisherigen Beobachtungen über invertierbare Matrizen können in folgendem Satz zusammengefasst werden (wir werden später diesen Satz noch erweitern und eine ganze Liste von weiteren, äquivalenten Eigenschaften finden):

invmatrbas

Theorem 18.

- *Die Spalten einer $n \times n$-Matrix $\boldsymbol{A}$ (über einem allgemeinen Körper $\mathbb{K}$) bilden eine Basis für den Vektorraum $\mathbb{K}^n$ genau dann, wenn die Matrix invertierbar ist.*
- *Die Abbildung $\mathbf{y} \mapsto \boldsymbol{A}^{-1}\mathbf{y}$ entspricht genau der Abbildung, die jedem Vektor $\mathbf{y} \in \mathbb{K}^n$ die (eindeutig bestimmten) Koordinaten bezüglich $\boldsymbol{A}$ zuordnet; später werden wir dafür die Schreibweise $\mathbf{y} = [\mathbf{x}]_{\boldsymbol{A}}$ verwenden;*
- *Die Abbildung $\mathbf{c} \to \boldsymbol{A} * \mathbf{c}$, also die Abbildung, die aus den Koordinaten in Bezug auf das Koordinatensystem, das durch die Spalten von $\boldsymbol{A}$ gegeben ist, ist dazu invers (später schreiben wir: $\mathbf{x} = \boldsymbol{A} * [\mathbf{x}]_{\boldsymbol{A}}$);*
- *Jede feste Basis vermittelt einen Automorphismus von $\mathbb{K}^n$, der die Standard-Darstellung eines Vektors $\mathbf{x} \in \mathbb{K}^n$ (bezgl. der üblichen Einheitsvektoren $\mathbf{e}_k, 1 \leq k \leq n$) in die $\boldsymbol{A}$-Koordinaten überführt.*

Proof. (a)[81] Zunächst sei $\boldsymbol{A} \in GL(n, \mathbb{K})$, d.h. eine invertierbare $n \times n$-Matrix über $\mathbb{K}$. Dann ist klar, dass die Koeffizienten $\mathbf{c} := \boldsymbol{A}^{-1} * \mathbf{x}$ erlauben $\mathbf{x}$ als Linearkombination der Spalten von $\boldsymbol{A}$ darzustellen, denn es gilt

$$\boldsymbol{A} * \mathbf{c} = \boldsymbol{A} * (\boldsymbol{A}^{-1} * \mathbf{x}) = (\boldsymbol{A} * \boldsymbol{A}^{-1}) * \mathbf{x} = \mathbf{1}_n * \mathbf{x} = \mathbf{x}.$$

Andererseits ist die Darstellung auch eindeutig, denn hat man irgendein $\mathbf{c} \in \mathbb{K}^n$ mit der Eigenschaft, dass $\mathbf{x} = \boldsymbol{A} * \mathbf{c}$, dann folgt (Anwenden von $\boldsymbol{A}^{-1}$ von links):

$$\mathbf{c} = (\boldsymbol{A}^{-1} * \boldsymbol{A}) * \mathbf{c} = \boldsymbol{A}^{-1} * \mathbf{x}.$$

(b) Wir haben nun die Umkehrung zu zeigen. Sei also eine Kollektion von n Vektoren, nennen wir sie $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in $\mathbb{K}^n$ gegeben, welche eine Basis für $\mathbb{K}^n$ bilden[82], dann ist klar, dass wir diese Vektoren zu einer $n \times n$-Matrix $\boldsymbol{A}$ zusammenfassen können. Die Frage ist, ob diese Matrix invertierbar ist. Klar ist, dass jeder Vektor $\mathbf{x}$ als Linearkombination darstellbar ist. D.h. die lineare Abbildung $\mathbf{c} \to \boldsymbol{A} * \mathbf{c} \in \mathbb{K}^n$ ist also surjektiv. Andererseits ist sie auch injektiv, als Konsequenz der Eindeutigkeit der Darstellung (ansonsten hätte man wenigstens einen Vektor $\mathbf{x} \in \mathbb{K}^n$ mit mehr als einer Darstellung, im Widerspruch zur Annahme). Da die Abbildung also surjektiv und injektiv ist, ist sie bijektiv, es gibt also eine Umkehr-Abbildung. Diese ist auch linear, d.h. mit Addition und sklararer Multiplikation verträglich, d.h. sie ist durch eine (inverse!) Matrix beschreibbar. Somit haben wir den Satz bewiesen.

□

[81] Empfehlung diese wichtigen Argumente einfach auch Auswendig zu Lernen bzw. immer wieder zu probieren, den Beweise selbst zu reproduzieren. Es ist wie eine Fingerübung/Konditionstraining.

[82] Später wird sich herausstellen, dass es genügt, die richtige Anzahl, d.h. n Vektoren zu haben, die ein Erzeugendensystem sind. Die Eindeutigkeit folgt dann aus Dimensionsgründen! > später!

18 Lineare Funktionale, Dualraum

Für jeden Vektorraum $\boldsymbol{V}$ über $\mathbb{K}$ spielen die linearen Funktionale, d.h. die Menge der $\mathbb{K}$-linearen Abbildungen von $\boldsymbol{V}$ in den Grundkörper $\mathbb{K}$ (also typischerweise $\mathbb{R}$ oder $\mathbb{C}$) eine wichtige Rolle. Wir verwenden alternativ zwei Symbole dafür, $\boldsymbol{V}'$ bzw. $\boldsymbol{V}^*$.

In der Funktionalanalysis werden die sogenannten stetigen (oder äquivalent, die sog. beschränkten) linearen Funktionale diskutiert, und geben dem Gebiet sogar den Namen. Man muss dies auch tun (und es ergibt sich eine spannende Mischung von Methoden der Theorie der allgemeinen Vektorräume und Überlegungen, wie sie in der Analysis eine Rolle spielt: Konvergenz von unendlichen Reihen, Stetigkeit, etc.), weil man sonst "pathologische Funktionale" bekäme, die uninteressant sind.

Im Fall von endlichdim. Vektorräumen[83] gibt es dieses Problem aber nicht. Die Existenz endlicher Basen macht viele Überlegungen übersichtlicher und im Rahmen der LA-Vorlesung abhandelbar.

Insbesondere kann man (für jede beliebig gewählte, aber dann festgehaltene Basis $\mathscr{B}$) jeden beliebigen Vektoren $\mathbf{v} \in \boldsymbol{V}$ durch seine Koeffizienten eindeutig beschreiben. Beschreiben wir die Basis wie folgt: $\mathscr{B} = (\mathbf{b}_k)_{k=1}^n$ haben wir konkret

$$\mathbf{v} = \sum_{k=1}^{n} c_k \mathbf{b}_k, \tag{128}$$ basvecV

d.h. die Vektoren sind durch die Koeffizienten-Folge $(c_k)_{k=1}^n$ in $\mathbb{K}^n$ eindeutig bestimmt. Um diese Funktionalität der Koeffizientenfolge $\mathbf{c} = (c_k)$ zu betonen schreiben wir ja auch $[\mathbf{v}]_{\mathscr{B}} = \mathbf{c}$. Daraus folgt aber auch, dass die Abbildung $\mathbf{v} \mapsto c_k = c_k(\mathbf{v})$ mit Addition und Skalar-Multiplikation verträglich (d.h. die Summe von Vektoren hat die Summe der entsprechenden Koeffizienten als eindeutig Koeffizienten, etc.), also eine *lineare* Abbildung ist, die von $\boldsymbol{V}$ in den Grundkörper $\mathbb{K}$ geht. Wir haben also für jedes feste $k, 1 \leq k \leq n$ ein lineares Funktional definiert (nennen wir es das $k-$te Koordinatenfunktional bzgl. der Basis). Nennen wir es für den Moment $\varphi_k : \mathbf{v} \mapsto c_k(\mathbf{v})$.

Somit haben wir n verschiedene lineare Funktionale gefunden, denn offensichtlich ist die Koeffizientenfolge die man für $\mathbf{v} = \mathbf{b}_k$ bekommt genau der $k-$te Einheitsvektor. Daraus läßt sich leicht schliessen, dass diese linearen Funktionale alle verschieden sind (Übungsaufgabe!). Wir werden aber noch mehr sehen.

Wir können mit den linearen Funktionalen auch rechnen, und insbesondere Linear-Kombination bilden, und werden sehen, dass sie insgesamt einen eigenen Vektorraum (von linearen Funktionalen auf $\boldsymbol{V}$) bilden, den sogenannten Dualraum zu $\boldsymbol{V}$, in dem die oben beschriebenen besonderen Funktionale eine Basis bilden, mit besonderem Bezug zur ursprünglich gewählten Basis $\mathscr{B}$. Sie wird deshalb (die *duale Basis* zu $\mathscr{B}$) genannt. Die definierende Eigenschaft wird durch die Gleichung (131) unten mit dem Kroneckersymbol formalisiert. dualbas2

Ein durchaus nützlicher (und nicht sehr schwer zu verstehender) Hilfssatz sagt, dass man lineare Abbildung durch ihre (gewünschte) Wirkung auf Basiselemente vorgeben kann.

[83] z.B. ein Raum von Polynomfunktionen, welche durchaus innerhalb des Raumes aller stetigen Funktionen, sagen wir auf dem Intervall $[0, 1]$ leben kann

Dieser Satz soll gegen Ende der Vorlesung vorgetragen werden! (falls zeitlich machbar).

linAbbBas **Theorem 19.**

- **Eindeutigkeitstest:** *Zwei lineare Abbildungen T_1, T_2 von einem n-dim. Raum $\boldsymbol{V}$ nach $\boldsymbol{W}$ sind identisch genau dann, wenn sie auf (irgend) einem (und daher auf jedem) Erzeugendensystem gleich sind.*
- **Vorgabe von Werten:** *Hat man eine linear unabhängige Menge von Vektoren in $\boldsymbol{V}$, sagen wir $\mathbf{v}_1, \ldots, \mathbf{v}_r, 1 \leq r \leq n$, und eine entsprechende Folge (von gewünschten Bildvektoren in $\boldsymbol{W}$) $\mathbf{w}_1, \ldots, \mathbf{w}_r$ in $\boldsymbol{W}$, dann gibt es jedenfalls eine lineare Abbildung $T : \boldsymbol{V} \to \boldsymbol{W}$ sodass $T(\mathbf{v}_j) = \mathbf{w}_j$ für $1 \leq j \leq r$ gilt.*
- **Charakterisierung***: Eine lineare Abbildung $T : \boldsymbol{V} \to \boldsymbol{W}$ ist eindeutig durch die Vorgabe ihrer Werte an irgendeiner Basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ in $\boldsymbol{V}$ gegeben.*

Proof. Der erste Punkt ist klar. Hat man Gleichheit auf einer Erzeugendenmenge, so kann man die Gleichheit für bel. Elemente von $\boldsymbol{V}$ durch Bilden von Linearkombinationen verifizieren (>: selber Details aufschreiben, das heißt von Worten zu Formeln kommen, und umgekehrt).

Im zweiten Punkt geht es darum, sicherzustellen, dass die zu definierende (lineare) Abbildung überhaupt auf ganz $\boldsymbol{V}$ definiert ist, und nicht nur auf dem linearen Erzeugnis $\boldsymbol{V}_M$ für $M = \{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, welcher natürlich r-dim. ist, weil die Vektoren nach Voraussetzung lin. unabh. sind, also eine Basis für $\boldsymbol{V}_M$. Es genügt daran zu erinnern, dass man jede lineare unabh. Menge durch Erweiterung ($n - r$ weitere Vektoren müssen hinzugefügt werden um zu einer Basis zu kommen) zu einer Basis von $\boldsymbol{V}$ zu machen. Wir können diesen zusätzlichen Vektoren einfach das Null-Element zuordnen (dann ist es nicht einmal wichtig, wie wir sie konkrete wählen, aber das ist im Moment nicht wichtig). Wir haben also jetzt eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ und Bildwerte $\mathbf{w}_1, \ldots, \mathbf{w}_n$. Da beliebige Elemente von $\boldsymbol{V}$ *eindeutig* als Linearkomb. der Basisvektoren in $\boldsymbol{V}$ darzustellen sind und da T Lin.Komb. respektieren soll, ist klar, dass die Bilder der allgemeinen Elemente die entsprechenden Linearkombinationen der Vektoren $\mathbf{w}_1, \ldots, \mathbf{w}_n$ sein müssen (und so auch eindeutig festgelegt sind).

Im letzten Fall ist keine Ergänzung nötig, somit ist dieser Fall auch gleich erledigt. □

extenluset **Lemma 37.** *Angenommen M ist eine linear unabhängige Menge in $\mathbb{K}^m$, und $\mathbf{x}_0 \in \mathbb{K}^m$ ein linear unabhängiges Element, dann ist $M' := M \cup \{\mathbf{x}_0\}$ eine linear unabhängige Menge.*

Angenommen es gibt eine lineare Relation zwischen bestimmten Elementen von M', mit einer von Null verschiedenen Koeffizientenfolge. Wenn der Koeffizient von $\mathbf{x}_0$ von Null verschieden ist, müßte $\mathbf{x}_0$ von M linear abhängig sein, d.h. eine Linear-Kombination von Elementen aus M sein, was aber ausgeschlossen wurde. Wenn andererseits (logische Alternative) der Koeffizient von $\mathbf{x}_0$ gleich Null ist, ist es de facto eine Linear-Kombination von Elementen von M, also ist die gesamte Koeffizientenfolge (auch in diesem Fall) trivial. Damit ist der Beweis beendet.

Am besten ist es, diese Konzept an unserem "halb-abstrakten Beispiel" zu betrachten, nämlich an den Räumen $\mathscr{P}_2(\mathbb{R})$ oder $\mathscr{P}_3(\mathbb{R})$. Typische lineare Funktionale sind dann

(noch eine Übungsaufgabe) Abbildungen wie $p(t) \mapsto p(t_0), t_0 \in \mathbb{R}$ beliebig aber fest. Normalerweise verwendet man für dieses lineare Funktional die Bezeichnung δ_{t_0}, also gilt $\delta_{t_0}(p(t)) = p(t_0)$ [84]. Weitere Beispiele sind $p(t) \mapsto \int_a^b p(t)dt$ für beliebige, feste Werte von $a, b \in \mathbb{R}$, aber auch $p(t) \mapsto p'(t)$ oder $p(t) \mapsto 3 \cdot p''(t) + p(4)$.

Hauptaufgabe ist nun also, den Dualraum $\boldsymbol{V}'$ von $\boldsymbol{V}$ zu studieren.

Dabei ist folgende einfache Überlegung hilfreich:

isodual **Lemma 38.** *Ist T ein Isomorphismus von Vektorräume $\boldsymbol{V}$ und $\boldsymbol{W}$, dann gibt es auch einen entsprechenden Isomorphismus T' von $\boldsymbol{W}'$ nach $\boldsymbol{V}'$.*

Proof. Es ist leicht zu sehen, dass für jedes lineare Funktional $\varphi \in \boldsymbol{W}'$, d.h. $\varphi : \boldsymbol{W} \to \mathbb{C}$, linear, ein neues Funktional (wir nennen es gleich $T'(\varphi)$) definiert werden kann durch

$$T'(\varphi)(\mathbf{v}) = \varphi(T(\mathbf{v})) \quad \mathbf{v} \in \boldsymbol{V}. \tag{129}$$ isodual

Diese Zuordnung ist linear. Es ist auch leicht zu zeigen dass für den inversen Isomorphismus $S = T^{-1}$ die zugehörige Abbildung $S' : \boldsymbol{V}' \to \boldsymbol{W}'$ die inverse Abbildung zu T' ist, womit die Behauptung verifiziert ist. □

Spezialfälle solcher Isomorphismen sind durch Basen $\mathscr{B}$ in einem Raum $\boldsymbol{V}$ gegeben. Für eine feste Basis $\mathscr{B}$ ist die Abbildung die $x = \sum_{k=1}^n c_k b_k$ den Koeffizienten $= [x]_{\mathscr{B}} = \mathbf{c} = (c_k)$, also $T(x) = \mathbf{c}$ ein solcher Isomorphismus zwischen dem Vektorraum $\boldsymbol{V}$ und $\mathbb{R}^n$.

Andererseits sind die linearen Funktionale auf $\mathbb{R}^n$, d.h. die Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}$ (der reelle Dualraum von $\mathbb{R}^n$) durch die $1 \times n$-Matrizen (d.h. durch die reellen Zeilenvektoren) gegeben ist. Da die *Einheitsvektoren* eine natürliche Basis für $\mathbb{R}^n$ (aufgefasst als Dualraum des $\mathbb{R}^n$, Zeilenvektoren als Funktionale, wirkend auf den Spaltenvektoren, wenn/weil man die übliche Konvention der Matrix-Multiplikation zur Umsetzung des abstrakten Prinzipes verwenden will) bilden, kommen den *entsprechenden* Funktionalen in $\boldsymbol{V}'$ eine besondere Rolle zu. Sie bilden insgesamt die **biorthogonale Basis** zu $\mathscr{B}$.

Wir können sie (eindeutig bestimmt) mit $\mathscr{B}'$ bezeichnen, und wenn wir sie mit $\varphi_1, \dots \varphi_n$ bezeichnen, so gilt die charakteristische Identität

$$\varphi_k(b_j) = \delta_{j,k}. \tag{130}$$ biorth

Wichtige Bemerkung: Betrachtet man (um konkreter zu sein) 4 Punkte auf der Zahlengeraden t_1, t_2, t_3, t_4, und berechnet man die Lagrange-Interpolations-Polynome für diese Folge in $\mathscr{P}_3(\mathbb{R})$ aus, so sind die Punktauswertungen $p(t) \mapsto p(t_i)$ genau das Biorthogonalsystem zum System dieser Funktionen, welche in der unnormierten Version als $L_1(t) = (t - t_2)(t - t_3)(t - t_4), \dots, L_4 = (t - t_1)(t - t_2)(t - t_3)$.

Ähnlich kann man behaupten, dass beispielsweise (siehe Übungen) $\mathscr{M}_3 := \{1, t, t^2, t^3\}$ als duale Basis die Funktionale $\varphi_1(p(t)) = p(0)$, $\varphi_2(p(t)) = p'(0)$, $\varphi_3(p(t)) = p''(0)/2$ und zuletzt $\varphi_3(p(t)) = p'''(0)/3$ hat (Beweis als Übungsaufgabe).

Man kann in jedem Falle zeigen, dass die sog. "duale Basis" in der Tat eine Basis für $\boldsymbol{V}'$ (der Name is also *gerechtfertigt*).

[84] Diese δ ist verschieden vom Kronecker-Delta!

Proof. Man muss zeigen, dass eine Linearkombination von Elementen der dualen Basis nur dann verschwinden kann, wenn die verwendeten Koeffizienten die Nullfolge sind. Andererseits ist zu zeigen, dass jedes lineare Funktional sich als Linearkombination dieser speziellen linearen Funktionale darstellen lässt. Dann bilden sie eine Basis für den Dualraum.

Befassen wir uns mit der Erzeugenden-Eigenschaft. Sei $\varphi \in \boldsymbol{V}'$ eine beliebige lineare Abbildung von $\boldsymbol{V}$ nach $\mathbb{C}$, und sie $\mathscr{B}$ eine feste Basis für $\boldsymbol{V}$. Dann gilt

$$\varphi(\mathbf{v}) = \varphi\left(\sum_{k=1}^{n} c_k b_k\right) = \sum_{k=1}^{n} c_k \varphi(b_k).$$

Wir können das lineare Funktional φ aber auch einfach auf die Basis-Elemente loslassen, und definieren $d_k := \varphi(b_k)$. Wegen der Biorthogonalität haben wir dann auch $\varphi(b_k) = d_j = \sum_{j=1}^{n} d_j \varphi_j(b_k), 1 \leq j \leq n$, und daher

$$\varphi(\mathbf{v}) = \sum_{k=1}^{n} c_k \sum_{j=1}^{n} (d_j\varphi_j)(b_k) = \sum_{j=1}^{n} d_j\varphi_j(\sum_{k=1}^{n} c_k b_k) = \sum_{j=1}^{n} (d_j\varphi_j)(\mathbf{v}).$$

Da diese Relation für alle $\mathbf{v} \in \boldsymbol{V}$ gilt, haben wir $\varphi = \sum_{k=1}^{n} d_j\varphi_j$, was zu zeigen war. □

Insbesondere folgt aus dieser Beobachtung:

dualdim **Corollary 7.** *Die Dimension des Dualraumes $\boldsymbol{V}'$ eines endlichdim. Vektorraumes ist gleich der Dimension des zugrundeliegenden Vektorraumes $\boldsymbol{V}$.*

Ohne Beweis merken wir an, dass man für jede Basis des Dualraumes eine Basis des Originalraumes finden kann (und zwar eindeutig bestimmt), sodass das Paar biorthogonal ist. D.h. innerhalb der Menge der Basen für $\boldsymbol{V}$ bzw. $\boldsymbol{V}'$ gibt es auf diese Weise eine Paarbildung.[85]

Der "Hauptsatz" für die Feststellung von Basen in $\mathbb{R}^n$ kann auch dazu verwendet werden zu zeigen, dass eine Folge von n linearen Funktionalen auf einem $n-$dim. Vektorraum $\boldsymbol{V}$ genau dann eine Basis für $\boldsymbol{V}$ ist, wenn sie injektiv ist, d.h. wenn aus $\varphi_j(\mathbf{v}) = 0$ folgt, dass $\mathbf{v} = \mathbf{0} \in \boldsymbol{V}$ gilt.

Entsprechender MATLAB Code:

```
bas5 = linspace(-.1,kk+0.1,501);  bas5 = bas5(:);
V5 = vander(0:kk); LV5 = inv(V5);
LAG5 = zeros(501,kk+1); figure;
for jj=1:kk+1; LAG5(:,jj) = polyval(LV5(:,jj),bas5); end;
plot(bas5, LAG5);  axis tight; grid; hold on;
plot(0:kk,ones(1,kk+1),'b*');  plot(0:kk,zeros(1,kk+1),'k*');
title('Lagrange Interpolation Polynomials');  shg
```

[85] Im Falle des Vektorraums der Polynomfunktionen wird eine solche Überlegung zur Beschreibung der Lagrange Interpolationspolynome führen.

18.1 Dualräume etwas anders beschrieben

cf. Greub (cite76-4), leider etwas andere Notation.

dualeBasis **Theorem 20.** *Es sei $\boldsymbol{V}$ ein endlichdim. Vektorraum mit Basis $\mathscr{B}$. Dann gibt es eine (eindeutig bestimmte) duale Basis $\mathscr{B}^*$ für den dualen Raum, d.h. den Vektorraum $\boldsymbol{V}^*$ aller linearen Funktionale von $\boldsymbol{V}$ nach $\mathbb{K}$ (also $\mathbb{R}$ oder $\mathbb{C}$)*[86] *Beide Basen haben gleich viele Elemente, also sind insbesondere $\boldsymbol{V}$ und $\boldsymbol{V}^*$ von gleicher Dimension.*

Im Prinzip sind die "Koeffizientenfunktionale" (d.h. die Abbildungen die den Vektoren die eindeutig bestimmten Koeffizienten zu einem der festen Basisvektoren zuordnen) genau diese duale Basis.

Proof. Besteht die ursprüngliche Basis aus den Elementen $\mathbf{b}_1, \ldots, \mathbf{b}_n$, so kann man die n linearen Funktionale $\mathbf{b}_j^*, 1 \leq j \leq n$ dadurch definieren, dass man fordert

$$\mathbf{b}_j^*(\mathbf{b}_k) = \delta_{j,k} \quad \text{Kronecker} - \delta. \tag{131}$$ dualbas2

Da bekanntlich eine lineare Abbildung eindeutig durch ihre Wirkung auf den Elementen einer Basis definiert ist, ist klar, dass dies n verschiedene lineare Funktionale, d.h. Elemente von $\boldsymbol{V}^*$ darstellt. In Wirklichkeit kennen wir diese Funktionale schon: Für jedes $j, 1 \leq j \leq n$ ist die Zuordnung von $\mathbf{v} = \sum_{k=1}^n c_k \mathbf{b}_k$ auf c_j eine lineare Abbildung (wegen der Eindeutigkeit), und in der Tat genau das Funktional $\mathbf{b}_j^*$, denn

$$\mathbf{b}_k^*(\mathbf{v}) = \mathbf{b}_k^*(\sum_{j=1}^n c_j \mathbf{b}_j) = \mathbf{b}_k^*(c_k \mathbf{b}_k)! = c_k \tag{132}$$ getcoeffs

Es bleibt zu zeigen, dass ein abstraktes lineares Funktional $\mathbf{v}^* \in \boldsymbol{V}^*$ als Linearkombination dieser Funktionale $\mathbf{b}_j^*$ darstellbar ist. Lassen wir $\mathbf{v}^*$ auf die Basiselemente von $\mathscr{B}$ los und definieren $a_k := \mathbf{v}^*(\mathbf{b}_k)$. Dann gilt für $\mathbf{v} \in \boldsymbol{V}$:

$$\mathbf{v}^*(\mathbf{v}) = \mathbf{v}^*(\sum_{k=1}^n c_k \mathbf{b}_k) = \sum_{k=1}^n c_k \mathbf{v}^*(\mathbf{b}_k) = \sum_{k=1}^n c_k a_k = \sum_{k=1}^n a_k \mathbf{b}_k^*(\mathbf{v}), \tag{133}$$ dualbasrek

womit bewiesen ist, dass $\mathbf{v}^* = \sum_{k=1}^n a_k \mathbf{b}_k^*$ (in $\boldsymbol{V}^*$ gerechnet) gilt. Da - leicht zu zeigen - die linearen Funktionale linear unabhängig sind (Ex.!) ist $\mathscr{B}^* := \{\mathbf{b}_k^* \,|\, 1 \leq k \leq n\}$ eine Basis für $\boldsymbol{V}^*$ ist. □

Insbesondere ist damit auch gezeigt, dass der Dualraum $\boldsymbol{V}^*$ die gleiche Dimension hat wie der ursprüngliche Raum $\boldsymbol{V}$[87].

Weniger als bedeutsame Fakten, sondern mehr um ein wenig konkreter mit diesen allgemeinen Begriffen umgehen zu lernen (und die allgemeine Bedeutung von Basen und den damit verbundenen Sätzen zu bekommen) eine paar illustrative Betrachtungen.

[86] Die Schreibweisen $\boldsymbol{V}^*$ und $\boldsymbol{V}'$ sind beide in der Literatur geläufig, und werden in diesem Skriptum beide verwendet. Die Verwendung des $*$-Symbols ist wohl weiter verbereitet, und hat den Vorteil nichts mit dem Transposition + Konjugationsoperator zu tun zu haben.

[87] Je nach Buch werden die Symbole $\boldsymbol{V}^*$ oder $\boldsymbol{V}'$ verwendet, als Symbol für den Dualraum.

Wir erinnern daran, dass die lineare Abbildung $f \mapsto f(t)$ z.B. auf dem Raum aller stetigen Funktionen $\boldsymbol{C}(\mathbb{R})$ oder $\boldsymbol{C}(I)$, für ein beliebiges Intervall $I \subseteq \mathbb{R}$ eine lineare Abbildung auf diesem Raum ist. Man verwendet üblicherweise das Symbol δ_t für das entsprechende lineare Funktional, d.h.

$$\delta_t(f) := f(t), \quad t \in I. \tag{134}$$ deltadef

Natürlich können wir uns das Funktional auch auf $\mathscr{P}_2(\mathbb{R})$ oder $\mathscr{P}_3(\mathbb{R})$ anschauen. Wir können dann zeigen (nehmen wir Grad = 2 her), dass für je drei verschiedene Zahlen $t_1, t_2, t_3 \in I$ die entsprechende Familie der linearen Funktionale $\delta_{t_1}, \delta_{t_2}, \delta_{t_3}$ eine Basis von $\mathscr{P}_2(\mathbb{R})'$ bilden. Hat man mehr als drei Punkte, so bilden die entsprechenden Punktauswertungs-Funktionale (δ_{t_j}) sicher eine linear abhängige Menge in $\mathscr{P}_2(\mathbb{R})'$.

Der Beweis ist einfach. Wir wissen, dass $\mathscr{P}_2(\mathbb{R})$ drei-dimensional ist. Also ist auch der Dualraum $\mathscr{P}_2(\mathbb{R})'$ drei-dimensional. Damit ist sofort klar, dass mehr als 3 Elemente linear abhängig sein müssen, wie immer sie konkret aussehen würden.

Andererseits genügt es zu zeigen, dass je drei (paarweise verschiedene) Punkte zu einer linear unab. Familie von Punkt-Funktionalen $\delta_{t_1}, \delta_{t_2}, \delta_{t_3}$ führen. Dazu erinnern wir uns an die Lagrange Interpolationspolynome, oder einfach (direkt, unnormiert hingeschrieben), an $L_1(t) := (t-t_2)(t-t_3)$ etc.. Hat man nämlich $\sigma := a_1\delta_{t_1} + a_2\delta_{t_2} + a_3\delta_{t_3} = 0$, d.h. ist eine Linearkombination das Nullfunktional, das bedeutet, dass $\sigma(f) = 0$ für alle $f \in \mathscr{P}_2(\mathbb{R})$, oder $\sum_{j=1}^{3} a_j f(t_j) = 0$, insbesondere für die zur Punktfolge gehörenden Folge L_1, L_2, L_3, woraus sich aber sofort $a_1 = 0$ ergibt. Analog für $j = 2, 3$. Da klarerweise $\delta_{t_j}(L_k) = \delta_{j,k}$ gilt (da haben wir beide Versionen von δ in einer Formel), ist sogar sofort klar, dass die Folge der Delta-Funktionalen *die duale Basis* zur Basis der Lagrange-Polynome ist (die ja auch offenbar linear unabh. sind).

Interessant ist es, dasselbe Problem für Polynome vom Grad 1 (ein 2-dim. Raum) zu betrachten. Auf diesem Raum müssen die linearen Funktional $\delta_0, \delta_1, \delta_2$ linear abhängig sein (beispielsweise). Es ist aber leicht zu sehen, dass für Funktionen (Geraden) der Form $f(x) = ax + b$ gilt: $f(1) = (f(0) + f(2))/2$, das ist aber gleichbedeutend mit $\delta_1 = 0.5(\delta_0 + \delta_2)$ in DIESEM Dualraum, d.h. als Elemente von $\mathscr{P}_1(\mathbb{R})'$ betrachtet.

Zuletzt wollen wir noch die zu einer Basis $\mathscr{B}$ von Spaltenvektoren im $\mathbb{R}^n$ (die wir uns in Form einer $n \times n$-Matrix $\boldsymbol{B}$ geschrieben denken können) die *duale Basis* bestimmen. Wir suchen also n lineare Funktionale auf $\mathbb{R}^n$, welche (131) erfüllen sollen. dualbas2 Da wir wissen, dass lin. Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}$ durch Zeilenvektoren (wirkend als $1 \times n$-Matrizen auf $\mathbb{R}^n$) dargestellt werden, suchen wir also Zeilenvektoren $\mathbf{z}_1, \ldots, \mathbf{z}_n$, mit $\mathbf{z}_j * \mathbf{b}_k = \delta_{j,k}$. Wir können diese Zeilenvektoren aber zu einer grossen Matrix $\mathbb{Z}$ zusammenfügen, und dann bedeutet (131), dass $\mathbb{Z} * \boldsymbol{B} = \mathbf{Id}_n$ gelten muss. dualbas2 Mit anderen Worten, $\mathbb{Z}$ ist notwendigerweise die inverse Matrix. Somit haben wir gezeigt: *Die duale Basis zu einer Basis von Spaltenvektoren ist die Kollektion der Zeilenvektoren der inversen Matrix!*.

Wir fassen zusammen:

dualeBasChar **Definition 24.** Ist $\mathscr{B}$ eine Basis mit n Elementen, dann ist die duale Basis für den Dualraum (den Raum aller Funktionale) durch die linearen Abbildungen $\phi_k^*, 1 \leq k \leq n$ gegeben, die gerade jedem $\mathbf{v} \in \boldsymbol{V}$ mit $\sum_{k=1}^{n} c_k \mathbf{b}_k$ den k-ten Koeffizienten (in linearer Weise) zuordnen, und das ist tatsächlich eine Basis des Dualraumes $\boldsymbol{V}^*$ aller *linearen Funktionale* auf $\boldsymbol{V}$ (lineare Abbildungen in den Grundkörper).

EINSCHUB (Mai 2014):

WIKI: Hinweiss dass die Ableitungen die duale Basis zur Monomial-Basis darstellt:

`http://de.wikipedia.org/wiki/Taylorsche_Formel#Taylorpolynom`

$$T_n f(x;a) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Man kann natürlich auch die Matrix-Multiplikation als Skalarprodukt interpretieren, dann spricht man davon, dass die Zeilenvektoren $\mathbf{z}_1, \ldots, \mathbf{z}_n$ ein Biorthogonalsystem bilden, weil

$$\langle \mathbf{z}_j, \mathbf{b}_k \rangle = \delta_{j,k},\ 1 \leq j,k \leq n. \tag{135}$$ biorth2

Wenn man im Komplexen agiert, muss man noch die Konjugation berücksichtigen. Dann ist es wohl sinnvoller die Situation wie folge zu beschreiben:

dualcompl **Lemma 39.** *Es sei $\boldsymbol{B}$ eine komplex-wertige invertierbare $n \times n$-Matrix. Betrachten wir die Spalten als Basis für $\mathbb{C}^n$, dann ist das Biorthogonalsystem genau die Familie der Spalten von $(\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$.*

Die Begründing ist einfach: Das Bilden von Skalarprodukten mit den Spalten der genannten Matrix is gleichbedeutend mit der gewöhnlichen Matrix-Multiplikation mit $\boldsymbol{A}^{-1}$.

Wir fassen die Begriffe nochmals zusammen:

biorthsequ **Definition 25.** Es sei $\boldsymbol{B}$ eine $m \times n$-Matrix, aufgefasst als Kollektion von Spaltenvektoren. Dann ist eine andere $m \times n$-Matrix ein Biorthogonalsystem (wir sagen auch, die Folge der Spaltenvektoren $\mathbf{c}_1, \ldots, \mathbf{c}_n$ ist **biorthogonal** zur Folge $\mathbf{b}_1, \ldots, \mathbf{b}_n$, wenn gilt $\mathbf{C}' * \boldsymbol{B} = Id_n$ [das ist eine symmetrische Relation].

Die obigen Ausführungen ergeben ein paar einfach Konsequenzen.

rankrmaps **Lemma 40.** *Eine lineare Abbildung $T : \boldsymbol{V} \to \boldsymbol{W}$, deren Bild r-dimensional (innerhalb von $\boldsymbol{W}$) ist, kann in der Form*

$$T(\mathbf{v}) = \sum_{k=1}^{r} \varphi_k(\mathbf{v})\mathbf{w}_k \tag{136}$$ rankrT

beschrieben werden, für passende lineare Funktionale $\varphi_k, 1 \leq k \leq r$ und passenden Elemente $\mathbf{w}_k, 1 \leq k \leq r$, welche eine Basis für den Bildraum von $T(\boldsymbol{V})$ in $\boldsymbol{W}$ bilden.

Im Standard-Fall $\boldsymbol{V} = \mathbb{R}^n, \boldsymbol{W} = \mathbb{R}^m$ kann man die linearen Funktionale natürlich in der Form von Skalarprodukten schreiben, d.h. es gibt Vektoren $\mathbf{y}_k, 1 \leq k \leq r$ sodass

$$T(\mathbf{v}) = \sum_{k=1}^{r} \langle \mathbf{v}, \mathbf{y}_k \rangle \mathbf{w}_k. \tag{137}$$ rankrTR

Schreibt man diese Vektoren[88] *als Zeilen in eine Matrix* $\mathbf{Y}$*, dann kann man die Martrix* $\boldsymbol{B} := \mathbf{Y}'$ *als eine Kollektion von* r *Spaltenvektoren ansehen, die* T *repräsentierende Matrix* $\boldsymbol{A}$ *schreiben als*[89] $\boldsymbol{A} = \mathbf{W} * \boldsymbol{B}'$.

Proof. Es genügt hier, den *Ansatz* anzudeuten: Die Voraussetzung bedeutet, dass der Bildraum $T(\boldsymbol{V})$ die Dimension r hat, d.h. jede (beliebige, und für den Beweis dann festgehaltene) Basis hat r Elemente: $\mathbf{w}_k = T(\mathbf{v}_k), 1 \leq k \leq r$. Wir verwenden nun die dazu duale Basis (vor allem aus notationstechnischer "convenience"): Für jedes $\mathbf{v} \in \boldsymbol{V}$: $T(\mathbf{v}) = \sum_{k=1}^{r} \mathbf{w}_k^*(T(\mathbf{v}))\mathbf{w}_k$. Auf diese Weise ist schneller/leichter zu sehen, dass die Abbildung $\varphi_k := \mathbf{w}^* \circ T$ ein lineares Funktional auf $\mathbf{V}$ ist, welches die gewünschte Darstellung liefert.[90] □

Ohne Beweis sei hier auch noch erwähnt, dass auch die Umkehrung dieses Proposition gilt, d.h. *wenn* sich $\boldsymbol{A}$ schreiben läß als $\boldsymbol{A} = \boldsymbol{B}' * \mathbf{W}$, mit einer $r \times m$-Matrix $\boldsymbol{B}$ und einer $r \times n$ Matrix $\mathbf{W}$, *und wenn beide Matrizen maximalen Rang haben*, dann hat auch $\boldsymbol{A}$ den Rang r. Dies kann man zeigen, indem man durch streichen passender Zeilen und Spalten auf zwei invertierbare $r \times r$ Matrizen übergeht.

Natürlich bietet dieser Ansatz eine einfache Methode, schnell eine zufällige $m \times n$-Matrix vom Rang r zu finden, indem man $\boldsymbol{B}$ (bzw. $\mathbf{Y}$) einfach zufällig wählt. Fast sicher (in einem wahrscheinlichkeitstheoritsch korrekten Sinne) ist eine solche zufällige Matrix von maximalem Rang (= Minimum der Dimensionszahlen).

In der Folge werden wir uns noch genauer ansehen, wie spezielle Abbildungen in dieser Form geschrieben werden können (man beachte, dass in einigen Fällen, d.h. wenn r sehr klein ist im Verhältnis zu m und n bei der Realisierung der linearen Abbildung/Matrixmultiplikation in dieser Form wesentlich weniger Multiplikationen anfallen: man erinnere sich an die Behauptung, dass $(\boldsymbol{A} * \boldsymbol{B}) * \mathbf{x}$ mehr Rechenaufwand erfordert als $\boldsymbol{A} * (\boldsymbol{B} * \mathbf{x})$ [warum?]).

Zuerst betrachten wir ein Orthonormalsystem in $\mathbb{C}^n$ (oder $\mathbb{R}^n$), d.h. eine Matrix $\mathbf{U}$ vom Format $n \times r$ $(1 \leq r \leq n!)$ mit $\mathbf{U}' * \mathbf{U} = Id_r$.

orthoproj

Theorem 21. *Es sei* $\mathbf{U}$ *ein Orthonormalsystem von* r *Vektoren im* $\mathbb{C}^n$*. Dann ist die Projektion auf den von den Spalten von* $\mathbf{U}$ *erzeugten* $r-$*dimensionalen Teilraum* $\boldsymbol{V}_{\mathbf{U}}$ *gegeben durch*

$$P_{\mathbf{U}}(\mathbf{x}) = \sum_{k=1}^{r} \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k, \tag{138}$$

orthprojskp

oder in Matrix-Schreibweise

$$\mathbf{PU} = \mathbf{U} * \mathbf{U}'. \tag{139}$$

orthprojmat1

Dementsprechend haben wir für die Projektion auf das orthogonale Komplement

$$n = max(size(\mathbf{U})); \quad \mathbf{PUO} = eye(n) - \mathbf{U} * \mathbf{U}'. \tag{140}$$

orthprojomat1

[88] Im Fall komplexer Zahlen muss man auf die Konjugation achten!

[89] Achtung: $\mathbf{W}$ bedeutet hier die von den Vektoren $\mathbf{w}_1, \ldots \mathbf{w}_r$ gebildete $r \times m$-Matrix.

[90] Ohne den Begriff der dualen Basis müßten wir nochmals die Eindeutigkeit der Darstellung zurate ziehen, um zu zeigen, dass die Abbildung, die jedem $\mathbf{v} \in \boldsymbol{V}$ die eindeutigen Koeffiziente des Bildes $T(\mathbf{v})$ bzgl. der Basis $\mathbf{w}_1, \ldots, \mathbf{w}_r$ zuordnet, ein lineares Funktional, eben genau das, was wir φ_k^* genannt haben, ist.

Dementsprechend ist $\mathbf{Id_n} - \mathbf{PU}$ *die Projektion*[91] *auf das Orthogonale Komplement von* $\boldsymbol{V}_{\mathbf{U}}^{\perp}$ *von* $\boldsymbol{V}_{\mathbf{U}}$.

Bevor wir diesen Satz beweisen, geben wir noch eine Charakterisierung von orthogonalen Projektionen:

projections1 **Definition 26.** Eine lineare Abbildung P von $\boldsymbol{V}$ in sich heißt **Projektion** wenn $P^2 = P$ gilt[92]. Die Menge der Bildpunkte einer Projektion, also $P(\boldsymbol{V})$ bilden den Raum, auf den (der ganze Vektorraum) $\boldsymbol{V}$ *projiziert wird*

Eine solche Projektion in $\mathbb{C}^n$ oder $\mathbb{R}^n$ heißt **orthogonal** wenn gilt: für jedes $\mathbf{x} \in \mathbb{C}^n$ liegt der Vektor $\mathbf{x} - P(\mathbf{x})$ im orthogonalen Komplement von $P(\boldsymbol{V})$.

rthprojchar1 **Lemma 41.** *Eine Matrix* $\mathbf{P}$ *beschreibt eine orthogonale Projektion in* $\mathbb{C}^n$ *genau dann wenn gilt* $\mathbf{P}^2 (= \mathbf{P} * \mathbf{P}) = \mathbf{P}$ *und* $\mathbf{P}' = \mathbf{P}$.

Proof. Wir führen nur die direkte Implikation hier aus.

Es geht darum, die Invarianzeigenschaft $\mathbf{P}' = \mathbf{P}$ mit der Orthogonalität der Projektion in Verbindung zu bringen. Sei $\mathbf{P} * \mathbf{z}$ ein beliebiges Element von $P(\boldsymbol{V})$. Dann gilt

$$\langle \mathbf{x} - \mathbf{P} * \mathbf{x}, \mathbf{P} * \mathbf{z} \rangle = \langle \mathbf{P}' * (\mathbf{x} - \mathbf{P} * \mathbf{x}), \mathbf{z} \rangle = \langle \mathbf{P} * \mathbf{x} - \mathbf{P} * \mathbf{x}, \mathbf{z} \rangle = 0. \tag{141}$$ xpxperp1

womit die Behauptung bewiesen ist. □

clospts1 **Corollary 8.** *Es sei* $\boldsymbol{M}$ *eine Teilraum von* $\mathbb{C}^n$ *und* P *eine (die!) orthogonale Projektion von* $\mathbb{C}^n$ *auf* $\boldsymbol{M}$. *Dann ist für jedes* $\mathbf{x} \in \mathbb{C}^n$ *das Elemente* $P(\mathbf{x})$ *dasjenige Element von* $\boldsymbol{M}$, *welches minimalen Abstand von* $\mathbf{x}$ *hat, d.h. für das gilt:*

$$\|\mathbf{x} - P(\mathbf{x})\| \leq \|\mathbf{x} - \mathbf{m}\| \quad \forall \mathbf{m} \in \boldsymbol{M}. \tag{142}$$ mindist1

Als Begründung verwendet man den Satz von Pythagoras, angewendet auf $\mathbf{m} - P(\mathbf{x})$ und $P(\mathbf{x})$ (und "Hypothenuse" $P(\mathbf{x}) - \mathbf{m}$).

projocompl **Corollary 9.** *Ist* P *die Projektion auf einen Teilraum* $\boldsymbol{V} \subseteq \mathbb{C}^n$, *dann ist* $\mathbf{Id}_n - P$ *die Projektion auf* $\boldsymbol{V}^{\perp}$ *(und natürlich umgekehrt).*

Dieses Prinzip ist insbesondere dann nützlich, wenn eine der beiden Projektionen einfach zu bestimmen ist, also niedrig-dimensional ist. Ist beispielsweise die *Codimension* gleich 1, d.h. ist die sog. *Hyperebene* $\boldsymbol{V} \subset \mathbb{C}^n$ durch eine einzige Gleichung gegeben, so ist sie auch in der Form $\boldsymbol{V} = \{\mathbf{v} \,|\, \langle \mathbf{v}, \mathbf{n} \rangle = 0\}$ für einen passend gewählten Richtungsvektor. Die Projektion auf das orthogonal Komplement, dass dann ja einfach aus dem eindim. Raum aller skalaren Vielfachen von $\mathbf{n}$ besteht, also wegen $\boldsymbol{V}^{\perp} = \{\lambda \mathbf{n}, \lambda \in \mathbb{C}\}$ (genauso über $\mathbb{R}$) ist also dann $\mathbf{x} \to \langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n}$, also ist $P_{\boldsymbol{V}} = Id_n - \mathbf{n} * \mathbf{n}'$ ($\mathbf{n}$ als Spaltenvektor geschrieben!).

[91] Je nach Situation ist es einfacher die Projektionsmatrix oder die Projektionsmatrix auf das orthogonale Komplement zu berechnen, denn oft ist einer der Rüme von deutlich kleinerer Dimension, und dann ist die entsprechende Rechnung einfacher bzw. schneller!

[92] im Vergleich dazu: eine Involution J erfüllt $J^2 = Id_{\boldsymbol{V}}$.

Wie man vorgeht, wenn die Erzeuger (idealerweise schon eine Basis) des Raumes nicht orthogonal (und dann am besten auch gleich normiert) sind, kann mit Hilfe der Grammatrix (und deren Inversion) und des Begriffes der Biorthogonal-Basis realisiert werden.

Kleine Zwischenzusammenfassung

Hat man eine linear unabhängige Menge von Vektoren im $\mathbb{R}^m$ (oder gleichbedeutend: eine $m \times n$-Matrix, mit $n \leq m$ mit reellen [analog komplexen] Eintragungen von maximalem Rang), dann kann man zur Lösung des Standardproblems der linearen Algebra, nämlich der Gleichung $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ dadurch kommen, indem man die ganze Gleichung zu einem (dann etwas kleineren) Gleichungssystem $\boldsymbol{A}'(\boldsymbol{A}\mathbf{x}) = \boldsymbol{A}'\mathbf{b}$ macht, und nach $\mathbf{x}$ auflöst (was ja wegen der Invertierbarkeit der Gram-Matrix geht!), um zu folgender Lösung zu kommen:

$$\mathbf{x} = \mathbf{G}^{-1} * (\boldsymbol{A}' * \mathbf{b}) =: \mathrm{pinv}(\boldsymbol{A}) * \mathbf{b}, \tag{143}$$ Normalform

mit der Fixierung

$$\mathrm{pinv}(\boldsymbol{A}) = \boldsymbol{A}^+ := (\boldsymbol{A}' * \boldsymbol{A})^{-1} * \boldsymbol{A}' \tag{144}$$ pinvchar1

oder alternativ, mit

$$\boldsymbol{H} := \mathrm{pinv}(\boldsymbol{A})' = \boldsymbol{A} * (\boldsymbol{A}' * \boldsymbol{A})^{-1}. \tag{145}$$ biorthog

Dann gelten einige interessante Identitäten, wie z.B.

$$\boldsymbol{A}^+ * \boldsymbol{A} * \boldsymbol{A}^+ = \boldsymbol{A}, \quad \boldsymbol{A}^+ * \boldsymbol{A} * \boldsymbol{A}^+ = \boldsymbol{A}^+ \tag{146}$$ pseudinv0

und insbesondere gilt dann für $\mathbf{P} = \boldsymbol{A} * \boldsymbol{A}^+$

$$\mathbf{P} * \mathbf{P} = \mathbf{P} \quad \text{und } \mathbf{P}' = \mathbf{P}. \tag{147}$$ projpinv1

ACHTUNG: hier stand bis vor kurzem: $\mathbf{P} = \boldsymbol{A} * \boldsymbol{A}'$!

Genaugenommen bekommt man sogar:

projcol1 **Lemma 42.** *Die ist die $n \times n$-Matrix $\mathbf{P} = \boldsymbol{A} * (\boldsymbol{A}' * \boldsymbol{A})^{-1} * \boldsymbol{A}'$ genau die Matrix zur Projektion auf den Spaltenraum von $\boldsymbol{A}$.*

Proof. Wir haben bereits gesehen, dass $\mathbf{P}$ die beiden charakteristischen Eigenschaften einer orthogonalen Projektion im $\mathbb{R}^n$ erfüllt. Andererseits folgt aus $\mathbf{P} * \boldsymbol{A} = \boldsymbol{A}$ dass die Spaltenvektoren von $\boldsymbol{A}$ einzeln invariant bleiben, damit aber auch ihre Linear- Kombinationen, also läßt $\mathbf{P}$ den ganzen Spaltenraum invariant (er liegt im Bild von $\mathbf{P}$).

Umgekehrt gilt $\mathbf{P} * \mathbf{x} = \boldsymbol{A} * \mathbf{y}$ für $\mathbf{y} = ...$, somit liegt das Bild von $\mathbf{P}$ innerhalb des Spaltenraumes von $\boldsymbol{A}$ (es werden ja Lin.Kombs. von Spalten von $\boldsymbol{A}$ gebildet...). Damit ist der Beweis fertig. □

BEMERKUNG: Sollten die Spalten von $\boldsymbol{A}$ nicht linear unabhängig sein, dann ist die Frage, ob $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$ lösbar, also ob $\mathbf{b}$ im Spaltenraum von $\boldsymbol{A}$ liegt, leicht *ersetzbar* durch eine andere Gleichung, die dadurch entsteht, dass man die linear abhängigen Spalten von $\boldsymbol{A}$ einfach wegläßt. Das ändert nichts am Spaltenraum und somit an der Lösbarkeit des Problems, stellt aber die Invertierbarkeit der Gram-Matrix her. Kurz gesagt, man bestimme die Pivotspalten (der ursprünglichen Matrix) und lasse diese (abhängigen, und somit unnötigen) Variablen weg, und wende dann das obige Prinzip an (später kommen noch elegantere Methoden!).

19 MATHEMATICA und Lineare Algebra

Untertitel: warum ich MATLAB und nicht MATHEMATICA verwende (ohne etwas allgemein über den Wert der beiden math. Softwarepakete zu sagen, jede/r MathematikerIn sollte beide kennen und wenigstens eines davon einigermaßen gut beherrschen, um kleine Rechnungen anzustellen).

Diverse LINKS (es gibt aber gute Gründe, warum MATLAB oder OCTAVE als relevante Software vorgeschlagen wird, selbst wenn es eine alte Version ist, und nicht das mächtigere - vor allem im symbolischen Bereich - aber doch etwas umständlichere Paket MATHEMATICA.

```
Google Suche nach "matrix MATHEMATICA"  oder Aehnliches:

http://www.ma.iup.edu/projects/CalcDEMma/linalg0/linalg0.html#linalg01
http://www.ma.iup.edu/projects/CalcDEMma/linalg0/linalg01.html
http://www.ari.uni-heidelberg.de/MathPhysI/mma-handb.pdf

Matrizen und Vektoren:  p.13
viele geschwungene Klammern, e.g.
m = {{2,1,0},{0,3,01},{0,0,3}}

D = DiagonalMatrix[{x1, x2, x3}]
versus MATLAB:  D = diag(x);    ..hgfei

MATHEMATICA:   Transpose[b]  versus  b'  in MATLAB

LinearSolve[{{1,2},{3,4},},{a,b}]    MATHEMATICA: symbols!!
    MATLAB:  A = [1,2;3,4]; b= [5;6];  x = A\b;

Inverse Matrix:
http://reference.wolfram.com/mathematica/ref/Inverse.html

This is a 22 matrix.
In[1]:= m = {{a, b}, {c, d}}  etc...
Here is the element m_(12).
In[3]:=
Click for copyable input
Out[3]=     schwer kopierbar!!


Nullraum:  Basis des Nullraumes:
    http://reference.wolfram.com/mathematica/ref/NullSpace.html
```

20 !! Charakterisierung von Basen !!

Das folgende Theorem wird noch mehr und mehr erweitert:

maininvmatr **Theorem 22.** *Die folgenden Eigenschaften sind äquivalent:*

1. *"Das Gauss'schen Eliminationsverfahren" liefert*[93] *führt für JEDE rechte Seite* $\mathbf{b} \in \mathbb{R}^n$ *zu einer eindeutigen Lösung von* $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$*;*
2. *Eine* $n \times n$ *Matrix* $\boldsymbol{A}$ *ist invertierbar;*
3. *Die Spalten von* $\boldsymbol{A}$ *sind eine Basis für* $\mathbb{R}^n$*;*
4. *Die* n *Spalten von* $\boldsymbol{A}$ *sind linear unabängig;*
5. *Die* n *Spalten von* $\boldsymbol{A}$ *sind ein Erzeugendensystem für* $\mathbb{R}^n$*;*
6. *Der Spaltenraum von* $\boldsymbol{A}$ *ist ganz* $\mathbb{R}^n$*;*
7. *Der Zeilenraum von* $\boldsymbol{A}$ *ist ganz* $\mathbb{R}^n$*;*
8. *Die Matrix* $\boldsymbol{A}'$ *oder* $\boldsymbol{A}^t$ *haben die entsprechenden Eigenschaften.*

Proof. BEWEIS: unsortierte Liste von Argumenten

- (triviale Beobachtungen) Ist die Matrix $\boldsymbol{A}$ invertierbar, so auch $\boldsymbol{A}^t$, somit sind die von $\boldsymbol{A}$ bzw. $\boldsymbol{A}^t$ induzierten Abbildungen surjektiv, somit gilt sowohl f.d. Spalten von $\boldsymbol{A}$ als auch die von $\boldsymbol{A}^t$ (gemeinhin als "Zeilen von $\boldsymbol{A}$" bezeichnet), dass ihr Erzeugnis mit ganz $\mathbb{R}^n$ übereinstimmen müßen, d.h. für invertierbare Matrizen gilt: Sowohl Spaltenraum als auch Zeilenraum sind der ganze $\mathbb{R}^n$.
- Ist die Matrix $\boldsymbol{A}$ invertierbar, dann gibt es eine (ebenfalls invertierbare) inverse Abbildung S zur Abbildung: $T : \mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$. Diese beiden Abbildungen ergeben $S \circ T = Id_{\mathbb{R}^n}$, daraus ergibt sich, dass die Abbildung T injektiv ist, d.h. die Spalten von $\boldsymbol{A}$ müssen linear unabhängig sein[94]. Andererseits gilt auch $T \circ S = Id_{\mathbb{R}^n}$, also muss T surjektiv sein, und somit sind die Spalten von $\boldsymbol{A}$ ein Erzeugendensystem von $\mathbb{R}^n$. Insgesamt ergibt sich auf diese Weise: Die Spalten von $\boldsymbol{A}$ sind eine Basis des $\mathbb{R}^n$.
- Ist der Zeilenraum ganz $\mathbb{R}^n$, dann kann man durch Lin. Komb. der gegebenen Zeilen alle Einheitsvektoren bilden, also findet man eine Links-Inverse Matrix, die aber ein Produkt von Elementar-Matrizen ist, somit selbst invertiertbar, also ist auch die Matrix selbst invertierbar, und die Links-Inverse ist letztendlich auch eine Rechtsinverse.

[93] Genauer gesagt sollte man es so formulieren: Es gibt jedenfalls eine [und daher viele] endliche Folge von Gauss'schen Eliminationsschritten, die zu einer Lösung führen. Ganz gleich welchen Weg man wählt, kommt man zu der eindeutig bestimmten, Lösung. Noch deutlicher: *Wenn das quadr. Glgsys. für jede rechte Seite lösbar ist, dann ist es auch eindeutig! lösbar.*

[94] Hier benutzen wir die einfache Aussage: Wenn $g \circ f$ eine bijektive Abbildung ist, dann muss f injektiv und g surjektiv sein!

- Ist der Spaltenraum der ganze $\mathbb{R}^n$, dann ist jeder Einheitsvektor $\mathbf{e}_k$ im Spaltenraum, es gibt also einen Vektor $\mathbf{b}_k \in \mathbb{R}^n$ mit $\boldsymbol{A} * \mathbf{b}_k = \mathbf{e}_k$.
- Ist der Zeilenraum ganz $\mathbb{R}^n$, dann ist der Spaltenraum von $\boldsymbol{A}'$ ganz $\mathbb{R}^n$, und natürlich umgekehrt.
- Ist der Spaltenraum von $\boldsymbol{A}$ ganz $\mathbb{R}^n$ (d.h. $\mathbf{Sp}(\boldsymbol{A}) = \mathbb{R}^n$), dann ist der Zeilenraum von $\boldsymbol{A}'$ ganz $\mathbb{R}^n$, und somit ist $\boldsymbol{A}'$ invertierbar! Dann ist aber wegen der Formel ... auch $\boldsymbol{A}$ invertierbar.
- Wenn der Zeilenraum ganz $\mathbb{R}^n$ ist, dann ist gesichert, dass das Gauss-Jordan Verfahren zur Idealgestalt, nämlich zur Einheitsmatrix führt, weil ja jeder Einheitsvektor ein im lineare Erzeugnis der (ursprünglichen) Zeilen liegt;
- Wenn der Spaltenraum von $\boldsymbol{A}$ ganz $\mathbb{R}^n$ ist, dann ist auch der Zeilenraum von $\boldsymbol{A}'$ ganz $\mathbb{R}^n$. Somit ist die Gauss-Elimination auf $\boldsymbol{A}'$ anwendbar, woraus sich die Invertierbarkeit von $\boldsymbol{A}'$ ergibt, durch Transponieren aber letztendlich die Invertierbarkeit von $\boldsymbol{A}$ (siehe (67)). #matrprodtransp
- Wenn der Zeilenraum nicht der ganze $\mathbb{R}^n$ ist, dann ist die Menge der Zeilenvektoren linear abhängig. Das bedeutet aber, dass wenigstens eine nicht-triviale Linear-Kombination der Zeilenvektoren existiert, die nicht null ist (d.h. *obwohl* die Koeffizientenfolge nicht aus Null besteht, ist die Summe Null. Klarerweise kann das nur bedeuten, dass mindestens zwei! Koeffizienten nicht Null sind, denn den trivialen Fall, dass einfach ein Zeilenvektor der Nullvektor ist, brauchen wir nicht genau zu diskutieren). Dann kann man aber durch Zeilen- Manipulationen diese spezielle Zeile zur Nullzeile machen, und durch Permutation der Zeilen zur letzten Zeile des Gleichungssystems machen. Somit ist aber klar, dass es mindestens einen freien Parameter gibt, somit unendlich viele Lösungen des linearen Gleichungssystems, d.h. die Gauss-Elimination führt eben *nicht* zu einer *eindeutigen* Lösung.[95]
- Wenn $\boldsymbol{A}$ eine Links-Inverse $\boldsymbol{B}$ hat, d.h. wenn es eine Matrix $\boldsymbol{B}$ gibt, mit $\boldsymbol{B} * \boldsymbol{A} = \mathrm{Id}_n$, dann ist $\boldsymbol{A}$ invertierbar, und $\boldsymbol{B} = \boldsymbol{A}^{-1}$, DENN: Die Komposition $\boldsymbol{B} * \boldsymbol{A}$ definiert eindeutig eine injektive Abbildung, also gilt dasselbe für die zuerst angewendete Abbildung, die Matrix-Vektor-Multiplikation mit $\boldsymbol{A}$. Das bedeutet aber, dass die Spalten von $\boldsymbol{A}$ *linear unabhängig* sind. Daraus folgt aber, dass es genau n Pivot-Elemente gibt, denn wären es höchstens $n-1$, dann müßte es ja wenigstens eine frei Variable geben, was im Widerspruch zur linearen Unabhängigkeit steht (keine nicht-triv. Lösung des homog. Systems!). Somit hat man aber n Pivotelement und folglich kann man im Zuge der Gauss-Elimination $\boldsymbol{A}$ in die Einheitsmatrix umformen. Das bedeutet aber, dass $\boldsymbol{A}$ invertierbar ist, und aus der Eindeutigkeit der inversen Matrix ergibt sich dann die Behauptung.

□

Es gibt das folgende interessante Korollar zu diesem Theorem:

[95] dieser Teil des Arguments ist sicherlich schon von den übrigen abgedeckt, ist also nur eine Art Zusatzüberlegung, um das Repertoire and Argumentationsmöglichkeiten zu erweitern.

Corollary 10. *Eine lineare Abbildung T von $\mathbb{R}^n$ nach $\mathbb{R}^n$ ist bijektiv (und hat dann natürlich eine inverse, ebenfalls lineare Abbildung) dann und nur dann wenn sie injektiv oder surjetiv ist. Insbesondere folgt in diesem speziellen Fall die Injektivität (und somit die Bijektivität) aus der Surjektivität, und umgekehrt die Surjektivität aus der Injektivität.*

Kurze Zusammenfassung: Was macht diese lange Liste von äquivalenten Eigenschaften möglich: es ist im Prinzip die Beobachtung, dass das Funktionieren des Gauss'schen Algorithmus garantiert ist, *wenn* gilt: $\mathbf{Z}(\boldsymbol{A}) = \mathbb{R}^n$ (der Zeilenraum von $\boldsymbol{A}$ ist der ganze $\mathbb{R}^n$, siehe (4)). Angewendet auf $\boldsymbol{A}'$ bzw. $\boldsymbol{A}^t$ zeigt dieses Argument, dass man aus der *Surjektivität* der Abbildung $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ (d.h. aus der Voraussetzung dass man n Vektoren im $\mathbb{R}^n$ hat[96] folgt schon, dass das Gaussverfahren zum gewünschten Ziel führt. Da die dabei verwendeten *Elementarmatrizen* alle invertierbar sind, ergibt sich aus der Existenz einer invertierbaren Links-Inversen (nun zur Matrix $\boldsymbol{A}'$) die Invertierbarkeit von $\boldsymbol{A}$ selber, aufgrund der einfachen Formel (67) (in Worten: das Produkt transponierter Matrizen ist gleich der Transponierten des Produktes der Matrizen, in umgekehrter Reihenfolge genommen).

[96] Wir werden sehen, bzw. haben gesehen, dass das die Minimalzahl von möglichen Erzeugern für den ganzen $\mathbb{R}^n$ ist!

Das folgende Lemma wird wohl später noch an einen anderen Platz verschoben:

takeone **Lemma 43.** *Eine Menge M ist genau dann linear abhängig, wenn man mindestens ein Element $m_0 \in M$ finden, sodass man dieses wegnehmen kann, ohne das lineare Erzeugnis zu ändern. In Symbolen: Es gibt eine echte Teilmenge $M_0 = M \setminus \{m_0\}$ (für ein passend gewähltes $m_0 \in M$) sodass*

$$\boldsymbol{V}_{M_0} = \boldsymbol{V}_M.$$

Man verwendet oft anstelle von $\boldsymbol{V}_M$ das Symbol $\langle M \rangle$, siehe z.B. Skriptum von Prof. Cap.

Proof. Es ist trivial, dass $\boldsymbol{V}_{M_0} \subset \boldsymbol{V}_M$ gilt, weil es mit mehr Elementen immer möglich ist, mehr Linearkombinationen zu bilden.

Wir müssen also nur zeigen, dass diese Inklusion unter den gegebenen Bedingungen, und bei passender Wahl des zu entfernenden Elements keine echte Inklusion liefert[97].

Gehen wir also von der linearen Abhängigkeit aus (und ignorieren wir den logisch denkbaren Fall, dass einer der Vektoren in M der Null-Vektor waere, den man immer weglassen kann ...).

Sei also eine nicht-triviale Linear-Kombination der Elemente von M, dann gibt es ja (weil die Koeffizientenfolge nicht identisch Null ist), darin einen Koeffizienten, der nicht Null ist (in Wirklichkeit mindestens zwei!!!). Das entsprechende Element kann dann entfernt werden, mit dem folgenden Argument (nehmen wir an, dass wir n Vektoren in M haben):

$$\sum_{k=1}^{n} \mu_k \mathbf{v}_k = \mathbf{0} \in \boldsymbol{V}, \tag{148}$$ lincombzero1

wobei aber nicht alle Koeffizienten μ_k null sind. Folglich gibt es mindestens einen Index $\mu_{k_0} \neq 0$, und das entsprechende Element, also m_{k_0}, wird sich als das entfernbare erweisen, mit folgender Begründung:

Hat man ein beliebiges Element im Linear Erzeugnis aller Elemente von M, als

$$\mathbf{v} = \sum_{k=1}^{n} \lambda_k m_k, \tag{149}$$ reprvv1

dann kann man den Term $\lambda_{k_0} m_{k_0}$ ja durch eine Linearkombination der restlichen Elemente darstellen, denn

$$m_{k_0} = \frac{1}{\mu_{k_0}} \sum_{k \neq k_0} \mu_k m_k, \tag{150}$$ pulloutm0

pulloutm0reprvv1

also hat man durch Einsetzen von (150) in (149)

$$\mathbf{v} = \sum_{k \neq k_0} (\lambda_k - \frac{\mu_k}{\mu_{k_0}}) m_k, \tag{151}$$ newpresvv1

□

[97] Wir geben den formalen Beweis hier für den Fall einer *endlichen Menge* M, der allgemeine Fall erfordert ein wenig mehr Buchhaltung (in der Vorlesung wurde das am 2.12.2010 so gemacht, vielleicht wird das später noch nachgeholt.

AustLemma2 **Lemma 44.** *[Austauschlemma] Hat man eine Basis $\mathscr{B}$ in einem Vektorraum $\boldsymbol{V}$ gegeben, und ist $\mathbf{v}_0 \in \boldsymbol{V}$ ein beliebiges Element. Dann kann man jeden Basis-Vektor $\mathbf{b}_{k_0}$, der zu dem Element $\mathbf{v}$ (in der Standarddarstellung $\mathbf{v} = \sum_{k=1}^{n} c_k \mathbf{b}_k$ muss also $c_{k_0} \neq 0$ gelten) durch $\mathbf{v}$ ersetzen, d.h. die neue, durch Austausch entstandene Familie $\mathscr{C}$ is ebenfalls eine Basis.*

Proof. Im Prinzip ist dies ein Korollar zum oben bewiesenen Lemma, wenn man als Menge $M = \mathscr{C} \cup \{\mathbf{v}_0\}$ nimmt. Dass diese Menge ein Erzeugendensystem für $\boldsymbol{V}$ ist klar. Nach dem Lemma (takeone 43) ist auch nach Entfernung der genannten Basis-Vektoren immer noch eine Erzeugendensystem für denselben Raum. Die Berechnung der Darstellung ist aber unter den genannten Voraussetzungen *eindeutig* (vorher wie nachher).

Das kann auch verbal leicht beschrieben werden. Hat man irgendeinen Vektor $\mathbf{v} \in \boldsymbol{V}$ dann ist seine Darstellung bzgl. $\mathscr{B}$ eindeutig. Der Anteil von $\mathbf{b}_{k_0}$ muss durch $\mathbf{v}_0$ ersetzt werden, was zu eindeutigen Korrekturen in allen anderen Koordinaten bzgl. $\mathscr{B}$ führt. □

Hierher gehört auch noch der Austauschsatz von Steinitz: Unter welchen Bedingungen kann man einen gegebenen Vektor eines Erzeugersystems in einem Raum $\boldsymbol{V}$ durch eine Linear-Kombination ersetzen, ohne das Lineare Erzeugnis, d.h. die Menge der Linear-Kombinationen zu verändern?

steinitz1 **Theorem 23.** *Es sei $\boldsymbol{V}$ ein Vektorraum über $\mathbb{K}$, und $(\mathbf{v}_j)_{j\in J}$ sei eine Basis von $\boldsymbol{V}$. Sind endlich viele, linear unabh. Vektoren in $\boldsymbol{V}$ gegeben, nennen wir sie $\mathbf{w}_1, \ldots, \mathbf{w}_m$, dann gibt es eine Indexfolge $j_1, \ldots, j_m$, sodass auch nach dem Austausch der Elemente $\mathbf{v}_{j_k}$ durch die Elemente $\mathbf{w}_k, 1 \leq k \leq m$ das neue System ebenfalls eine Basis von $\boldsymbol{V}$ ist.*

Insbesondere kann es in der Folge $j_1, \ldots, j_m$ keine Wiederholungen geben, also muß die ursprüngliche Basis mindestens m Elemente haben[98].

Proof. Beweis durch Induktion nach der Anzahl m der auszutauschenden Elemente. Details vorerst noch ausgelassen, obwohl die Aussage als solche sehr wichtig ist. □

In der Praxis ist es durchaus eine interessante (oft auch leicht *durch geometrische Anschauung* lösbare Aufgabe, zu sehen, welche Elemente entbehrlich sind, d.h. *von den anderen linear abhängig* sind, d.h. entfernt werden können, ohne das lineare Erzeugnis zu verändern.

Das im Lemma angesprochene Prinzip ist natürlich auch von praktischem Nutzen, weil es Zeit, dass ein Erzeugendensystem entweder schon eine Basis ist, oder aber lediglich zuviele Elemente enthält, sodass man nur eine (passende) Auswahl von Elementen entfernen muss, um zu einer Basis zu kommen. Konkrete Beispiele dazu sollen im Proseminar behandelt werden.

Man kann diese Aufgabe auch praktisch/numerisch mit Hilfe der Gauss-Elimination realisieren, denn die Vereinigung von zwei Erzeugenden Systemen (Spaltenvektoren) ist natürlich ein Erzeugenden-System, und durch Wahl der Reihenfolge kann man steuern,

[98] Eine andere Art der Formulierung: In einem $n-$dimensionalen Raum, d.h. einem Vektorraum $\boldsymbol{V}$ mit einer [und daher allen] Basis mit n Elementen (Anzahl der Elemente der abstrakten Indexmenge), sind maximal n Elemente linear unabhängig. Noch anders ausgedrückt: Jede Menge von mehr als n Elementen in einem n-dim. Vektorraum ist (notwendigerweise) linear abhängig.

welche der Spaltenvektoren "möglichst früh" (als Pivot-Spalten) vorkommen, und die anderen (linear abhängigen) werden verwendet.

Für diesen Teil kann man folgendes MATLAB Experiment machen, Details fehlen noch. May 29, 2014:

```
>> V =  round(6*rand(5,3));
>> rank(V),  % typically linear independent
>> U =  round(6*rand(5,5));
>> rank(U),  % it is a basis
>> U(:,1) =   4* V(:,1) - 3*V(:,3);
>> U(:,3) =   V(:,2) + 2*V(:,3);
>> rank(U),  % it is still a basis!!!
% hence the linear span of  U and V together is the whole R^5!
>> rref([V,U])
```

20.1 Wie bestimmt man Basen?

1. Wie bestimmt man die Basis des Nullraumes von $\boldsymbol{A}$, d.h. für die allgemeine Lösungsmenge des homogenen Gleichungssystems $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$?
   ANTWORT: Man wende die Gauss-Elimination an, um zu sehen, wieviele frei Parameter es gibt. Fasst man die Vielfachheiten dieses Parameters in den einzelnen Komponenten zu jeweils einem Vektor zusammen, ergibt sich ein linear unabh. Erzeugendensystem[99]

2. Wie bestimmt man eine Basis für den Zeilenraum von $\boldsymbol{A}$?
   ANTWORT: Man wendet die Gauss-Elimination an, um festzustellen, wieviele Pivot-Elemente, d.h. (o.B.d.A. führende) Einsen es gibt.

   Daraus ergibt sich schon eine einfache Dimensionsformel (genauere Interpretation später!): Hat man m Gleichungen in n Variablen, und r Pivot-Elemente, dann gilt: $r = \dim(\mathbf{Z}(\boldsymbol{A}))$ und

   $$\dim(\mathbf{Z}(\boldsymbol{A})) + \dim(Kern(\boldsymbol{A})) = n \quad \text{bzw.} \quad \text{rank}(\boldsymbol{A})) + \dim(\text{Null}(\boldsymbol{A})) = n \quad (152)$$ dimFormel

   d.h. die Anzahl der freien Variablen ist die Summe aus *Nullität* von $\boldsymbol{A}$ (d.i. die dimension des Nullraumes) plus der Dimension des Bildraumes $\dim(\mathbf{Sp}(A))$ (der ja die gleiche Dimension hat wie der Zeilenraum, nämlich die Anzahl der Pivot-Elemente!).

3. Wie bestimmt man die Basis des Spaltenraumes einer Matrix $\boldsymbol{A}$?
   ANTWORT: Man kann die Basis bestimmen, indem man die Gauss-Elimination an der transponierten Matrix (und transponiert zum Schluss das End-Ergebnis wieder in die Ausgangslage zurück) durchführt. Wenn man sowohl die Basis für $\mathbf{Z}(\boldsymbol{A})$ *als auch* $\mathbf{Sp}(\boldsymbol{A})$ benötigt, kann man sich diese doppelte Arbeit sparen: Es genügt, die Spaltennummer zu notieren, in denen sich die Pivot-Element befinden.

[99] Unabhängigkeit ist hier leicht zu verstehen, denn die Parameterwerte λ, μ etc. können tatsächlich unabhängig voneinander variieren.

Geht man dann zu den *entsprechenden Spalten* von $\boldsymbol{A}$, dann hat man eine Basis für $\mathbf{Sp}(\boldsymbol{A})$. Die Begründung dafür folgt separat.

21 APPENDIX: diverse MATLAB files

21.1 Einfache Routinen, zum Lernen und Benützen

```
% kreuz.m
% berechnet das Kreuzprodukt von 2 Vektoren im Raum.
% a,b...Vektoren
% nv....Normalvektor
%
% (c) in 1997 by Andreas M. Wenth
% Zur Diplomarbeit:
% "Bearbeitung und Visualisierung ausgewdhlter Kapitel der linearen
% analytischen Geometrie mit dem Softwarepaket MATLAB"
%
% Aufruf: [nv] = kreuz(a,b); oder
%          nv  = kreuz(P,Q,R); mit 3 gegebenen Punkten

function [nv] = kreuz(a,b);

if nargin == 3;   a = P - Q;  b = R - Q; end;
a = a(:), b = b(:); % um Zeilen oder Spalten-input zu erlauben
% Berechnung des Normalvektors:
nv(1,1)=  a(2,1)*b(3,1)-a(3,1)*b(2,1);
nv(2,1)=-(a(1,1)*b(3,1)-a(3,1)*b(1,1));
nv(3,1)=  a(1,1)*b(2,1)-a(2,1)*b(1,1);
```

21.2 Solve quadratic equation

```
% SOLVQU.M   Sept. 99, HGFei, Rev. May 2014
%
% solve quadratic equation:  ax^2+bx+c=0
%                   or  x^2 + p*x + q = 0;
%
% Usage:    [x1,x2] = solvqu(a,b,c);
%  resp:    [x1,x2] = solvqu(p,q);
%    or:    [x1,x2] = solvqu([c1,c2,c3]);
%    or:    [x1,x2] = solvqu(cof);   % cof in R^3

function   [x1,x2] = solvqu(a,b,c);

if nargin == 1;  % input is a coefficient vector
     if length(a) == 3;
```

```
     c = a(3); b = a(2); a = a(1);
     [x1,x2] = solvqu(a,b,c);    %   return;
     else; disp('format not suitable');
end; end;

if nargin == 2;   p = a;  q = b;
discr =  ((p^2/4) - q);
  if discr < 0;
disp(' the quadratic equation has only complex roots');
  end;
u = sqrt(discr);
x1 = - p/2 + u;
x2 = - p/2 - u;
end;

if nargin == 3;
  discr =  b^2 - 4*a*c;
  if discr < 0;
  disp(' the quadratic equation has only complex roots');
  end;
u = sqrt(discr);
x1 = (-b + u)/(2*a);
x2 = (-b - u)/(2*a);
end;
```

21.3 Winkelbestimmung im allgemeinen Dreieck

Aus heutiger Sicht (2014) wäre es besser, mit Vektorrechnung, d.h. mit Skalarprodukten zu arbeiten. Dann könnte man genauso Dreiecke im $\mathbb{R}^n$ nehmen, z.B. $n = 3$. So wurde es "damals" gelöst:

```
% cossatz.m  HGFei  ,  Sept. 99, from  sss.m (Wenth)
% Trigonometrie im schiefwinkeligen Dreieck, als Angabe dienen
% die drei Seiten des Dreiecks
%
% (c) in 1997 by Andreas M. Wenth, Zur Diplomarbeit:
% "Bearbeitung und Visualisierung ausgewaehlter Kapitel
% der linearen analytischen Geometrie mit Softwarepaket MATLAB"
%
% Aufruf: [winkel1] = cossatz(Seite1,Seite2,Seite3);

function [winkel1] = sss(Seite1,Seite2,Seite3);

% Berechnung von Winkel1 mit Hilfe des Cosinussatzes

winkel1 = acos((Seite1^2-Seite2^2-Seite3^2)/((-2)*Seite2*Seite3));
```

```
% Winkel2 = asin(sin(Winkel1)*Seite2/Seite1);
Winkel1 = winkel1*180/pi;
x=['Winkel1 = ',num2str(Winkel1),'in Grad, output radians '];
disp(x);
```

Ein Bespiel einer Winkelumwandlung + Demo

```
% WACOS.M  A(rcus)COS in terms of natural angels
% output in degrees (2 * pi = 360 )! (not radians)
%
% Usage:  wt = wacos(alpha);

function   wt =  wacos(alpha);

if nargin == 0;
    a = sqrt(2), b = 1/a;  wt = wacos(b),
end;

wt = acos(alpha) * 180/pi;
```

21.4 Matrix zum Translationsoperator in $\mathscr{P}_n(\mathbb{R})$

```
% TRANSMAT.M    HGFei  , Nov. 2010
%
% Usage:  H = transmat(degree);

function H = transmat(degree);

order = degree + 1;
H = zeros();
for jj=1:order; H(jj,jj:order) = binom(order - jj); end;
% change signs appropriately
% trivally done via:
H = inv(H);
H = H';
```

21.5 Matrix zum Dilation-Operator in $\mathscr{P}_n(\mathbb{R})$

Dilation oder Streckung ist ebenfalls eine lineare Operation of $\mathscr{P}_n(\mathbb{R})$: $T : p(t) \to p(\alpha \cdot t)$ mit entsprechendem MATLAB code (siehe Formel 37).

```
% DILMAT.M  hgfei  May 2014
%
% Usage:   DM  = dilmat(c,a);
```

```
%      or      DM  = dilmat(ord,a);

function  DM =  dilmat(deg,alph);

% note: symbol used as input does not have to match

if nargin == 0;
    disp('doing a demo');
    a = 2, c = rand(5,1),  DM = dilmat(c,a),
    x = rand(1,4), disp('  polyval(c,a*x)  ');
    polyval(c,a*x), disp(' polyval(DM*c,x)  ');
    polyval(DM*c,x),
    return;
end;

if nargin < 2;  alph = 2; end;
if length(deg) > 1; deg = length(deg) -1; end;

DM =  diag( alph.^(deg :-1 : 0));
```

21.5.1 Verschieben und Differenzieren vertauschen

Betrachten wir die Ableitung einer Funktion, d.h. den linearen Operator $p(t) \mapsto p'(t)$ so ist das eine Abbildung auf jedem der Polynomräume $\mathscr{P}_k(\mathbb{R})$ (das Ergebnis liegt sogar in $\mathscr{P}_{k-1}(\mathbb{R})$, aber das soll uns jetzt nicht interessieren). Man kann also eine Polynomfunktion ableiten und dann die Ableitung verschieben, oder zuerst die Polynomfunktion dem Verschiebungsoperator T_z unterwerfen und nachher ableiten. Begrifflich ist nicht sofort klar, dass diese beiden linearen Operatoren miteinander vertauschen (damit alle auf $\mathscr{P}_k(\mathbb{R})$ definiert sind, betrachten wir sie als Teil der $(k+1) \times (k+1)$-Matrizen über $\mathbb{R}$). Andererseits ist aus der Deutung der Ableitung durch die Tangente klar, dass die verschobene Funktion als Tangente einfach die "gleichartig verschobene" Tangente hat. Da beide lineare Operatoren durch entsprechende $(k+1) \times (k+1)$-Matrizen dargestellt werden können, ist klar, dass diese Matrizen (bei feste gewälter, aber im Prinzip beliebiger Basis) miteinander vertauschen müssen.

Für den Fall $\boldsymbol{V} = \mathscr{P}_2(\mathbb{R})$ kann man den Shift-Operator (Rechts-Shift: $T_1 : p(t) \mapsto p(t-1)$) in der Monomialbasis $\mathscr{M}^2 = \{t^2, t^1, t^0\}$ durch die Matrix

$$[T_1]_{\mathscr{M}^2} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \tag{153}$$

und für den Differentiationsoperator haben wir in derselben Basis

$$[\mathrm{Diff}]_{\mathscr{M}^2} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{154}$$

diffP2

Das Produkt (in beiden Reihenfolgen) ist in der Tat $\mathbf{H} = [\mathrm{Diff} \circ T_1]_{\mathscr{M}^2} = [T_1 \circ \mathrm{Diff}]_{\mathscr{M}^2}$:

$$\mathbf{H} = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ -2 & 1 & 0 \end{pmatrix} \tag{155}$$

unabhängig von der Reihenfolge. Will man also die Koeffizienten von $p'(t-1)$ aus den Koeffizienten $[a;b;c]$ von $p(t) = at^2 + bt + c$ berechnen, so ergibt sich das durch Anwendung er Matrix $\mathbf{H}$ auf $[a;b;c]$, also $[0;2a;b-2a]$, bzw. (was man durch Anwenden der Differentiationsregeln ohnehin erwarten wird): $p'(t-1) = 2at + (b-2a)$. Man kann das Vertauschen der Operatoren durch die Analysis oder aber durch das Vertauschen der beiden Matrizen begründen (dann gilt es hier für $p(t) \in \mathscr{P}_2(\mathbb{R})$).

Was kann man über das Vertauschen des Shift/Verschiebungs-Operators T_z mit dem Streckungs/Dehnungs-Operator sagen?

21.6 MATLAB plotting example

```
%  POLYPLOT.M  hgfei   3.5.2014
%
%  Usage:    polyplot(cof,a,b);
%    or:     polyplot(cof);    % with default [0,1]
%
% plotting polynomial with coefficient vector "cof"  over [a,b]

function   polyplot(cof,a,b);
if nargin == 0; cof = rand(1,4) - 0.5; end;
if  nargin < 2;   a = 0; b = 1;  end
bas = linspace(a,b,512);
plot(bas,polyval(cof,bas),'r');
% private refinement:
grid;
ax20;  % private adjustment
shg;    % show graphics

% DEMO
% hold on;  for jj = 1 : 100 ;
% polyplot(rand(1,4)-0.5); end;  hold off;
```

21.7 Gauss Elimation: Typical Steps

```
%  ADDROW.M   HGFei
%
%  Usage:  B = addrow(B, nrchg, nrop, lam);

function   B = addrow(B, nrchg, nrop, lam);
```

```
if nargin < 4;   lam = 1, end ;  % default value
% eigentlicher Schritt:
B(nrchg,:) = B(nrchg,:) +  lam*B(nrop,:) ;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ROWMULT.M     - Rowwise multiplication of a matrix.
%
% Input : M = a matrix
%                 v = a vector
%              lam = multiplication factor (if only one!)
%
% Output : MN =  matrix
%
% Usage : MN = rowmult(M,v);  % different input formats!
%            = rowmult(v,M);
% or         = rowmult(M, rownr, lam);
%
% See also  : COLMULT

% HGFei     , 5.5.1991, May 2014

function  MN = rowmult(M,v,lam);

if nargin == 0;
 help rowmult;
 return;
end;
[m n] = size(M);
[m1 n1] = size(v);

if nargin == 3;
    ndx = v;
    v = ones(1,m);  v(ndx) = lam;
    [m1 n1] = size(v);
end;

 % exchange if format is wrong
if (m1~=1) & (n1~=1)
 aux = M;  M = v;  v = aux;  % changing order
end;
[m n] = size(M);  l = length(v);

if l ~= m  % warning
 display ('The length of the vector should be equal to
          the number of rows of the matrix');
 return;
```

```
end;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A = zeros(m,n);
for j=1:m
 A(j,:) = M(j,:).*v(j);   % MAIN STEP!
end;
MN = A;


% PIVOTRECON.M   HGFei   May 29th, 2014
%
% Reconstruction of the matrix A from its Pivot-Columns
% and the information coming out of  the command
% [RA, pivnd] = rref(A);
%
%  Usage:    AREC = pivotrecon(PIVA,RA,pivnd);

function  AREC = pivotrecon(PIVA,RA,pivnd);

if nargin == 0;
    A= rand(5,7);  A(:,3) = A(:,1) -  2*A(:,2);
    A(:,5) =  3*A(:,2) - A(:,4);
    [RA,pivnd] = rref(A);
    PIVA = A(:,pivnd);
end;

[m,n] = size(RA);
r = length(pivnd);
nonpiv = 1:n;  nonpiv(pivnd) = [];
% PIVA =  A(:,pivnd);
AREC =   zeros(m,n);
AREC(:,pivnd) = PIVA;
AREC(:,nonpiv) =  PIVA*RA(1:r,nonpiv);
if nargin == 0;
norm(AREC-A);
% quotient of norms: norm(x)/norm(y) = 1
% difference of normalized versions  = 1.2413e-16
end;
```

21.8 more

Literatur

an98 [1] H. Anton. *Lineare Algebra. Einführung, Grundlagen, Übungen. Aus dem Amerikanischen von Anke Walz. (Linear algebra. Introduction, foundations, exercises. Transl. from the American by Anke Walz).* Spektrum Akademischer Verlag, Heidelberg, 1998.

fi08 [2] G. Fischer. *Linear algebra. An introduction for beginners. (Lineare Algebra. Eine Einführung für Studienanfänger.) 16th revised and enlarged ed.* Vieweg Studium: Grundkurs Mathematik. Wiesbaden: Vieweg. xxii, 384 p. EUR 19.90, 2008.

fu12 [3] P. Fuhrmann. *A Polynomial Approach to Linear algebra 2nd Ed.* Universitext. New York, NY: Springer. xvi, 2012.

mu06 [4] H. J. Muthsam. *Lineare Algebra und ihre Anwendungen.* Spektrum Akademischer Verlag, Heidelberg, 2006.

olsh06 [5] P. Olver and C. Shakiban. *Applied Linear Algebra.* Pearson Education Inc., 2006.

scst09 [6] H. Schichl and R. Steinbauer. *Introduction to Working Mathematically (Einführung in das Mathematische Arbeiten).* Springer Berlin, 2009.

frinsp00 [7] L. Spence, A. Insel, and S. Friedberg. *Elementary Linear Algebra: A Matrix Approach.* Prentice Hall, 2000.

st03-8 [8] G. Strang. *Lineare Algebra.* Springer-Lehrbuch. Springer, Berlin, 2003.


```
Vorlesungsmanuskript zu
Lineare Algebra I
Werner Balser
Institut fuer Angewandte Analysis
Wintersemester 2007/08

http://cantor.mathematik.uni-ulm.de/m5/balser/Skripten/LA1.pdf

p.14: Austauschsatz von Steinitz, basierend auf dem AUSTAUSCHLEMMA 1.7.1 : p.13
```

An dieser Stelle werden im Laufe der Weihnachtsferien typische Prüfungsfragen f.d. Kolloqium aufgelistet.

Ebenso ist beabsichtigt, Querverbindungen zum Buch noch deutlicher zu machen!

22 Linearkombinationen von Linearkombinationen

Die Kurzfassung diese Abschnittes ist die folgende Aussage: *(Neue) Linearkombinationen von (alten) Linearkombinationen sind wieder als (ganz neue) Linearkombinationen der ursprünglichen Elemente darstellbar, und zwar in einer Weise, die gut mit der üblichen Matrixmultiplikation verträglich ist.* Die formalen Änderungen sind gering[100], man muß nur in dem Beweis von (71) matr-assoc1 die entsprechenden Änderungen vornehmen (siehe unten).

Proof. Die Hauptproblem für den formalen (allgemeinen) Beweis ist eine vernünftige Wahl der Indizierung. Es ist naheliegend (damit auch die Summierungen über passende Index-Symbole erfolgen kann) die Eintragung der Matrizen mit $(a_{p,q}), (b_{q,r})$ bzw. $(c_{r,s})$ zu bezeichnen, woraus sich in natürlicher Weise die Index-Bezeichnungen $(e_{p,r}), (f_{q,s})$ bzw. $(u_{p,s})$ und $(v_{p,s})$ ergeben.

In formalisierter Form (und um die Analogie zum Beweis der Assoziativät der Matrixmultiplikation, Formel (71) matr-assoc1 herzustellen), wollen wir die Koeffizienten, die benutzt werden um aus einer Familie $\mathbf{a}_1, \ldots, \mathbf{a}_n$ eine passende (erste) Linearkombination zu bilden mit $\boldsymbol{B}$ bezeichnen. Um einsetzbar zu sein, muss $\boldsymbol{B}$ vom Format $n \times k$ sein, für ein k (die Zahl der neu gebildeten Linear-Kombinationen), jede der k Spalten von $\boldsymbol{B}$ enthält die n Koeffizienten, die genommen werden, um dann "auf einmal" die entsprechenden Korrekturen durchzuführen. $\mathbf{C}$ sei eine weitere Kollektion von Spaltenvektoren, der Höhe k und von Breite l. Wir wollen auch $\mathbf{F} := \boldsymbol{B} * \mathbf{C}$ setzen, das ist also die Behauptung, dass die iterierte Anwendung (zuerst $\boldsymbol{B}$ und dann $\mathbf{C}$) logisch das gleiche Ergebnis haben sollte wie die globale Variante, d.h. aus den urspünglichen Vektoren eine Kollektion von Lin.Kombinationen zu machen, und zwar mit der Matrix $\mathbf{F} := \boldsymbol{B} * \mathbf{C}$, welche jedenfalls das Format $n \times l$ hat.

Gehen wir nun an die technischen Details: Fangen wir mit der Wirkung der Matrix $\boldsymbol{B}$ (rechts angeschrieben) auf die Kollektion der n *ursprünglichen* Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_n$ in $\boldsymbol{V}$, d.h. wir bilden m neue Linearkombinationen, die wir mit $\mathbf{e}_1, \ldots, \mathbf{e}_m$ bezeichnen[101] wollen. Bilden wir danach aus diesen neuen Vektoren mit Hilfe der Koeffizienten einer $k \times l$-matrix $\mathbf{C}$ eine Folge von l Vektoren $\mathbf{u}_1, \ldots \mathbf{u}_l$ in $\boldsymbol{V}$, d.h.

$$\mathbf{e}_r = \sum_{q=1}^{n} b_{q,r}\mathbf{a}_q \quad , \quad \mathbf{u}_s = \sum_{r=1}^{k} c_{r,s}\mathbf{e}_r. \tag{156}$$ lincombB

Setzt man die Darstellung von $\mathbf{e}_r$ in die rechte Gleichung ein, bekommt man

$$\mathbf{u}_s = \sum_{r=1}^{k} c_{r,s}\mathbf{e}_r = \sum_{r=1}^{k} c_{r,s}\left(\sum_{q=1}^{n} b_{q,r}\mathbf{a}_q\right) = \sum_{q=1}^{n}\left(\sum_{r=1}^{k} c_{r,s}b_{q,r}\right)\mathbf{a}_q, \tag{157}$$ lincombit

einfach durch Vertauschen der Summationsreihenfolge. Eleganter ausgedrückt bekommt man eine Darstellung der Vektoren $\mathbf{u}_1, \ldots, \mathbf{u}_l$ mit Hilfe der Koeffizienten einer Matrix

[100] Trotzdem heikel und gut zu studieren!

[101] Wir halten uns an die Symbole im Beweis des Assoziativgesetzes, es sind also nicht die Einheitsvektoren damit gemeint!

$\mathbf{F}$, deren Eintragungen genau durch die bereits bekannte Matrix-Multiplikation von $\boldsymbol{B}$ mit $\mathbf{C}$ entsteht, also

$$\mathbf{u}_p = \sum_{q=1}^{n} f_{p,s}\mathbf{a}_p \quad , \quad f_{q,s} = \sum_{r=1}^{k} b_{q,r}c_{r,s}. \tag{158}$$ lincombmmult

Die einzelnen Spalten von F, also die Eintragungen $f_{q,s}$ für festes s sind natürlich genau die durch die Anwendung der gewöhnlichen Matrix-Multiplikation mit $\boldsymbol{B}$, also durch $\vec{\mathbf{y}} := \boldsymbol{B} * \vec{\mathbf{x}}$, also $y_q = \sum_{r=1}^{k} b_{q,r}x_r$ gegeben.

Wie man leicht sieht, ist reduziert sich diese Argumentation auf die in der Herleitung des Assoziativgesetzes gemachte Überlegung, wenn man als Vektorraum $\boldsymbol{V} = \mathbb{R}^m$ nimmt, und die Vektoren $\mathbf{a}_p$ koordinatenweise durch ihre (Spalten)-Koordinaten $(a_{p,q})$ beschreibt.

□

Wir haben also bewiesen:

Corollary 11. *Hat man eine Basis $\mathscr{B} = \{\mathbf{b}_1, \dots \mathbf{b}_n\} \subset \boldsymbol{V}$, und bildet man daraus eine neue Kollektion von Vektoren, indem man ebenfalls n Linear-Kombinationen bildet*[102]*. Dann ist die neue Familie eine Basis genau dann wenn die zugehörige Koeffizienten-Matrix $\mathbf{C}$ invertierbar ist.*

Wir werden das Argument in der Form einer "eleganten Schreibweise" verpacken. Diese Schreibweise weicht ein wenig von der üblichen ab, ist also "sehr abstrakt", aber auch sehr "kompakt" und in gewissem Sinne sehr einfach[103].

Auch um zu demonstrieren, wie sehr die Wahl von Symbolen (in diesem Fall eine selbstgestrickte), versuchen wir das Bewiesene durch eine besondere symbolische Schreibweise sichtbar zu machen[104]:

Für eine Kollektion von Vektoren, wie etwa $\mathscr{B} = \{\mathbf{b}_1, \dots \mathbf{b}_n\} \subset \boldsymbol{V}$, und eine Matrix $\mathbf{C}$ passender Größe (d.h. eine $\mathbf{n} \times k$ Matrix) $\mathbf{C}$ definieren wir den **Linear-Kombinations Operator** (mit dem Symbol "$\bullet$"):

$$\mathscr{B} \bullet \mathbf{C} := \{\mathbf{f}_1, \dots, \mathbf{f}_k\}, \quad \text{mit} \quad \mathbf{f}_q = \sum_{k=1}^{n} c_{q,k}\mathbf{b}_k, \ 1 \le q \le k. \tag{159}$$ setmatrix

Dann haben wir in kompakter Form soeben bewiesen (mit $*$ für Matrix-Mult.):

$$\mathscr{B} \bullet (\mathbf{C} * \boldsymbol{E}) = (\mathscr{B} \bullet \mathbf{C}) \bullet \boldsymbol{E}. \tag{160}$$ bullastmatr

[102] Da jeder der neuen Vektoren eine Linearkombination von n Vektoren ist, und wir n Linearkombinationen bilden wollen, ist es klar, dass man sämtliche benötigten Koeffizienten in der Form einer quadratischen Matrix $\mathbf{C} \in \mathcal{M}_{n,n}$ schreiben kann.

[103] Erinnern wir uns daran, das die mathematischen Symbole viel mit einer Art Kurzschrift zu tun hat, so wie sich untern Musikern die Notenschrift entwickelt hat, wobei man auch bedenken sollte, dass es in anderen Traditionen andere als bei bei uns üblichen [klassichen] Tonarten und Notenschriften gibt.

[104] So wie andere Leute die Notenschrift verwenden, um musikalische Sachverhalte zu beschreiben.

Proof. Wir können nun den Beweis ganz einfach formulieren: Wenn $\mathscr{B}$ eine Basis ist und $\mathscr{A} = \mathscr{B} \bullet \mathbf{C}$ durch eine invertierbare Matrix erstellt wird, dann ist auch die neue Familie $\mathscr{A}$ eine Basis. Denn ist $\mathbf{v} \in \boldsymbol{V}$ ein beliebiger Vektor, so hat er eine eindeutige Darstellung, was wir in der Form $\mathbf{v} = \mathscr{B} \bullet \mathbf{x}$ schreiben können. Es ist naheliegend, das in die Form

$$\mathbf{v} = \mathscr{B} \bullet (\mathbf{C} * \mathbf{C}^{-1} * \mathbf{x}) = (\mathscr{B} \bullet \mathbf{C}) \bullet (\mathbf{C}^{-1} * \mathbf{x}) = \mathscr{A} \bullet \mathbf{y}$$

äquivalent umzuschreiben, mit $\mathbf{y} := \mathbf{C}^{-1} * \mathbf{x}$. Diese Zuordnung ist auch eindeutig, und liefert somit sogar ein Rezept: Die (eindeutigen) Koeffizienten von $\mathbf{v}$ bzgl. $\mathscr{B} \bullet \mathbf{C}$ sind die Koeffizienten von $\mathbf{v}$ bzgl. $\mathscr{B}$, unter der inversen Matrix, d.h. $\mathbf{C}^{-1} * \mathbf{x}$.

Umgekehrt, habe eine Matrix eine feste (man sieht insgesamt, dass es dann für eine beliebige Basis dieselbe Eigenschaft hat) Basis $\mathscr{B}$ durch die Bildung von $\mathscr{A} = \mathscr{B} \bullet \mathbf{C}$ eine neue Basis entsteht. Wäre $\mathbf{C}$ nicht invertierbar, gäbe es eine lineare Beziehung zwischen den Spalten von $\mathbf{C}$, d.h. es gäbe eine nichttriviale Lösung des linearen Gleichungssystems $\mathbf{C} * \vec{\mathbf{y}} = \vec{\mathbf{0}}$. Dann sind aber auch die neugebildeten Vektoren von $\mathscr{A}$ in $\boldsymbol{V}$ linear abhängig, denn klarerweise gilt

$$\mathscr{A} \bullet \mathbf{y} = \mathscr{B} \bullet (\mathbf{C} * \vec{\mathbf{y}}) = \mathscr{B} \bullet \vec{\mathbf{0}} = \mathbf{0} \in \boldsymbol{V}.$$

□

Anmerkung: Die Verwendung dieser Symbole ist nicht Standard im Bereich der allgemeinen Literatur und wird bei schriftlichen Prüfungen nicht verlangt, obwohl es sehr empfohlen wird, sich damit auseinanderzusetzen, denn er beschreibt wichtige Aussagen der linearen Algebra in sehr kompakter Form.

Weitere Anmerkung: Das fortgesetzte Abbilden ist natürlich nicht nur Linearkombinationen mit einer festen Zahl von Vektoren möglich, genausogut kann man auch rechteckige Matrizen bilden, um immer wieder hintereinander neue Linear-Kombinationen zu bilden.

Ein mögliche verbale Zusammenfassung der obigen (in Formeln gepackten) Aussage könnte folgendermassen lauten: Bildet man aus einer Kollektion $\mathscr{B}$ von Vektoren eine Kollektion von Linear-Kombinationen (d.h. bildet man mit Hilfe der Koeffizienten aus der Matrix $\mathbf{C}$) und danach eine weitere Linearkombination (dieser neuen Familie von Vektoren, die wir $\mathscr{B} \bullet \mathbf{C}$ genannt haben), nun mit einer Matrix $\boldsymbol{E}$ (von passendem Format), so lässt sich diese *Linearkombination von Linearkombinationen* auch durch Umsummieren als Kollektion von (einfachen) Linear-Komb. der *ursprünglichen* Vektoren aus $\mathscr{B}$ darstellen, wobei die entsprechenden Koeffizienten einfach der Produktmatrix $\mathbf{C} * \boldsymbol{E}$ zu entnehmen sind.

Man könnte beispielsweis aus n Vektoren $\mathscr{B} = \{\mathbf{b}_1, \ldots \mathbf{b}_n\} \subset \boldsymbol{V}$ eine Anzahl von $k < n$ Linearkombinationen bilden, aus denen man wieder n neue Linear-Kombinationen bilden kann. In einer derartigen Situation könnte man dann ganz allgemein zeigen, dass diese n (ganz) neuen Vektoren einen Teilraum von einer Dimension $\leq k$ aufspannen können. In der Tat, jede Kollektion von mehr als k Vektoren müßte als Bild einer entsprechenden Anzahl von Vektoren in einem $k-$dim. Raum entstehen, ist also notwendigerweise linear abhängig.

rankcontrol **Lemma 45.** *Es sei $T : \boldsymbol{V} \to \boldsymbol{W}$ eine lineare Abbildung. Dann ist $T(\boldsymbol{V})$ ein Teilraum von $\boldsymbol{W}$, für dessen Dimension gilt:*

$$\dim(T(\boldsymbol{V})) \leq \dim(\boldsymbol{V}).$$

Gleichheit gilt genau dann wenn T trivialen Kern hat (d.h. T ist injektiv).

Proof. Sei $r = \dim(\boldsymbol{V})$. Die Annahme $\dim(T(\boldsymbol{V})) > r$ bedeutet, dass es jedenfalls $r+1$ linear unabhängige Element der Form $T(\mathbf{v}_k)$ gibt, $1 \leq k \leq r+1$. Dann müssen auch die Urbilder linear unabhängig sein, denn wären sie linear abhängig, müssten die Bilder auch linear abhängig sein (T bewahrt Linearkombinationen). Das ist aber nicht möglich, weil $r = \dim(\boldsymbol{V})$ die Maximalzahl von linear unabhängigen Vektoren in $\boldsymbol{V}$ ist. □

SS14:MATRIXDARST.TEX: 31.MAI 2014

23 Wechsel von Koordinaten-Systemen

Vorüberlegung:

changbas **Proposition 2.** *Hat man eine invertierbare $n \times n$-Matrix $\mathbf{C}$ und eine Basis $\mathscr{B}$ für einen Vektorraum $\boldsymbol{V}$, dann ist aus $\mathscr{B}' := \mathscr{B} \bullet \mathbf{C}$ ein Basis. Dabei ist:*

$$\mathbf{b}'_j := \sum_{k=1}^{n} c_{j,k}\mathbf{b}_k \quad 1 \leq j \leq n. \tag{161}$$ baschang1

Proof. Offenbar erzeugt die Relation (161) eine Kollektion von $\mathscr{B}'$ von n Vektoren in $\boldsymbol{V}$, welche durch Linear-Kombination aus den Basis-Vektoren $\mathscr{B}$ gewonnen wurden. Jede Spalte $\vec{\mathbf{c}}_j$ von $\mathbf{C}$ enthält die nötigen Koeffizienten, um das einen Vektor $\mathbf{b}'_j$ zu erzeugen. Nun geht es darum, zu zeigen, dass diese neuen Vektoren linear unabhängig sind, denn dann ist schon gesichert dass sei eine Basis für $\boldsymbol{V}$ bilden.
Angenommen wir finden einen Vektor $\vec{\mathbf{x}} \in \mathbb{K}^n$ mit $\sum_{k=1}^{n} x_k\mathbf{b}'_k = \mathbf{0}$, also

$$\mathscr{B}' \bullet \vec{\mathbf{x}} = (\mathscr{B} \bullet \mathbf{C}) \bullet \vec{\mathbf{x}} = \mathscr{B} \bullet (\mathbf{C} * \vec{\mathbf{x}}) = \mathbf{0}. \tag{162}$$

Da $\mathscr{B}$ eine Basis ist, folgt $\mathbf{C} * \vec{\mathbf{x}} = \vec{\mathbf{0}}$, aber da $\mathbf{C}$ invertierbar ist, folgt daraus $\vec{\mathbf{x}} = \vec{\mathbf{0}}$, d.h. wir haben gezeigt, dass die Familie $\mathscr{B}'$ linear unabhängig ist.
Es wäre auch leicht, direkt zu zeigen, dass $\mathscr{B}'$ ein Erzeugendensystem ist. Denn klarerweise gilt $\mathscr{B} = \mathscr{B}' \bullet \mathbf{C}^{-1}$, also kann jeder Vektor $\mathbf{v}$, der geschrieben werden kann als

$$\mathbf{v} = \mathscr{B} \bullet \vec{\mathbf{y}} = \sum_{k=1}^{n} y_k\mathbf{b}_k$$

auch als

$$\mathbf{v} = (\mathscr{B}' \bullet \mathbf{C}^{-1}) \bullet \vec{\mathbf{y}} = \mathscr{B}' \bullet (\mathbf{C}^{-1} * \vec{\mathbf{y}})$$

geschrieben werden, was auch sofort klar macht:

$$[\mathbf{v}]_{\mathscr{B}'} = \mathbf{C}^{-1} * [\mathbf{v}]_{\mathscr{B}},$$

d.h. die Matrix-Multiplikation mit $\mathbf{C}^{-1}$ macht auch $\mathscr{B}$-Koordinaten die $\mathscr{B}'$-Koordinaten.
Im Sinne er verwendeten "abstrakten Terminologie" haben wir

$$\mathbf{C} = [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B}' \hookleftarrow \mathscr{B}}.$$

□

Zusammenfassung:

$$\mathscr{B}' = \mathscr{B} \bullet \mathbf{C} \quad \Leftrightarrow \quad [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B} \leftarrow \mathscr{B}'} = \mathbf{C}^{-1}, \quad \Leftrightarrow [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B} \bullet \mathbf{C} \leftarrow \mathscr{B}} = \mathbf{C}, \tag{163}$$

baschang5

bzw. gleichwertig

$$[\mathbf{v}]_{\mathscr{B}} = \mathbf{C} * [\mathbf{v}]_{\mathscr{B}'} \ , \qquad [\mathbf{v}]_{\mathscr{B}'} = \mathbf{C}^{-1} * [\mathbf{v}]_{\mathscr{B}} \ , \quad \mathbf{v} \in \boldsymbol{V}. \tag{164}$$

baschang6

24 TEST: Matrix Multiplikation

Ein MATLAB file, das die Matrix-Multiplikation nachbildet! Unterschied zwischen elementweiser (zeilenweiser) und spaltenweiser Vorgehensweise.

```
function  b = matrowmul(A,x,'st');
[m,n] = size(A);
[h,w] = size(x);
if abs(h-n) > 0; disp('format not OK'); return; end;
if nargin == 3;
b = zeros( m, 1);
for ii = 1 : m;
  for jj = 1 : n;
      b(ii) = b(ii) + x(jj)*A(ii,jj);
  end;
end;
else  b = zeros( m, 1);
    for jj = 1 : n ;
    b = b + x(jj)*A(:,jj);
end;
norm(b  - A*x) =  fast null!


A = rand(10000);  x = rand(10000,1);
tic; matrowmul(A,x,'standard'); toc
Elapsed time is 4.393293 seconds.
tic; A*x; toc   % built in method
Elapsed time is 0.168529 seconds.
and with columnwise operation:
tic; matrowmul(A,x); toc  % fast columnwise
Elapsed time is 0.028302 seconds.
```

Zum Thema Assoziativität der Matrix bzw. Matrix-Vektor Multiplikation:

```
>> A = rand(5000); B = rand(5000); x = rand(5000,1);
>> tic; b1 = A*B*x; toc;
Elapsed time is 17.506211 seconds.
>> tic; b2 = A*(B*x);  toc;
```

```
Elapsed time is 0.089481 seconds.
>> norm(b1-b2)
ans =    6.2101e-007
```

25 Basis-Wechsel

25.1 Matrix-Darstellung und Basis-Wechsel: Version SS11

Dieser Teil wird noch im Detail im Fortsetzungssemester behandelt, ist aber relativ zentral innerhalb der linearen Algebra. Jedenfalls sollte klar sein, wie man von den Koordinaten in einem System zu den Koordinaten in einem anderen System übergeht, d.h. wie man die entsprechende lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$ findet, anhand konkreter Beispiele.

Wir haben bisher gesehen, dass bei festgehaltenen Basen $\mathscr{B}_1$ bzw. $\mathscr{B}_2$ in zwei Vektorräumen $\boldsymbol{V}_1$ bzw. $\boldsymbol{V}_2$ jede lineare Abbildung T durch eine $m \times n$-Matrix (Standardannahme: $\boldsymbol{V}_1$ ist n-dimensional und $\boldsymbol{V}_2$ ist m-dimensional) $\boldsymbol{A}$ oder $[T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_1}$ darstellbar. Im Falle $\boldsymbol{V}_1 = \boldsymbol{V} = \boldsymbol{V}_2$ und $\mathscr{B}_1 = \mathscr{B} = \mathscr{B}_2$ schreiben wir $[T]_{\mathscr{B} \leftarrow \mathscr{B}}$. Aber wie verhalten sich z.B. die Darstellungen ein und derselben linearen Abbildung in bezug auf zwei verschiedene Basis für denselben Raum (als einfaches Demonstrationsbeispiel), d.h. was ist der Zusammenhang zwischen $[T]_{\mathscr{B}_1 \leftarrow \mathscr{B}_1}$ und $[T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_2}$

Zu diesem Zweck erinnern wir uns an die charakteristische Eigenschaft dieser Matrizen. Es sei $\mathbf{v} \in \boldsymbol{V}$ ein Vektor, und er habe die Koeffizentenfolge (c_k) bzgl. $\mathscr{B}_1$ und die Koeffizientenfolge (d_k) bzlg. $\mathscr{B}_2$, d.h. $\mathbf{c} = [\mathbf{v}]_{\mathscr{B}_1}$ und $\mathbf{d} = [\mathbf{v}]_{\mathscr{B}_2}$ und folglich $[T\mathbf{v}]_{\mathscr{B}_2} = [T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_2} * \mathbf{d}$ und $[T\mathbf{v}]_{\mathscr{B}_1} = [T]_{\mathscr{B}_1 \leftarrow \mathscr{B}_1} * \mathbf{c}$.

Da jede der beiden Basen $\mathscr{B}_1$ und $\mathscr{B}_2$ einen Isommorphismus zwischen $\boldsymbol{V}$ und dem kanonischen Raum $\mathbb{K}^N$ herstellt, ist auch die Zuordnung zwischen $\mathbf{c}$ und $\mathbf{d}$ ein linearer Isomorphismus (als Zusammensetzung zweier Isomorphismen), und $\mathbf{C}$ sie die zugehörige $n \times n$-Matrix. Dann gilt natuerlich dass die Umkehrabbildung, also $\mathbf{C}^{-1}$ den Übergang von $\mathscr{B}_2$-Koordinaten zu $\mathscr{B}_1$-Koordinaten realisiert. Somit gilt $\mathbf{d} = \mathbf{C} * \mathbf{c}$ und $[T\mathbf{v}]_{\mathscr{B}_1} = \mathbf{C}^{-1} * [T\mathbf{v}]_{\mathscr{B}_2}$, und insgesamt

$$[T\mathbf{v}]_{\mathscr{B}_1} = \mathbf{C}^{-1} * [T\mathbf{v}]_{\mathscr{B}_2} = \mathbf{C}^{-1} * [T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_2} * \mathbf{C} * \mathbf{c}, \tag{165}$$ koord-wechs

was aber für beliebiges $\mathbf{v} \in \boldsymbol{V}$ gilt, somit muss gelten:

$$[T]_{\mathscr{B}_1 \leftarrow \mathscr{B}_1} = \mathbf{C}^{-1} * [T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_2} * \mathbf{C} \tag{166}$$ koordtrans1

oder äquivalent

$$[T]_{\mathscr{B}_2 \leftarrow \mathscr{B}_2} = \mathbf{C} * [T]_{\mathscr{B}_1 \leftarrow \mathscr{B}_1} * \mathbf{C}^{-1}. \tag{167}$$ koordtrans2

Man sagt, dass die beiden Matrizen durch die *Konjugation* mit der *Basis-Wechsel-Matrix* $\mathbf{C}$ (bzw. $\mathbf{C}^{-1}$) auseinander hervorgehen.

Der Begriff der Ähnlichkeit (zwei Matrizen werden ähnlich genannt, wenn sie durch die Konjugation mit einer invertierbaren Matrix auseinander hervorgehen) ist eine Äquivalenz-Begriff.

Vergleiche auch das Skriptum von Prof. Andreas Cap (ca. Seite 55).

Wichtig ist auch noch der Verträglichkeit des Matrizen-Kalküls (Matrix-Multiplikation) mit der Komposition von linearen Abbildungen (mit entsprechenden, passenden Basen).

25.2 Ein konkretes Beispiel

Nehmen wir $\boldsymbol{V} = \mathscr{P}_2(\mathbb{R})$, $\mathbf{v} = p(t)$, das quadratische Polynom $p(t) = 3x^2 + 2x + 1$, d.h. für $\mathscr{B} = \mathscr{M}^2$ haben wir die Koeffizientenfolge $[\mathbf{v}] = [3; 2; 1]$. Verschieben wir alle drei Basis Elemente um eins nach *rechts*, d.h. wenden wir den linearen Isomophismus $T : p(t) \mapsto q(t) := p(t+1)$ an, so ist klar, dass die verschobenen Monome weiterhin eine Basis sind. Weiters ist die inverse lineare Abbildung S natürlich die *Links-Verschiebung*, realisiert durch $S : p(t) \mapsto p(t-1)$.

Wenn wir die neue Basis in den alten (d.h.$\mathscr{M}^2$) Koordinaten beschreiben wollen

In der Basis $\mathscr{M}^2$ hat der lineare Operator nun die Matrix-Darstellung $[T]_{\mathscr{B}\hookleftarrow\mathscr{B}}\mathscr{B}$, so haben wir einfach aufgrund des Binomischen Lehrsatzes die Matrix

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \tag{168}$$

Weil $T(t^2) = (t+1)^2 = 1 \cdot t^2 + 2 \cdot t + 1 \cdot 1$, und somit ist die erste Spalte von $\mathbf{C}$ gerade $[1, 2, 1]$.

In unserer abstrakten Sprache wäre das

$$\mathscr{B}' := \{(t+1)^2, (t+1), (t+1)^0\} = \{t^2, t, 1\} \bullet \mathbf{C} = \mathscr{M}^2 \bullet \mathbf{C} = \mathscr{B} \bullet \mathbf{C}.$$

Andererseits ist entweder (per Analogie, aus $+1$ wird im binomischen Lehrsatz -1) oder durch numerische Inversion folgendes verifiziert

$$\mathbf{C}^{-1} = [S]_{\mathscr{B}\hookleftarrow\mathscr{B}} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \quad \text{also} \quad \mathscr{B} = \mathscr{B}' \bullet \mathbf{C}^{-1}. \tag{169}$$

Man kann das auch leicht feststellen, indem man die Basis-Elemente von $\mathscr{B}$ durch $\mathscr{B}'$ Elemente darstellt und z.B. überprüft, dass gilt

$$t^2 = (t+1)^2 - 2t - 1 = 1 \cdot (t+1)^2 - 2 \cdot (t+1) + 1\cdot, 1$$

usw., also auch konkret

$$3x^2+2x+1 = 3(x+1)^2-4x-2 = 3(x+1)^2-4(x+1)+4-2 = 3\cdot(x+1)^2-4\cdot(x+1)+2\cdot 1.$$

Das bedeutet nun direkt dass $[p(t)] = [3; -4, 2]$ ist, weil das die (eindeutigen) Koeffizienten sind, die man benötigt, um $p(t)$ in der $\mathscr{B}'$-Basis auszudrücken.

Ein kleine Probe ergibt aber natürlich, dass gilt

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} * \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -4 \\ 2 \end{pmatrix}$$

was die konkrete Realisierung von

$$[p(t)]_{\mathscr{B}'} = [\mathbf{Id}_{\boldsymbol{V}}]_{\mathscr{B}\hookleftarrow\mathscr{B}'} * [p(t)]_{\mathscr{M}^2} = [\mathbf{Id}_{\boldsymbol{V}}]^{-1}_{\mathscr{B}'\hookleftarrow\mathscr{B}} * [p(t)]_{\mathscr{M}^2}.$$

Man kann aber auch mit unbestimmtem Ansatz aus $a(t+1)^2 + b(t+1) + c = 3t^2 + 2t + 1$ mittels 3×3-Gleichungssystem nach $[a; b; c]$ auflösen. Selber tun zeigt die Analogie!

26 Allgemeine Konstruktionen mit Vektorräumen

26.1 Produkt-Räume

Gegeben zwei Vektorräume $\boldsymbol{V}$ und $\boldsymbol{W}$ über demselben Körper $\mathbb{K}$, so verwenden wir die übliche Schreibweise $\boldsymbol{V} \times \boldsymbol{W}$ für die Menge aller Paare $(\mathbf{v}, \mathbf{w})$, mit $\mathbf{v} \in \boldsymbol{V}, \mathbf{w} \in \boldsymbol{W}$.

prodspace **Lemma 46.** *Mit der naheliegenden "koordinatenweisen Addition (so wie man die Addition in $\mathbb{R}^2$ definiert) ist die Menge ein Vektorraum über $\mathbb{K}$.*

proddim1 **Lemma 47.** *Ist $\mathscr{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_r)$ eine Basis für $\boldsymbol{V}$ und $\mathscr{C} = (\mathbf{c}_1, \ldots, \mathbf{c}_s)$ eine Basis für $\boldsymbol{W}$ dann ist die Vereinigung der Menge $(\mathbf{b}_j, \mathbf{0}_{\boldsymbol{W}}), 1 \leq j \leq r$ mit $(\mathbf{0}_{\boldsymbol{V}}, \mathbf{c}_l), 1 \leq 1 \leq s$ eine Basis für $\boldsymbol{V} \times \boldsymbol{W}$. Insbesondere gilt die Dimensionsformel*

$$\dim(\boldsymbol{V} \times \boldsymbol{W}) = \dim(\boldsymbol{V}) + \dim(\boldsymbol{W}). \tag{170}$$ pordadddim

Neben die "direkten Summe" gibt es auch die "komplexe" Summe (auch Minkowski-Summe genannt), wenn beide Räue in einem großen Raum liegen.

Minksum **Definition 27.** Seien M_1, M_2 zwei Mengen in $\boldsymbol{W}$, dann definiert man

$$M_1 + M_2 := \{m \,|\, m = m_1 + m_2, \quad m_1 \in M_1, m_2 \in M_2\}. \tag{171}$$ sumsets

Es ist leicht zu sehen, dass die Summe von zwei Teilräumen $\boldsymbol{V}_1 + \boldsymbol{V}_2$ in einem Vektorraum $\boldsymbol{W}$ selbst wieder ein Teilraum ist. Man spricht dann von einer *direkten Summe* wenn $\boldsymbol{V}_1 \cap \boldsymbol{V}_2 = \{\mathbf{0}\}$.

dirsum **Theorem 24.** *Genau dann, wenn $\boldsymbol{V} = \boldsymbol{V}_1 + \boldsymbol{V}_2$ eine direkte Summe ist, kann man einen Isomorphismus zwischen $\boldsymbol{V}$ und $\boldsymbol{V}_1 \times \boldsymbol{V}_2$ herstellen.*

Proof. Wenn die Summe direkt ist, dann für jedes $\mathbf{v} \in \boldsymbol{V}$ wenigstens eine Darstellung der Form $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ möglich. Um die Eindeutigkeit zu beweisen, nehmen wir an, wir hätten zwei Darstellungen:

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v} = \mathbf{v}_1' + \mathbf{v}_2'$$

Bringt man die Elemente mit gleichem Index auf eine Seite ist dies äquivalent zu

$$\mathbf{v}_1 - \mathbf{v}_1' \,(\text{ klar in } \boldsymbol{V}_1) = \mathbf{v}_2' - \mathbf{v}_2 \in \boldsymbol{V}_2,$$

also liegt diese Differenz in beiden Räumen, ist also nach Voraussetzung gleich 0, womit die Eindeutigkeit geklärt ist. Daraus folgt aber auch sofort, dass die Zuordnung von $\mathbf{v} \mapsto (\mathbf{v}_1, \mathbf{v}_2)$ linear sein muss, und auch offenbar surjektiv ist! (weil $\mathbf{v}_1 + \mathbf{v}_2 \in \boldsymbol{V}$, für bel. Paare, aufgrund der Definition). Also haben wir einen Isomorphismus.

Umgekehrt (weniger wichtig für uns) kann man argumentieren, dass im Falle, dass $\boldsymbol{V}_1 \cap 3\boldsymbol{V}_2 \neq \{\mathbf{0}\}$ gilt, die (nun nicht mehr direkte Summe) sicherlich von zu kleiner Dimension ist (siehe Lemma (47). proddim1 □

Bemerkung: Die soeben beschriebene Situation ist insbesondere dann erfüllt, wenn die beiden Räume jeweils ein Erzeugendensystem haben, welches linear unabhängig ist. Insbesondere kann man eine ONB nehmen $(\mathbf{b}_k)_{k=1}^n$ und ein zwei disjunkte Teilmengen zerlegen. Dann bilden ihre linearen Erzeugnisse sogar ein paar von komplementären Teilräumen (d.h. zueinander orthogonal liegend, aber insgesamt den Raum aufspannend).

26.2 Quotientenräume und Homomorphiesatz

QuotraumDef **Definition 28.** Es sei $\boldsymbol{V}$ ein endlichdim. Vektorraum, und $\boldsymbol{W}$ ein Teilraum. Dann nennt man $\mathbf{v}_1$ und $\mathbf{v}_2$ *äquivalent modulo* $\mathbf{W}$ wenn ihre Differenz in $\boldsymbol{W}$ liegt[105]. Dann bildet der **Quotientenraum** $\boldsymbol{V}/\boldsymbol{W}$ von $\boldsymbol{V}$ modulo $\boldsymbol{W}$, also die Menge aller Äquivalenzklassen der Form

$$\boldsymbol{V}/\boldsymbol{W} := \{\tilde{\mathbf{v}} := \mathbf{v} + \boldsymbol{W} \,|\, \mathbf{v} \in \boldsymbol{V}\}, \tag{172}$$ quotdefform

einen Vektorraum bzgl. der naheliegenden Vektorraum-Struktur $\tilde{+}$ und $\tilde{\cdot}$ gegeben durch:

$$\tilde{\mathbf{v}_1}\tilde{+}\tilde{\mathbf{v}_2} := (\mathbf{v}_1 + \mathbf{v}_2)\tilde{}, \;\; \lambda\tilde{\cdot}\tilde{\mathbf{v}} := (\lambda\mathbf{v})\tilde{}. \tag{173}$$ vektquotdef

Einer der wichtigsten Sätze ist der folgende Homomorphie bzw. Isomorphiesatz:

homoqot0 **Theorem 25.** *Es sei T eine surjektive, lineare Abbildung von $\boldsymbol{V}_1$ nach $\boldsymbol{V}_2$, und $\mathbf{W} := T^{-1}\{0\}$ der Nullraum von T. Dann gilt folgende natürliche Isomorphie:*

$$\boldsymbol{V}_1/\boldsymbol{W} \approx \boldsymbol{V}_2. \tag{174}$$ quotism00

Proof. Der *natürliche* Isomorphismus ist gegeben durch $J(\mathbf{v}_1 + \boldsymbol{W}) := T(\mathbf{v}_1)$. Man muß dann nur(!) zeigen, dass die tatsächlich eine bijektive, lineare Abbildung ist. Die Umkehrabbildung ordnet jedem $\mathbf{v}_2 \in \boldsymbol{V}_2$ das Urbild $T^{-1}(\mathbf{v}_2)$ zu, d.h. $J^{-1} = T^{-1}$ "in gewissem Sinne, wobei J^{-1} eine echte inverse Abbildung ist, während T^{-1} nur die Urbild-Operation (die im Grunde mengen-wertig ist) darstellt. Durch die Bildung von Äquivalenzklassen wird die Urbildoperation aber zu einer Abbildung, die "Elementen" (von $\boldsymbol{V}_2$) Elemente (von $\boldsymbol{V}/\boldsymbol{W}$) zuordnet. □

Ein typischer Fall ist der Fall $\boldsymbol{V} = \mathbb{R}^n$, und $T : \mathbb{R}^n$ eine (durch eine Matrix $\boldsymbol{A}$ gegebene) lineare Abbildung nach $\mathbb{R}^m$, sowie $\boldsymbol{W} := Null(T)$ (Kern von $\boldsymbol{A}$). Dann ist die Identifizierung von oben nichts anderes als die Identifizierung des Bildraumes (= Spaltenraum von $\boldsymbol{A}$) mit dem Quotientenraum aus $\mathbb{R}^n$ nach $Null(T)$. Insbesondere der Fall unserer "Testmatrix" $\boldsymbol{A}$ vom Format 3×3 mit Rang $= 2$ zeigt, dass es sich um Geradenbüschel handelt ($Null(\boldsymbol{A})$ ist eindimensional). Dann ist es aber naheliegend, den Quotientenraum mit dem Zeilenraum von $\boldsymbol{A}$ (wieder: eigentlich Spalten von $\boldsymbol{A}'$) zu identifizieren. D.h. das Urbild mit dem eindeutigen Erzeuger aus dem Zeilenraum (der MNLSQ-Lösung > Methode der kleinsten Quadrate). Details dazu im kommenden SS 2011.

Illustration anhand von geometrisch intuitiven Beispielen in $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ (Geraden in der Ebene oder im Raum als Elemente).

[105] Man schreibt: $\mathbf{v}_1 \equiv \mathbf{v}_2$ wenn $\mathbf{v}_1 - \mathbf{v}_2 \in \boldsymbol{W}$

27 Affine Teilräume

In den Übungsbeispiele (59) und (60) (SS 2014) wurden *affine Abbildungen* kurz gestreift. Wir haben auch schon bemerkt, dass die in der Schule als *linear* bezeichneten Abbildungen $x \mapsto f(x) = ax + b$ (eher $y = kx + d$) im internationalen bzw. math. Sprachgebrauch als affine Abbildungen auf $\mathbb{R}^n$ bezeichnet werden, welche allgemein als

$$\vec{\mathbf{x}} \mapsto \boldsymbol{A} * \vec{\mathbf{x}} + \vec{\mathbf{b}}_0$$

geschrieben werden können (und genau dann invertierbar sind, wenn die $n \times n$-Matrix $\boldsymbol{A}$ nichtsingulär, d.h. invertierbar ist.

Wir wollen nun zeigen, dass der Übergang von *affinen Koordinaten* (hier anhand einer Ebene), von einer Parameterdarstellung zu einer anderen, eben eine affine Abbildung der Parameter entspricht, d.h. als affine Abbildung realisiert werden kann.

Wie wir schon in Beispiel (42) gesehen haben, hat die Matrix die interessante Eigenschaft, das der Zeilenraum mit dem Spaltenraum übereinstimmt. Weiters ist klar, dass je zwei dieser 6 Vektoren (3 Zeilenvektoren in $\mathbb{R}^3$, 3 Spaltenvektoren im $\mathbb{R}^3$) gemeinsam eine Basis f.d. Raum, nennen wir ihn $\boldsymbol{W}$, darstellen (es ist eine Ebene durch den Nullpunkt $\vec{\mathbf{0}} \in \mathbb{R}^3$.

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \tag{175}$$

Nimmt man also einen beliebigen Punkt und erzeugt die zu $\boldsymbol{W}$ parallele Ebene im $\mathbb{R}^3$ so hat man sofort eine Parameterdarstellung einer allgemeine Ebene. Man kann jede solche ebene aber auch einfach durch die Normalgleichung bestimmen (in unnormierter, dafür ganzzahliger Normalvektor wäre $\vec{\mathbf{n}} = [1; -2; 1]$).

Hat man nun zwei solche Darstellungen (sozusagen in affinen Koordinaten), etwa der Punkt $P = (1, 0, 1)$ als Koordinatenzentrum und die ersten beiden Spaltenvektoren als Richtungsvektoren, und einen passenden anderen Punkt in dieser Ebene (z.B. $Q = (6, 7, 10)$, bitte selber überprfen, dass dieser Punkt in der Ebene liegt!) und die beiden ersten Zeilenvektoren als Richtungsvektoren, um also dieselbe Ebene mit einer anderen Parameter-Darstellung zu beschreiben.

Was kann man dann über den Zusammenhang zwischen Koordinaten eines allgemeinen Punktes im ersten (mittels P) bzw. zweiten (mittels Q beschriebenen) *"Koordinatensystems"* sagen? (Stichwort: affine Abbildung, aber was sind die sinnvollen Fragen, was kann/muss man berechnen, um diesen Zusammenhang zu verstehen?).

28 Rechteckige Matrizen

Wir betrachten als Prototypen die Matrizen

```
AP = zeros(2,3);  AP([1,3]) = 1; AE = AP';
```

$$\boldsymbol{A}_P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{bzw.} \quad \boldsymbol{A}_E = \boldsymbol{A}_P^t := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{176}$$

Bei der ersten Abbildung handelt es sich um die (orthogonale) **P**rojektion von $\mathbb{R}^3$ auf die $x-y$-Ebene, welche auf natürliche Weise mit $\mathbb{R}^2$ identifiziert wird. Die Abbildung $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ ist selbstverständlich surjektiv ist, während im Gegensatz dazu die transponierte Matrix $\boldsymbol{A}_E$ (**E**inbettung) eine isometrische Einbettung, via $[x; y] \mapsto [x; y; 0]$ darstellt, welche klarerweise die lineare Unabhängigkeit von Vektoren bewahrt.

Man kann diese beiden natürlich diese beiden Matrizen in der einen oder anderen Reihenfolge zusammensetzen, und stellt dabei fest:

$$\boldsymbol{A}_P * \boldsymbol{A}_E = \mathbf{Id}_2 \quad \text{bzw.} \quad \boldsymbol{A}_E * \boldsymbol{A}_P = \mathbf{P}_2. \tag{177}$$

ProjEmb01

wobei $\mathbf{P}_2$ die Matrix der Projektionsabbildung (als Abbildung von $\mathbb{R}^3$ nach $\mathbb{R}^3$ aufgefasst) darstellt, also die Diagonalmatrix $\mathbf{P}_2 = diag([1; 1; 0])$.

Somit ist klar, dass $\boldsymbol{A}_E$ eine *Links-Inverse* hat, während hingegen $\boldsymbol{A}_P$ eine *Rechts-Inverse* hat (und trivialerweise beide Matrizen nicht invertierbar sein können).

28.1 Eigenschaften von rechteckigen Matrizen

Wir erinnern hier nochmals an Theorem 17.

linunabhmat

rechtsinv2 **Lemma 48.** *Eine Matrix $\boldsymbol{A}$ hat ein Rechts-Inverse $\mathbf{C}$, d.h.*

$$\boldsymbol{A} * \mathbf{C} = \mathbf{Id}_m$$

genau dann, wenn die Spalten von $\boldsymbol{A}$ ein Erzeugendensystem für $\mathbb{K}^m$ sind.

In der obigen Situation muß gelten $n \leq m$, denn alle Spaltenvektoren liegen im $\mathbb{K}^m$ und mehr als m Vektoren sind in diesem Vektorraum stets linear abhängig. Daher hat $\boldsymbol{A}$ maximalen Rang $r = min(m, n) = m$. Diese Eigenschaft bleibt unter Transposition erhalten, wobei dann klar ist, dass die transponierte Matrix dann $r = n$ erfüllen muss. So gesehen kann man das folgende Lemma 49 aus dem vorangehenden Lemma 48 leicht durch Transposition (deshalb auch Umkehrung der Reihenfolge!) herleiten.

linksinv2 **Lemma 49.** *Eine Matrix $\boldsymbol{A}$ hat ein Links-Inverse $\boldsymbol{B}$, d.h.*

$$\boldsymbol{B} * \boldsymbol{A} = \mathbf{Id}_n$$

genau dann, wenn die Spalten von $\boldsymbol{A}$ linear unabhängig sind.

29 Neustart im Januar 2011: ! Revision 2014 nötig

Auffrischung des Begriffes des Skalarproduktes auf $\mathbb{R}^n$ und $\mathbb{C}^n$. Beachte, dass $\langle \vec{\mathbf{x}}, \vec{\mathbf{x}} \rangle > 0$ genau dann wenn $\vec{\mathbf{x}} \neq \mathbf{0} \in \mathbb{C}^n$ gilt.

xxx **Lemma 50.** *Man überzeuge sich davon, dass für eine $m \times n-$matrix $\boldsymbol{A} = [\mathbf{a}_1; \ldots; \mathbf{a}_n]$ (aus n Spaltenvektoren in $\mathbb{C}^m$ bestehend) die folgende wichtige Identität gilt:*

$$\langle \boldsymbol{A}' * \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{y}, \boldsymbol{A} * \mathbf{x} \rangle, \quad \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m \tag{178}$$ adjungMATprop

und natürlich gilt aus Symmetriegründen gilt natürlich auch (setze $\boldsymbol{B} = \boldsymbol{A}'$)

$$\langle \boldsymbol{B} * \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{y}, \boldsymbol{B}' * \mathbf{x} \rangle, \quad \mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m \tag{179}$$ adjungMATprop

Wiederholung: WIEDERHOLUNG:

- Zwei Vektoren $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \mathbb{R}^n$ heißen **orthogonal zueinander** wenn

  $$\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle = \sum_{k=1}^{n} x_k \overline{y_k} = 0 \tag{180}$$ orthskal

  gilt. Man verwendet das Symbol $\mathbf{x} \perp \mathbf{y}$.

- Zwei Mengen $M, N \subset \mathbb{R}^d$ heißen *orthogonal zueinander* wenn je zwei Vektoren $\mathbf{m} \in M$ und $\mathbf{n} \in N$ zueinander senkrecht stehen.

- Sei $M \subset \mathbb{R}^d$ eine Menge, dann bezeichnet man die Menge aller auf M senkrecht stehenden Vektoren das *orthogonale Komplement* zu M:

  $$M^{\perp} := \{ \mathbf{z} \mid \langle \mathbf{m}, \mathbf{z} \rangle = \mathbf{0} \quad \forall \mathbf{m} \in \mathbf{M} \}. \tag{181}$$ orthcompl

Wichtig ist auch die Verträglichkeit des Skalarproduktes mit Linearkombinationen, d.h.

$$\left\langle \sum_{k=1}^{n} c_k \mathbf{a}_k \, , \mathbf{y} \right\rangle = \sum_{k=1}^{n} c_k \langle \mathbf{a}_k, \mathbf{y} \rangle. \tag{182}$$ skalplincomb

Diese wird in Lemma 23? benutzt um u.a. folgendes Lemma zu zeigen:

nullorth **Lemma 51.** *Der Nullraum der Matrix $\boldsymbol{A}$ ist das orthogonale Komplement des Zeilenraumes von $\boldsymbol{A}$ (richtiger: des Spaltenraumes von $\boldsymbol{A}'$).*

Zur Erinnerung: Man geht von einer Uminterpretation des durch eine Matrix $\boldsymbol{A}$ beschriebenen homogenen linearen Gleichungssystems aus. Jede Zeile (rechte Seite ist ja 0!) wird als eine Orthogonalitätsrelation betrachtet. Die Aussage, dass ein Vektor orthogonal auf alle Zeilen steht ist nichts anders als eine Uminterpretation dessen, was die "allgemeine Lösung eines homogenenen Gleichungssystems" ist. Steht ein Vektor orthogonal auf alle Zeilenvektoren, so auch auf deren Linearkombinationen (das ist aber genau der Zeilenraum!). Umgekehrt ist trivialerweise eine Vektor der senkrecht auf den Zeilenraum auch senkrecht auf jeden seiner Erzeuger (logischer Spezialisierung).

Wir wissen nun aber aus der Gauss Elimination schon, dass die Dimension des Zeilenraumes gerade der (Zeilen = Spalten-) Rang der Matrix $\boldsymbol{A}$ ist. Somit gilt für jeden r-dimensionalen Raum des $\mathbb{R}^n$, dass sein orthogonales Komplement $n-r$-dimensional ist (Zahl der freien Parameter).

Darauf ergibt sich aber folgende ebenfalls sehr wichtig Aussage:

dopporthog

Proposition 3. *Für jeden Teilraum $\boldsymbol{W} \subset \mathbb{R}^n$ und jede Menge $M \subseteq \mathbb{R}^n$ gilt*

$$(\boldsymbol{W}^\perp)^\perp = \boldsymbol{W} \quad \textit{bzw.} \quad (M^\perp)^\perp = \boldsymbol{V}_M. \tag{183}$$

dopporthform

Proof. Es ist eine (wirklich einfache) Übungsaufgabe (die aber logisch eine gewisse Konzentration erfordert! bitte selber machen!) zu sehen dass für jede Menge M gilt $M \subseteq (M^\perp)^\perp$. Klarerweise ist die rechte Seite eine Vektorraum, daher gilt: $\boldsymbol{V}_M \subseteq (M^\perp)^\perp$, denn nach Definition ist $\boldsymbol{V}_M$ der *kleinste* Teilraum von $\mathbb{R}^n$ der M enthält (wir wissen, er besteht genau aus den Linearkombinationen der Elemente von M).

Wenn also $M = \boldsymbol{W}$ ein Teilraum ist, sie haben wir einen r-dimensionalen (sagen wir) Raum $\boldsymbol{W}$, dessen orthog. Komplement $n-r$-dimensional ist, also hat $(\boldsymbol{W}^\perp)^\perp$ die Dimension r, also ist der Raum gleich gross wie $\boldsymbol{W}$. Konkret: Jede Basis von $\boldsymbol{W}$ hat r Elemente, und ist daher ein linear unabh. System in dem r-dim. Raum $(\boldsymbol{W}^\perp)^\perp$, d.h. es ist ein max. lin. unab. System, also eine Basis. Damit ist der erste Teil der Aussage erfüllt.

Für den allgemeinen Fall können wir davon ausgehen, dass die Menge M endlich ist (und nicht mehr als n Elemente hat, denn ansonsten können wir ohne den Raum $\boldsymbol{V}_M$ zu ändern alle lin. abh. Elemente von M entfernen!). Dann können wir M aber als eine Menge von Zeilenvektoren einer $m \times n$-Matrix $\boldsymbol{A}$ auffassen, der von M erzeugt Raum ist also dann der Zeilenraum $\mathbf{Z}(\boldsymbol{A})$. Seine Dimension ist dann aber gerade $r = \mathit{Rang}(\boldsymbol{A})$, sein orthog. Komplement ist dann aber $n-r$-dimensional (es ist der Nullraum von $\boldsymbol{A}$!). das doppelte orthog. Komplement hat dann aber genau die Dimension $n-(n-r) = r$, somit ist $\mathbf{Z}(\boldsymbol{A}) = \boldsymbol{V}_M \subseteq (M^\perp)^\perp$ eine Inklusion von r-dim. Räumen, und wieder ergibt sich Gleichheit[106]. □

Eine der letzten (und im Detail noch auszuführenden) Aussagen vor Weihnachten lautete (kurz zusammengefasst) wie folgt:

Proposition 4. *Die folgenden Eigenschaften sind äquivalent für n Vektoren $\mathbf{a}_1, \dots \mathbf{a}_n$ in $\mathbb{R}^n$ (oder $\mathbb{C}^n$), aufgefasst als Spalten oder Zeilen einer $n \times n$-Matrix $\boldsymbol{A}$:*

- *Die Matrix $\boldsymbol{A}$ (bzw. $\boldsymbol{A}^t$ bzw. $\boldsymbol{A}'$) ist invertierbar bzw. die Vektoren bilden eine Basis;*
- *Eine (dann eindeutig bestimmte) inverse Matrix kann mit Hilfe der Gauss'schen Eliminationsmethode bestimmt werden;*
- *Die Vektoren sind ein Erzeugendensystem für $\mathbb{C}^n$, d.h. sie spannen $\mathbb{C}^n$ auf;*
- *Die Vektoren sind linear unabhängig*

[106] Kann gut sein, dass sich eine etwas kürzere Argumentation für diesen Sachverhalt finden läßt. Vorschläge in diese Richtung sind willkommen!

Details dazu folgen noch!

Daraus ergibt sich unter anderem folgende interessante Konsequenz:

onesidinv **Corollary 12.** *Hat eine $n \times n$-Matrix $\boldsymbol{A}$ eine Links- oder eine Rechts-Inverse Matrix $\boldsymbol{B}$, dann ist $\boldsymbol{B}$ auch die (zweiseitig) inverse Matrix.*

Proof. Wir müssen zwei Fälle betrachten: Nehmen wir an, dass $\boldsymbol{B} * \boldsymbol{A} = Id_n$ gilt, dann ist die Komposition der Matrix-Multiplikation von $\boldsymbol{A}$ (zuerst angewendet) mit $\boldsymbol{B}$ (danach) eine (offenbar) injektive Abbildung. Daraus folgt, dass auch $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ eine injektive Abbildung sein muss, d.h. die Spalten von $\boldsymbol{A}$ sind linear unabhängig, sind also dann doch eine Basis, und $\boldsymbol{A}$ muss invertierbar sein. Die Links-Inverse Matrix $\boldsymbol{B}$ muss also (wegen der bekannten Eindeutigkeit des inversen Elements in einer allgemeinen Gruppe) auch eine Rechts-Inverse sein.

Im anderen Fall geht man von der Annahme $\boldsymbol{A} * \boldsymbol{B} = Id_n$ aus, und verwendet das Surjektivitätsargument. Die Spalten von $\boldsymbol{A}$ müssen also ein Erzeugendensystem für den $\mathbb{C}^n$ sein, also ist wieder die Invertierbarkeit von $\boldsymbol{A}$ gegeben (aufgrund der Proposition oben), also haben wir analog zu oben: $\boldsymbol{B}$ muss auch gleichzeitig ein links-inverses Element sein! □

Für Beispiele bedeutet es: Das orthogonale Komplement (m) von Vektoren zu beschreiben in $\mathbb{C}^n$ bedeutet es

In gleicher Weise ist das orthog. Komplement des Spaltenraumes $\mathbf{Sp}(\boldsymbol{A})$ von $\boldsymbol{A}$ genau der Nullraum der linearen Abbildung $S : \mathbf{y} \mapsto \mathbf{y} * \boldsymbol{A}'$. Sehen wir uns diese kurz an (eine *einfache* aber *wichtige* Beobachtung:):

sdf **Lemma 52.** *Die Abbildung $S : \mathbf{y} \mapsto \mathbf{y} * \boldsymbol{A}'$ bildet $\mathbf{y}$ auf die Folge*

$$(\langle \mathbf{y}, \mathbf{a}_k \rangle)_{k=1}^n.$$

Daraus ergeben sich sofort interessante Eigenschaften für unitäre (im Reellen: orthogonale) Matrizen:

Angewendet auf unitäre Matrizen haben wir

sdfsdf **Lemma 53.** *Ein Matrix erfüllt $\mathbf{U} * \mathbf{U}' = Id_n$ genau dann wenn gilt $\mathbf{U}' * \mathbf{U} = Id_n$.*

Die zweite Aussage ist in der Tat äquivalent zu der geometrischen Interpretation eines *Orthonormalsystems*, die Vektoren erfüllen $\|\mathbf{u}_k\|^2 = 1$, haben also euklidische Länge 1, und stehen paarweise zueinander orthogonal ist nur eine Uminterpretation der Aussage dass $\mathbf{U}' * \mathbf{U} = Id_n$ gilt.

Das ist aber aufgrund des obigen Corollaries äquivalent zu der Aussage $\mathbf{U} * \mathbf{U}' = Id_n$, was andererseits wieder dasselbe bedeutet wie

$$\mathbf{x} = Id_n * \mathbf{x} = \mathbf{U} * (\mathbf{U}' * \mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{C}^n.$$

Diese letzte Gleichung aber ist gleichbedeutung mit der Tatsache, dass die **U**-Koeffizienten von $\mathbf{x} \in \mathbb{C}^n$ gerade die Skalarprodukte $\langle \mathbf{x}, \mathbf{u}_k \rangle$ sind, oder aber (weiterhin gleichbedeutend) mit der sehr einfachen und einprägsamen Darstellung

$$\mathbf{x} = \sum_{k=1}^{n} \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k, \quad \forall \mathbf{x} \in \mathbb{C}^n. \tag{184}$$

orthexpans

Es fehlt im Moment noch: $\mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{u}_k \rangle$ ist die Best-Approximation von $\mathbf{u}$ durch Vielfache von $\mathbf{u}_k$, d.h. für jeden anderen Wert von λ gilt

$$\|\mathbf{x} - \lambda \mathbf{u}_k\| \geq \|\mathbf{x} - \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k\|.$$

- Stoff für den Test am 17. Jan. 2011, Trainingsfragen
- Gibt es (aufgrund des Lernens in den Weihnachtsferien) [konkrete] Fragen?
- Ergänzung von Lücken (z.B. Erzeugendensysteme mit n Elementen in einer $n \times n$-Matrix bilden automatisch eine Basis, und es gibt sogar eine Garantie, dass eine (passende gewählte) Form von Gauss'schen Eliminationsschritten zum Inversen Matrix führen kann!
- Geometrische Interpretation von Matrizen (als "Bewegungen" im Raum);
- noch weiterer, neuer Stoff, z.B. affine Räume sind genau die Lösungsmengen von inhomogenen linearen Gleichungssystemen;

`http://de.wikipedia.org/wiki/Affiner_Raum#Definition_der_linearen_Algebra`

`http://de.wikipedia.org/wiki/Cauchy-Schwarzsche_Ungleichung`

Wir werden der dort gegebenen Argumentation folgen, die die **Cauchy-Schwarz Ungleichung** aus der arith-geom. Mittelwerts-Ungleichung herleitet, die da lautet: Für $u, v \geq 0$ gilt:

$$\sqrt{uv} \leq \frac{u+v}{2}. \tag{185}$$

AGUngl

d.h. in Worten, das geometrische Mittel zweier positiver Zahlen ist immer kleiner gleich dem arithmetischen Mittel (Beweis: Übungsaufgabe, die aber auch beim Kolloqium gefragt werden kann).

Für den Beweis der sogenannten Dreiecksungleichung für die **Euklidische Norm** (auch *Länge des Vektors* genannt) ist die Cauchy-Schwarz'sche Ungleichung wichtig. Zur Erinnerung

CSUngLemma **Lemma 54.** *Für beliebige Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ gilt die CS Ungleichung.*

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|x\| \|\mathbf{y}\| \tag{186}$$

CSUngl

Proof. Offenbar ist die CS-Ungleichung trivialerweise richtig wenn einer der beiden Vektoren der Null-Vektor in $\mathbb{C}^n$ ist. Wenn also beide positive Länge, und man kann (186) CSUngl äquivalent umformen, indem man durch die rechte Seite dividiert, und die (pos.) Konstanten in das Skalarprodukt *hineinzieht*. D.h. wir haben dann zu beweisen dass gilt:

$$\left|\left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\rangle\right| \leq 1 \tag{187}$$ CDUnglNorm

Die Abschätzung beginnt für komplexe Vektoren mit der Abschätzung:

$$|\langle \vec{\mathbf{x}}, \vec{\mathbf{y}} \rangle| = |\sum_{k=1}^{n} x_k y_k| \leq \sum_{k=1}^{n} |x_k||y_k|, \tag{188}$$ abs-estim1

was aber weiter abgeschätzt werden kann
(weil ja $2uv \leq u^2 + v^2$, für $u = \sqrt{|x_k|}$ und $v = \sqrt{|y_k|}$) mit

$$1/2 \cdot \sum_{k=1}^{n} (|x_k|^2 + |y_k|^2) \leq 1/2(\sum_{k=1}^{n} |x_k|^2) + 1/2(\sum_{k=1}^{n} |y_k|^2) = 1. \tag{189}$$ prod-sum1

□

Die CS-Ungleichung erlaubt es auch, den Winkel zwischen zwei Vektoren mit Hilfe des COS-Funktion zu definieren. In der Tat, die Funktion $t \mapsto \cos(t)$ ist bijektiv zwischen $[0, \pi]$ und $[-1, 1]$, d.h. für jedes $s \in [-1, 1]$ (und für reelle Vektoren sind die Skalarprodukte der normierten Vektoren genau in dem Intervall, wegen CS!) gibt es einen eindeutig bestimmten Winkel in $\alpha \in [0, 180°]$ sodass $\cos(\alpha) = s$.

Ein zugehöriges MATLAB Program sieht etwa wie folgt aus:

```
% winkel.m  hgfei  Jan. 2011
% Usage:  phi = winkel(a,b);
% Input: a,b   Vektoren in C^n, Zeilen oder Spalten! r

function phi = winkel(a,b);

% Berechnung von phi
phi = acos(b(:)'*a(:)/(norm(a)*norm(b)));
phi = 180 * phi /pi;        % Umrechnung in Grad
```

Man kann leicht zeigen, dass Orthogonalität eine Bedingung ist, die stärker ist als lineare Unabhängigkeit:

orthlinindep **Lemma 55.** *Jede Folge von (von Null verschiedenen) Vektoren $\mathbf{x}_k, 1 \leq k \leq s$ in $\mathbb{C}^n$ die ein Orthogonalsystem bildet, d.h. die Bedingung*

$$\langle \mathbf{x}_j, \mathbf{x}_k \rangle = 0 \quad \textit{for} \quad k \neq j \tag{190}$$ orthdef

ist auch linear unabhängig.

Proof. Wir haben zu zeigen, dass eine Linearkombination diese Vektoren, die Null ist, nur in trivialer Weise (d.h. lauter Null-Koeffizienten) entstehen kann. Sei also $\mathbf{z} = \sum_{k=1}^{s} c_k \mathbf{x}_k = \mathbf{0}$. Um für beliebiges $j \in \{1, \ldots, s\}$ zu zeigen, dass $c_j = 0$ gilt, müssen wir nur feststellen, dass

$$\langle \mathbf{z}, \mathbf{x}_j \rangle = \langle \mathbf{0}, \mathbf{x}_j \rangle = 0$$

gilt. Anderseits gilt aber aufgrund der Linearität des Skalarproduktes in der ersten Variablen

$$\langle \mathbf{z}, \mathbf{x}_j \rangle = \sum_{k=1}^{s} c_j \langle \mathbf{x}_k, \mathbf{x}_j \rangle = c_j \|\mathbf{x}_j\|^2.$$

Insgesamt (und unter Verwendung der Annahme dass $\|\mathbf{x}_j\| \neq 0$ gilt, folgt also $c_j = 0$, und zwar für jedes einzelne $j \in \{1, \ldots, s\}$. □

pythag **Lemma 56.** *Seien* $\mathbf{a}, \mathbf{b}$ *zwei zueinander orthogonale Vektoren in* $\mathbb{C}^n$*, dann gilt*

$$\|\mathbf{a} + \mathbf{b}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2.$$

Proof. Es genügt, die linke Seite (definitionsgemäß) durch ein Skalarprodukt auszudrücken und dieses dann auszumultiplizieren (Distributivgesetz), um die Orthogonalität ins Spiel zu bringen. Dieser Beweis sollte jederzeit reproduzierbar sein. □

Corollary 13. *Für jede Orthonormalbasis* $\mathbf{U}$ *gilt für* $\mathbf{x} = \sum_{k=1}^{n} c_k \mathbf{u}_k$*:*

$$\|\mathbf{x}\| = \sqrt{\sum_{k=1}^{n} |c_k|^2}.$$

Proof. Man wende (genaugenommen unter Anwendung des Induktionsprinzip, nach n) den Satz von Pythagoras auf die zueinander orthogonalen Komponenten von $\mathbf{x}$, nämlich $c_k \mathbf{u}_k = \langle \mathbf{x}, \mathbf{u}_k \rangle \mathbf{u}_k$ an. Klarerweise gilt $\|c_k \mathbf{u}_k\|^2 = |c_k|^2 \|\mathbf{u}_k\|^2 = |c_k|$. □

bestapp1 **Lemma 57.** *Die Bestapproximation eines Vektors* $\mathbf{x}$ *durch Vielfache eines gegebenen Richtungsvektors*[107] $\mathbf{n}$ *(also mit* $\|\mathbf{n}\| = 1$*) ist gegeben durch*

$$\mathbf{y} = \langle \mathbf{n}, \mathbf{x} \rangle \mathbf{n}.$$

Proof. Hier ist der Beweis vielleichte genauso wichtig wie die Aussage selber. Wir betrachten den allgemeinen Ansatz, d.h. wir bestimmen für variables λ den Abstand von $\mathbf{x}$ zu $\lambda \mathbf{n}$. Mit Hilfe der Analysis könnte man diese Funktion minimieren, indem man ein Minimum der Ableitung sucht, was zu $\lambda = \langle \mathbf{x}, \mathbf{n} \rangle$ führen würde.

Eine *gute Übungsaufgabe* ist die folgende (wir wollen den Beweis hier nicht einfach vorführen):

Frobnormest **Lemma 58.** *Es sie* $\boldsymbol{A}$ *eine (reelle oder komplexe)* $n \times n$*-Matrix. Die Frobenius-Norm von* $\boldsymbol{A}$ *ist gegeben durch*

$$\|\boldsymbol{A}\|_F := \sqrt{\sum_{k,j} |a_{k,j}|^2},$$

dann gilt für jedes $\mathbf{x} \in \mathbb{C}^n$*:*

$$\|\boldsymbol{A} * \mathbf{x}\| \leq \|\boldsymbol{A}\|_F \|\mathbf{x}\|. \tag{191}$$ Frobinequ

[107] G. Strang nennt das den Einheitsvektor in die Richtung von $\mathbf{x}$.

In MATLAB kann man die Frobenius-Norm auf verschiedene Arten:

```
norm(A(:)) , oder   norm(A,'fro')   oder  sqrt(sum(sum(abs(A).^2)));
```

Wir wollen mehr geometrisch argumentieren, indem wir feststellen, dass genau für diesen speziellen Wert von λ gilt:

$$\mathbf{x} - \lambda\mathbf{n} \quad \perp \quad \mathbf{n}.$$

In der Tat, die Orthongonalität bedeutet umgedeutet

$$\langle \mathbf{x}, \mathbf{n} \rangle \; = \; \lambda \langle \mathbf{n}, \mathbf{n} \rangle = \lambda.$$

also genau $\lambda = \langle \mathbf{x}, \mathbf{n} \rangle$.
Um zu zeigen, dass der Abstand von $\mathbf{x}$ zu $\mu\mathbf{n}$ für jeden anderen Wert von μ echte größer ist, Verwenden wir den Satz von Pythagoras, angewendet auf $\mathbf{x}-\mathbf{y}$ und $\mathbf{y}-\mu\mathbf{n} = (\mu-\lambda)\mathbf{n}$, die ja auch aufeinander senkrechte stehen! Also gilt

$$\|\mathbf{x} - \mu\mathbf{x}\|^2 = \|\mathbf{x} - \mathbf{y}\|^2 + |\mu - \lambda|^2.$$

Man sieht sofort, dass das der Minimalabstand ($\mathbf{x} - \langle \mathbf{x}, \mathbf{n} \rangle \mathbf{n}\|$) genau dann angenommen wird, wenn $\mu = \lambda = \langle \mathbf{x}, \mathbf{n} \rangle$ gilt. □

Eine Umdeutung dieses Resultates läuft auf folgende Beobachtung hinaus, wobei der Begriff der sog. *Gram-Matrix* ins Spiel kommt.

Definition 29. Es sei $\boldsymbol{A}$ eine $m \times n$-Matrix, aufgefasst als eine Folge von n Spaltenvektoren in $\mathbb{C}^m$. Dann ist $\mathbf{G} := \boldsymbol{A}' * \boldsymbol{A}$ die zugehörige Gram-Matrix, deren Eintragungen genau die paarweisen Skalarprodukte der Spalten von $\boldsymbol{A}$ sind, d.h. mit

$$g_{j,k} = \langle \mathbf{a}_k, \mathbf{a}_j \rangle, \quad 1 \leq j \leq n.$$

Dann sind die Orthogonalsystem genau diejenigen Folgen von s Vektoren im $\mathbb{C}^m$ für die die Gram-Matrix eine invertierbare Diagonal-Matrix ist (man erinnere sich daran, dass eine Diagonalmatrix genau dann invertierbar ist, wenn alle Diagonal-Eintragungen von Null verschieden sind, und dass die inverse Matrix einfach die Matrix mit den Kehrwerten der Diagonaleintragungen ist).
Die lineare Unabhängigkeit ist im Endeffekt ein Spezialfall einer (viel) allgemeineren Aussage:

Gramtest **Lemma 59.** *Eine Familie von Spaltenvektoren (zusammengefaßt in einer Matrix $\boldsymbol{A}$) ist genau dann linear unabh. wenn die zugehörige Gram-Matrix $\boldsymbol{A}' * \boldsymbol{A}$ invertierbar ist.*

Proof. Wir beginnen mit der Beobachtung, dass der Nullraum von $\boldsymbol{A}$ gleich dem Nullraum von $\boldsymbol{A}' * \boldsymbol{A}$ ist.
Klarerweise impliziert $\boldsymbol{A} * \mathbf{x} = \mathbf{0}$ auch $\boldsymbol{A} * \boldsymbol{A} * \mathbf{x} = \mathbf{0}$, oder kurz $Null(\boldsymbol{A}) \subseteq Null(\boldsymbol{A}' * \boldsymbol{A})$.
Sei aber umgekehrt $\mathbf{z} \in Null(\boldsymbol{A}' * \boldsymbol{A})$, dann gilt auch

$$\langle \boldsymbol{A}' * \boldsymbol{A} * \mathbf{z}, \mathbf{z} \rangle = 0,$$

aber das ist gleichbedeutend mit

$$\langle \boldsymbol{A} * \mathbf{z}, \boldsymbol{A} * \mathbf{z} \rangle = 0,$$

was wiederum bedeutet, dass auch $Null(\boldsymbol{A}' * \boldsymbol{A}) \subseteq Null(\boldsymbol{A})$ gilt.

Sind also die Spalten von $\boldsymbol{A}$ linear unabhängig (was gleichbedeutend ist mit) $Null(\boldsymbol{A}) = \{\mathbf{0}\}$, dann ist auch die Gram-Matrix invertierbar.

Ist aber umgekehrt die Gram-Matrix invertierbar, dann gilt $inv(\boldsymbol{A}' * \boldsymbol{A}) * \boldsymbol{A} * \boldsymbol{A} = Id_n$. Da diese zusammengesetzte Abbildung also bijektiv ist, muss sich auch injektiv sein, also muss die zuerst angewendete Abbildung $\mathbf{x} \mapsto \boldsymbol{A} * \mathbf{x}$ auch injektiv sein, was wiederum bedeutet, dass die Spalten von $\boldsymbol{A}$ eine linear unabh. Folge in $\mathbb{C}^m$ bilden. □

Nachtrag (ein Beispiel zum Üben!)

- Aufgaben aus der 5.-ten Klasse: Lösung eines quadratischen Gleichungssystems: $x^2+px+q = 0$ bedeutet: Lösen eines quadratischen Gleichungssystems. Die übliche Methode: ergänze auf eine vollständiges Quadrat: $x^2 + px + (p/2)^2 = (p/2)^2 - q$. Für $y := x + p/2$ gilt somit $y^2 = \pm\sqrt{p^2/4 - q}$ oder zurückübersetzt: $x_{1,2} = -p/2 \pm \sqrt{p^2/4 - q}$. Zur Probe verifizert man, dass $(x - x_1) \cdot (x - x_2) = x^2 + px + q$ gilt (Vieta)!

  Aufgabenstellung: Man beschreibe den Graphen der Funktion $y = x^2 + 6x - 4$ und bestimme die Nullstellen der Funktion. Welcher Bereich der Funktion ist interessant (Kurvendiskussion!?).

30 Das Kreuzprodukt [speziell im $\mathbb{R}^3$]

Das *Kreuzprodukt* ist eine Besonderheit des $\mathbb{R}^3$, für die es in anderen Dimensionen nichts Entsprechendes gibt. Für $\vec{u}, \vec{v} \in \mathbb{R}^3$ definiert man das Kreuzprodukt von $\vec{u}, \vec{v}$ als

$$\vec{u} \times \vec{v} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}. \tag{192}$$

Das Kreuzprodukt wird auch als *äußeres Produkt* oder als *Vektorprodukt* bezeichnet. Die Bezeichnung Kreuz*produkt* rührt von dem Umstand her, dass man es als bilineare Abbildung auffassen kann:

$$(\vec{u}, \vec{v}) \mapsto \vec{u} \times \vec{v}. \tag{193}$$

Diese Abbildung ist bilinear

$$(\alpha \vec{u}_1 + \beta \vec{u}_2) \times \vec{v} = \alpha(\vec{u}_1 \times \vec{v}) + \beta(\vec{u}_2 \times \vec{v}), \tag{194}$$

$$\vec{u} \times (\alpha \vec{v}_1 + \beta \vec{v}_2) = \alpha(\vec{u} \times \vec{v}_1) + \beta(\vec{u} \times \vec{v}_2), \tag{195}$$

für beliebige reelle Zahlen α, β und Vektoren $\vec{u}_1, \vec{u}_2, \vec{u}, \vec{v}_1, \vec{v}_2, \vec{v}$.

Im Gegensatz zur Skalarmultiplikation oder dem inneren Produkt ist das Kreuzprodukt nicht kommutativ, aber es gilt

$$\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}, \tag{196}$$

welches man als *Anti-Kommutativität* des Kreuzproduktes bezeichnet. Aus der Anit-Kommutativität ergibt sich

$$\vec{u} \times \vec{u} = \vec{0}, \quad \text{für alle} \quad \vec{u} \in \mathbb{R}^3. \tag{197}$$

Damit gilt auch für jedes $\lambda \in \mathbb{R}$:

$$\vec{u} \times \lambda \vec{u} = \vec{0}, \quad \text{für alle} \quad \vec{u} \in \mathbb{R}^3. \tag{198}$$

Man kann auch zeigen (> Determinanten) dass auch die Umkehrung gilt, d.h. man kann die lineare Abhängigkeit von zwei Vektoren durch das Kreuzprodukt checken. Man kann zeigen (später), dass die Länge von $\vec{u} \times \vec{v}$ gleich der Fläche des von $\vec{u}$ und $\vec{v}$ aufgespannten Parallelogrammes ist (und diese ist genau dann strikt positiv wenn ...).

Wichtig ist dies auch zur die Beschreibung einer Ebene durch den Normalvektor, denn wenn zwei Vektoren $\vec{u}$ und $\vec{v}$ linear unabhängig sind, spannen sie eine Ebene auf, die in Normalform durch eine Gleichung der Form $\langle \vec{x}, \vec{n} \rangle = \gamma$ für eine $\gamma \in \mathbb{R}$ beschrieben werden kann. Man nehme als $\vec{n}$ einfach den Richtungsvektor zu $\vec{u} \times \vec{v}$.

Man sagt $(\mathbb{R}^3, \times)$ ist eine reelle Algebra. Schreibt man $\vec{u}, \vec{v}$ als Linearkombination der Einheitsvektoren $\{e_1, e_2, e_3\}$ des $\mathbb{R}^3$, so sieht man, daß$\vec{u} \times \vec{v}$ allein durch die Produkte $\vec{e}_i \times \vec{e}_j$ festgelegt ist.

Konkret gilt: $\vec{e}_i \times \vec{e}_j = \vec{e}_k$ für alle i, j, k, die durch zyklisches Vertauschen von $1, 2, 3$ hervorgehen, also:

$$\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2, \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1 \,. \tag{199}$$

Neben der Anti-Kommutativität gibt es noch einen weiteren Unterschied zur Skalarmultiplikation. In der Algebra $(\mathbb{R}^3, \times)$ gilt **nicht** das *Assoziativgesetz*, z.B. ist

$$(\vec{e}_1 \times \vec{e}_1) \times \vec{e}_2 = 0 \quad \text{aber} \quad \vec{e}_1 \times (\vec{e}_1 \times e_2) = -\vec{e}_2 \ .$$

NICHT Prüfungsstoff: Jedoch gilt folgende Identität von Grassmann

$$\vec{u} \times (\vec{v} \times \vec{w}\,) = \langle \vec{u}, \vec{w}\,\rangle\, \vec{v} - \langle \vec{u}, \vec{v}\,\rangle\, \vec{w}\,. \tag{200}$$

Es gibt zwei Möglichkeiten diese Identität zu beweisen, entweder man prüft sie für die Einheitsvektoren und verwendet die Bilinearität des Kreuzproduktes für den allgemeinen Fall, oder man rechnet sich die Komponenten aus.[108]

Vertauscht man in der Grassmann Identität die Vektoren zyklisch und addiert sie auf so erhält man

$$\vec{u} \times (\vec{v} \times \vec{w}\,) + \vec{w} \times (\vec{u} \times \vec{v}\,) + \vec{v} \times (\vec{w} \times \vec{u}\,) = 0 \tag{201}$$

einen Ersatz für das Assoziativgesetz, welche unter dem Namen *Jacobi-Identität* in der Literatur zu finden ist.

Einfache Folgerungen aus der Definition des Kreuzproduktes und der Grassmann-Identität:

1. Wenn $\vec{u} \times \vec{v} = 0$ für alle $\vec{u}$ ist, dann ist $\vec{v} = 0$.
2. $\vec{u} \times \vec{v} = 0 \Leftrightarrow \quad \vec{u}, \vec{v}$ linear abhängig sind.
3. $\langle \vec{u}, \vec{u} \times \vec{v}\,\rangle = 0$, d.h. das Kreuzprodukt ist orthogonal zu $\vec{u}$.
4. $\langle \vec{v}, \vec{u} \times \vec{v}\,\rangle = 0$, d.h. das Kreuzprodukt ist orthogonal zu $\vec{v}$.
5. $\vec{u} \times \vec{v}$ ist orthogonal zu der von $\vec{u}, \vec{v}$ aufgespannten Ebene.

Die folgende Identität erlaubt es einem dem Vektorprodukt eine geometrische Deutung zu geben:

$$\langle \vec{u}, \vec{v}\rangle^2 + \|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \|\vec{v}\|^2. \tag{202}$$

grassmannf

Dies sieht man am einfachsten in dem man es "stur" ausrechnet. Weiters ergibt sich daraus auch, dass das Kreuzprodukt mit orthogonalen Transformationen verträglich ist (siehe MATLAB Experiment), denn **im** $\mathbb{R}^3$ ist der Normalvektor (bis auf das Vorzeichen) eindeutig bestimmt, und seine Länge aus Formel (202). D.h.

$$(\mathbf{U} * \vec{u}) \times (\mathbf{U} * \vec{v}) = \mathbf{U} * (\vec{u} \times \vec{v}). \tag{203}$$

UinvKrz

Da $\|\vec{u} \times \vec{v}\|$ stets positiv ist, erhalten wir einen einfachen Beweis der Ungleichung von Cauchy-Schwarz und überdies eine Aussage überdies eine Aussage über die Gleichheit in der Ungleichung von Cauchy-Schwarz. Es besteht genau dann Gleichheit in der Ungleichung, wenn in der obigen Gleichung das Kreuzprodukt verschwindet, wie wir oben gesehen haben, ist dies genau dann der Fall, wenn die beiden Vektoren $\vec{u}, \vec{v}$ linear abhängig sind. Dann (und nur dann gilt) $|\langle \vec{u}, \vec{v}\rangle^2| = \|\vec{u}\|^2 \|\vec{v}\|^2$.

Das Kreuzprodukt ist insbesondere unter orthogonalen Transformation invariant (und auf diese Weise im Prinzip eindeutig festgelegt). Wir "verifizieren" das zuerst einmal nur durch eine kleines MATLAB Experiment:

[108] Für die Prüfung werden nur die elementaren Eigenschaften und die Definition des Kreuzproduktes verlangt.

```
>> U = orth(rand(3));  % erzeugt zufaellige Orthonormalbasis (orthog. Matrix)
>> u = rand(3,1); v = rand(3,1); w = kreuz(u,v);
>> norm( U*w - kreuz(U*u,U*v))
Ergibt:  ans =   1.4752e-016,  d.h. numerisch gleichwertig
unter Verwendung des selbst-geschriebenen M-files

function [nv] = kreuz(a,b);
% Berechnung des Normalvektors:
nv(1,1)=  a(2,1)*b(3,1)-a(3,1)*b(2,1);
nv(2,1)=-(a(1,1)*b(3,1)-a(3,1)*b(1,1));
nv(3,1)=  a(1,1)*b(2,1)-a(2,1)*b(1,1);
```

31 Gilbert Strang: Material

Im Prinzip sollen bis zum Ende des ersten Semesters ca. die ersten 200 Seiten des Buches von Gilbert Strang abgearbeitet sein (mit kleinen Ausnahmen), plus Kernteile von Paragraph 7 (lineare Abbildungen und deren Matrix-Darstellung).

Zusätzlich (teilweise auch als Illustration gedacht) wuerde einige Dinge über (quadratische und kubische) Polynomfunktionen gemacht (z.B. Vandermonde-Matrix).

Auch im Skriptum von Andreas Cap finden sich noch die Kernteile der Vorlesung, ebenso natürlich in fast allen Büchern zum Thema, von denen viele (englisch-sprachige) auch an der NuHAG entlehnbar sind.

Regeln für Matrixoperation: Block-Matrizen (wird wohl nicht gemacht), p.67-69.

LU-Faktorisierung: p. 88 - 99 bzw. 108 - 113.

Dafür gibt es etliche andere/neue Aussagen über den Vektorraum von (quadratischen bzw. kubischen) Polynomfunktionen.

Section 3.6: Dimension der 4 Unterräume (wird noch gemacht!)

wichtig: Skizze auf Seite p.185: four spaces

Wichtig: Fundamentalsatz der LA, Teil 2.

Gram-Schmidt: wird erst im Sommersemester (G.Strang: Section 4.4), p.229 - 243. ebenso G.Strang: Determinanten! (section 5)

ZUSAMMENFASSUNG: page: BUCH bis page 200 PLUS Section 7 von G. Strang: Lineare Abbildungen und Matrix.

32 Determinanten, Regel von Sarrus

Die allgemeine Diskussion über Determinanten und ihre Eigenschaften erfolgt im zweiten Teil der Lin.Agl.-Vorlesung (die sog. LA 1). Für den Fall von 2×2 bzw. 3×3-Matrizen $\boldsymbol{A}$ ist die Determinante duch ein einfaches Rezept zu beschreiben.

tbtdet2 **Lemma 60.** *Für allgemeine* 2×2 *Matrizen* $\boldsymbol{A} = [a, c; b, d]$; *gilt:* $\boldsymbol{A}$ *ist genau denn invertierbar, wenn die Determinant, d.h.*

$$\det(\boldsymbol{A}) = ac - bd \tag{204}$$ tbtdetdef

von Null verschieden ist.

Proof. Wir wissen, dass eine quadratische Matrix genau dann invertierbar ist, wenn die Spalten linear unabhängig sind, also NICHT Vielfache voneinander sind. Linear Abhängigkeit ist aber gleichbedeutend mit "gleichen Proportionen", d.h. $a : b = c : d$. Invertierbarkeit ist somit gleichzusetzen mit dem Nichtverschwinden der Determinante. □

Für den Fall von 2×2-Matrizen gibt es eine ähnliche Regel. Im Skriptum von A. Cap wird die Regel von Sarrus damit begründet, dass für jeden möglichen, festgelegten Wert von x_1 die Matrix, die die Bestimmung von x_2 und x_3 bewerkstelligen soll, von Null verschieden sein soll. Andere Begründungen basieren auf der Berechnung des Volumens des Parallel-Epipeds, welches von den drei Spalten von $\boldsymbol{A}$ im $\mathbb{R}^3$ aufgespannt wird, und das genau dann nicht gleich Null ist, wenn die drei Vektoren nicht in einer Ebene (oder gar auf einer Geraden) liegen. Die Regel von Sarrus funktioniert wie folgt: siehe A.Cap Skriptum ...

33 WICHTIGE repoduzierbare Beweise (wird 2014 noch erweitert)

1. Gauss Elimination als Matrixmultiplikation mit Elementar Matrizen;
2. Zu jeder Menge M gibt es einen kleinsten Vektorraum $\boldsymbol{V}_M$ der M enthält, nämlich die Linearkombinationen
3. Zeige, dass man aus einer linear abhängigen Menge jeweils wenigstens einen Vektor wegnehmen kann, ohne das lineare Erzeugnis zu verändern;
4. Wie kann man die Matrix zu einer linearen Abbildung finden, bzw. wie kann man beweisen, dass diese Matrix die *richtige ist*
   item Satz (invmatrbas (18)): äquivalente Eigenschaften zur Invertierbarkeit;
5. Eine orthogonale Menge ist jedenfalls linear unabhängig;
6. Wie geht man den Beweis der Assoziativität der Matrix-Multiplikation an (keine Index-Details);
7. Wie zeigt man aus den (minimalen) Axiomen, dass eine lineare Abb. beliebige Linear-Kombinationen respektiert?
8. Stimmen zwei lineare Abb. auf einer Erzeugermenge üeberein, dann sind sie als Abb. gleich; (Lemma (identlinabb (15)))
9. Die inverse zu einer bijektiven linearen Abbildung ist automatisch wieder linear;
10. Der Übergang von $\boldsymbol{V}$ mit einer Basis $\mathscr{B}$ zur Koeffizientenfolge $(c_n(\mathbf{v}))$ ist linear;
11. Falls $n \neq m$, dann ist die von der $m \times n$-Matrix $\boldsymbol{A}$ induzierte Abbildung nicht bijektiv;
12. Ist $\mathbb{R}^n$ zu $\mathbb{R}^m$ isomorph, dann ist notwendigerweise $n = m$.
13. Warum sind die den Pivot-Spalten der ZST-Form entsprechenden Spalten der Matrix $\boldsymbol{A}$ eine maximal linear unabhängige Familie;
14. Dimensionsformel: Rang der Matrix + Nullität (= Dim. des Nullraumes)...
15. Wie ist der Zusammenhang zwischen Orthonormalität (unitäre Matrizen) und Darstellbarkeit eines Vektors;
16. Allgemeine Lösung eines inhomogenen linearen Glsys. $\boldsymbol{A} * \mathbf{x} = \mathbf{b}$;
17. Wo tritt die Vandermonde Matrix auf?
18. Was sind die Lagrange Interpolationspolynome?
19. Wieviele Werte eines Polynoms $p(t)$ vom Grad r braucht man, um es eindeutig bestimmen zu können.

Fertigkeiten

1. Beispiele von Vektorräumen und deren Basen;
2. Realisierung der Gausseleminination;
3. Aufstellen (und Invertieren) von Matrizen;
4. Wie kann man die Konsistenz eines inhomog. Glsys. verwenden;
5. Was ist die Rolle der Vandermonde Matrix;
6. Matrix-Multiplikation = Bilden von Linear-Kombinationen;
7. Basis f. Zeilenraum oder Spaltenraum bestimmen;
8. Interpretation von Matrix-Multiplikation *von rechts* (Bildung von Linearkombinationen von Zeilen);
9. Definition von Matrix-Multiplikation (vs. Matrix-Vektor-Mult.);
10. Liegt ein Vektor im Spaltenraum oder Zeilenraum einer Matrix;
11. Welche Regeln gelten für die Inversen/Transponierten/Transp-Konj. Matrizen;
12. Welche Verbindungen gibt es zwischen Skalarprodukten und Matrix-Multiplikation;
13. Bestimmung des Nullraumes einer linearen Abbildung (Basis, Defekt);
14. Zusammensetzbarkeit von linearen Abbildungen (sog. *Dominoprinzip*);
15. Aufstellen und Verwenden der Vandermonde-Matrix
16. Lösung von Interpolationsproblemen durch Polynomfunktionen

34 Appendix

34.1 Gruppentheorie

groupdef1 **Definition 30.** Eine Gruppe $(\mathscr{G}, \bullet, e)$ ist eine Menge mit einer Abbildung $(g, h) \mapsto g \bullet h$, welche *assoziativ* ist, d.h. folgende Gleichheit erfüllt

$$(g \bullet h) \bullet k = g \bullet (h \bullet k), \quad \forall g, h, k \in \mathscr{G}.^{109}$$

Es wird auch die Existenz eines *neutrales Element* $e \in \mathscr{G}$ gefordert, mit der Eigenschaft

$$e \bullet g = g = g \bullet e, \quad \forall g \in \mathscr{G}.$$

Ausserdem wird für jedes $g \in \mathscr{G}$ die Existenz eines *inverses Element* $h \in \mathscr{G}$ gefordert:

$$g \bullet h = e = h \bullet g.$$

Falls $g \bullet h = h \bullet g$, $\forall h, g \in \mathscr{G}$ heißt die Gruppe *kommutativ* (oder Abelsch).

Das neutrale Element ist durch seine Eigenschaft eindeutig bestimmt, denn wenn es zwei neutrale Elemente e, e' gäbe, hätte man

$$e' = e' \bullet e = e,$$

unter zweimaliger Verwendung der Eigenschaften eines neutralen Elementes.

Genauso kann man zeigen, dass auch das inverse Element eindeutig bestimmt ist, ja sogar, wenn h ein Rechts-Inverses zu g ist und k ein Links-Inverses, dann sind die beiden gleich, denn

$$h \bullet g = e = g \bullet k$$

impliziert (durch Multiplikation mit einem inversen Element):

$$h \bullet (g \bullet k) = h \bullet e = h,$$

aber andererseits gleich

$$(h \bullet g) \bullet k = e \bullet k = k.$$

Wir schreiben für das eindeutig inverse Element g^{-1}. Weiters folgt aus der Eindeutigkeit des inversen Elementes, dass gilt $(g^{-1})^{-1} = g$, denn $g \in \mathscr{G}$ erfüllt die Eigenschaften des (nach Voraussetzung) in $\mathscr{G}$ existierenden inversen Elementes zu g^{-1}.

Eng mit diesem Begriff verbundene Begriffsbildungen sind die der Untergruppe bzw. des (Gruppen-)Homomorphismus.

untergr1 **Definition 31.** Gegeben eine Gruppe $(\mathscr{G}, \bullet, e)$ bezeichnet man eine Teilmenge $\mathcal{H}$ als **Untergruppe** von $\mathscr{G}$ wenn diese Menge, ausgestattet mit der von $\mathscr{G}$ ererbten "Multiplikation" selbst wieder eine Gruppe ist.

[109] Bis hierher haben wir eine *Halbgruppe*.

Man kann/muss zeigen, dass das notwendigerweise vorhandene neutrale bzw. inverse Element einfach das schon aus $\mathscr{G}$ bekannte Element ist. Da die Assoziativität der Multiplikation in ganz $\mathscr{G}$ gilt, ist also zur Überprüfung der Untergruppeneigenschaft (> eigenes Lemma?) ausreichend, nachzuweisen, dass $\mathcal{H}$ bzgl. Multiplikation und Inversion abgeschlossen ist, d.h. (in Kurzschreibweise), dass gilt:

$$\mathcal{H} \bullet \mathcal{H} \subset \mathcal{H} \quad \text{und} \quad \mathcal{H}^{-1} := \{h^{-1} \,|\, h \in \mathcal{H}\} \subseteq \mathcal{H}. \tag{205}$$

subgrtst1

grphomdef1 **Definition 32.** Gegeben seien zwei Gruppen, $(\mathscr{G}_1, \bullet)$ und $(\mathscr{G}_2, \odot)$. Eine Abbildung $\varphi : \mathscr{G}_1 \to \mathscr{G}_2$ heisst **Gruppenhomomorphismus** wenn gilt

$$\varphi(g_1 \bullet g_2) = \varphi(g_1) \odot \varphi(g_2) \quad \forall g_1, g_2 \in \mathscr{G}. \tag{206}$$

grphomdef

Klarerweise ist jede lineare Abbildung von $\boldsymbol{V}_1$ nach $\boldsymbol{V}_2$ ein Gruppenhomomorphismus zwischen den additiven Gruppenstrukturen dieser Vektorräume (aber eben noch mehr).

Folgende abstraken Fakten sind hilfreich bei der Diskussion von Invertierbarkeit von Matrizen aufgrund der Existenz von Links-Inversen (Produkte von elementaren Matrizen):

gruppe03 **Lemma 61.** *Es sei $(H, \diamond)$ eine Halbgruppe, und*

$$G := \{g \in H \,|\, \exists g' \in H : g' \diamond g = e = g \diamond g'\}$$

sei die Menge der invertierbaren Elemente in H. Dann ist $(G, \diamond)$ eine Gruppe.

Proof. Es genügt zu zeigen, dass die Menge G abgeschlossen gegenüber Multiplikation ist, und dass jedes Element ein inverses Element hat (das ist aber praktisch Definition). Wegen $e \diamond e = e$ ist auch klar, dass $e \in G$ gilt. □

elemmatr03 **Lemma 62.** *Es sei $\boldsymbol{A}$ eine $n \times n$-Matrix, sodass ein Produkt von Elementar-Matrizen $\boldsymbol{B} = \prod_{k=1}^{s} \boldsymbol{E}_{s+1-k}$ eine Links-Inverse liefert, d.h. $\boldsymbol{B} * \boldsymbol{A} = \mathbf{Id}_n$ erfüllt. Dann ist $\boldsymbol{A}$ invertierbar und $\boldsymbol{B} = \boldsymbol{A}^{-1}$.*

Proof. Da Elementarmatrizen offenbar invertierbar sind (jeder Schritt in der Gauss-Elimination ist invertierbar, und die inverse Operation auf den Zeilen der Matrix ist durch die inverse Elementarmatrix beschreibbar) ist aufgrund von Lemma (61) gruppe03 auch $\boldsymbol{B}$ invertierbar. Da $\boldsymbol{A}$ eine Rechts-inverse von $\boldsymbol{B}$ ist, muss (wegen der Eindeutigkeit des inversen Elementes in jeder Gruppe) auch gelten $\boldsymbol{A} = \boldsymbol{B}^{-1}$, oder äquivalent: $\boldsymbol{B} = \boldsymbol{A}^{-1}$, und somit ist $\boldsymbol{A}$ invertierbar und $\boldsymbol{B}$ ist nicht nur Links-Inverse, sondern auch Rechts-Inverse zu $\boldsymbol{A}$. □

Abstrakt könnte man die Überlegung auch so zusammenfassen.

invgrel04 **Lemma 63.** *In der Situation von Lemma (61) gruppe03 gelte für ein $h \in H$: $\exists g \in G$ mit $g \diamond h = e$. Dann gilt: $h \in G$ und $h = g^{-1}$.*

Proof.

$$g \diamond h = e, \quad h^{-1} \diamond h = e$$

impliziert durch Anwenden von h^{-1} von rechts:

$$g \diamond (h \diamond h^{-1}) = h^{-1} \diamond (h \diamond h^{-1})$$

oder äquivalten: $g = h^{-1}$. □

Siehe Streaming vom 21. Mai 2014

Wir verwenden dieses abstrakte Argument im Zusammenhang mit der Berechnung des inversen Elementes

Protokolle einer früheren Vorlesung:

`http://www.univie.ac.at/NuHAG/FEICOURS/ss06/LAiNotes1.pdf`

Siehe auch homepage von Prof. A. Cap

`http://www.mat.univie.ac.at/~cap/lectnotes.html`