

Einige Notizen zur Linearen Algebra (HGFei, SS06)

Version of July 14th:
Prüfungsfragen am Ende !!
Weitere updates folgen in nächster Zeit!!

GROSSE THEMEN: Worum geht es in der Linearen Algebra???

Antwort: Um lineare Räume (auch Vektorräume genannt), bzw. um lineare Abbildungen, also um Abbildungen, die den Begriff der Linear-Kombination respektieren!!!

Was hat das mit Matrizenrechnung oder linearen Gleichungssystemen zu tun???? (Antwort bitte selber überlegen).

Konkrete Abbildungen: (orthogonale) Projektion auf eine Ebene, am besten vorzustellen anhand der orthog. Projektion im $\mathbb{R}^3$ auf die $x - y$ -Ebene (als Teilraum des $\mathbb{R}^3$ aufgefasst).

THEMA: MATRIX-VEKTOR und MATRIX-MATRIX Multiplikation

Die “Standardsichtweise” interpretiert eine Matrix $\mathbf{A}$ als eine Abbildung, die aus einem Koeffizientenvektor $\mathbf{c}$ auf $\mathbf{A} * \mathbf{c}$ abbildet, also eine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ realisiert (also im Vergleich zu den Formatangaben, die die $m \times n$ -Matrix als eine Matrix mit m Zeilen und n Spalten beschreibt, gerade die “umgekehrte” Reihenfolge). *Alternativ* kann es aber auch so gesehen werden, dass der Vektor $\mathbf{c}$ bestimmt, wieviel von dem jeweiligen Spaltenvektor in die zu bildende Linear-Kombination eingeht, denn es ist

$$\mathbf{A} * \mathbf{c} = \sum_{k=1}^n c_k \mathbf{a}^k.$$

Am deutlichsten erkennt man dieses Prinzip, wenn man als $\mathbf{c}$ einen der Einheitsvektoren $\mathbf{e}^k$ wählt, denn $\mathbf{A} * \mathbf{e}^k = \mathbf{a}^k$.

Die Matrix-Matrix-Multiplikation $\mathbf{A} * \mathbf{B}$ wird einfach so definiert, dass man die “rechte” Matrix (also $\mathbf{B}$) als Kollektion von Spaltenvektoren auffasst, und so die Matrix $\mathbf{A}$ *spaltenweise* auf die Spalten von $\mathbf{B}$ losläßt. Anders ausgedrückt, um die 3-te Spalte von $\mathbf{A} * \mathbf{B}$ zu bestimmen benötigt man nur die 3-te Spalte von $\mathbf{B}$ (und die ganz Matrix $\mathbf{A}$).

FORMAL: Schreiben wir $\mathbf{C}$ für die Produkt-Matrix, dann sind die individuellen Eintragungen von $\mathbf{C} = (c_{j,l})$ durch folgende Formel gegeben:

Matrixmult

$$(1) \quad c_{j,l} = \sum_k a_{j,k} b_{k,l},$$

wobei die Matrix Multiplikation prinzipiell nur möglich ist, wenn die beiden Matrizen gemäß dem (von mir als) **Domino-Prinzip** (bezeichneten Format-Test) entsprechen müssen, und die Summe genausoviele Summanden hat wie die wegfallende Formatgröße.

Das Produkt einer 3×5 Matrix mit einer 5×7 Matrix ist eine 3×7 Matrix $\mathbf{C}$, wobei die Multiplikationsformel oben eben 5 Multiplikationen pro Eintragung benötigt.¹

¹Die Gesamtzahl an Multiplikationen sind also für diese Beispiel $3 \cdot 7 \cdot 5$ Multiplikationen. Dementsprechend “kostet” die Matrix Multiplikation von zwei quadratischen $n \times n$ Matrizen insgesamt

Praktisch kann man es sich das so vorstellen: Um eine bestimmte Eintragung in der Zeile j und der Spalte k zu bestimmen, muss man nur die Elemente der j -ten Zeile des ersten Faktors paarweise mit den Elementen der k -ten Spalte des zweiten Faktors multiplizieren und dann aufaddieren. Die Summe (mittels einer “gebundenen Variablen”) hat genau n Terme (das ist die “gemeinsame” Dimension der beiden multiplizierten Matrizen).

Aus dieser Beobachtung lassen sich weitere Aussagen herleiten. Die Spalten des Matrix Produktes $\mathbf{C} := \mathbf{A} * \mathbf{B}$ sind Linearkombinationen der Spalten von $\mathbf{A}$ (die Koeffizienten für die k -te Spalte von $\mathbf{C}$ kommen von der k -ten Spalte von $\mathbf{B}$), aber genauso kann man $\mathbf{B}$ als eine Kollektion von Zeilenvektoren ansehen, und dann zeigt sich, dass die Zeilen von $\mathbf{A}$ angeben, welche Linearkombination von Zeilen von $\mathbf{B}$ genommen werden. Insbesondere ist der Spaltenraum von $\mathbf{A} * \mathbf{B}$ immer im Spaltenraum von $\mathbf{A}$ enthalten (Gleichheit falls $\mathbf{B}$ invertierbar ist, wie wir später sehen werden), und analog der Zeilenraum von $\mathbf{A} * \mathbf{B}$ im Zeilenraum von $\mathbf{B}$ enthalten.

Wir merken auch noch an, dass zwei Matrizen (gleicher Dimension) definitionsgemäß gleich sind genau dann wenn sie an denselben Stellen die gleichen Werte haben. Es ist auch leicht zu zeigen, dass zwei Matrizen $\mathbf{A}$ und $\mathbf{B}$ genau dann gleich sind, wenn $\mathbf{A} * \mathbf{x} = \mathbf{B} * \mathbf{x}$ für alle Vektoren $\mathbf{x}$ gilt.

Man kann Matrizen auch addieren (wieder *komponentenweise*) bzw. mit Skalaren multiplizieren (alle Komponenten werden mit derselben Zahl multipliziert). Auf diese Weise werden für jedes Paar natürlicher Zahlen $m, n \in \mathbb{N}$ die $m \times n$ -Matrizen selbst zu einem “Vektorraum” (siehe später), und die “Einheitsmatrizen” ($E^{(n,k)}$, die lauter Nullkoordinaten haben und einen Einser an einer Stelle), sind ein Erzeugendensystem.

Wir können aber einen umgekehrten Standpunkt einnehmen. Wenn mir jemand einen Spaltenvektor im $\mathbb{R}^m$ gibt, dann kann ich den als Koeffizientenvektor für eine Linearkombination von m Vektoren (in irgendeinem Vektorraum, z.B. in $\mathbb{R}^m$) verwenden. Eine $m \times n$ Matrix kann daher als ein rechteckiges Zahlenschema angesehen werden, das alle nötigen Koeffizienten enthält, um aus diesen m Vektoren n neue Vektoren zu machen! Diese Sichtweise hat den Vorteil, dass nun die “richtige Reihenfolge” herrscht: eine $m \times n$ -Matrix dient dazu, aus m Vektoren, n neue Vektoren zu machen, wobei die Eintragungen der j -ten Spalte angeben, welche Koeffizienten man für das Zusammensetzen des j -ten (neuen) Vektors benötigt.

Was nun interessant ist, ist die Frage: Nimmt man diese n neuen Vektoren her, und bildet daraus p weitere Linearkombinationen der (neuen) Vektoren, so kann man die dazu verwendeten Koeffizienten wieder in eine Matrix schreiben, die demgemäß das Format einer $n \times p$ Matrix bilden. Versucht man nun, die (ganz neuen) k Vektoren, die auf diese Weise entstehen, wieder als Linear-Kombination der ursprünglichen m Vektoren darzustellen, so ist offnebar (wenn’s überhaupt geht) eine Matrix $\mathbf{C}$ vom Format $m \times k$ benötigt. Eine wichtige AUSSAGE ist nun: Ja, das geht, und diese Matrix ist genau die Produktmatrix $\mathbf{A} * \mathbf{B}$.

Die Details dazu werden in der Vorlesung gebracht, wobei in der Vorlesung vor allem die “ganz allgemeine” Sichtweise zur Geltung kommen soll, während im Proseminar bzw.

n^3 Multiplikationen, d.h. die Verdoppelung der Größe von Matrizen (Übergang von n auf $2n$) ergibt eine Verachtfachung $8 = 2^3$ des Rechenaufwandes!.

im Einzelstudium bzw. in der von mir empfohlenen Arbeit in kleinen Gruppen vielleicht konkrete Beispiele am besten helfen, das Prinzip der Matrix-Matrix bzw. Matrix-Vektor Multiplikation zu verstehen.

Natürlich sind wir mit diesem (vermutlich als *steil empfundenen*) Einstieg der Begriff einer “inversen Matrix” schon sehr nahe: Eine Matrix $\mathbf{B}$ heißt **inverse Matrix** zu einer $n \times n$ Matrix ² wenn $\mathbf{B} * \mathbf{A} = I = I_n = \mathbf{A} * \mathbf{B}$ gilt, wobei I_n bzw. kurz I (wenn das Format klar ist) die **Einheitsmatrix** der Größe $n \times n$ ist, deren Spalten als ganz einfach die Einheitsvektoren $\mathbf{e}^k$, $1 \leq k \leq n$ sind.

Eine der Dinge ist die Überlegung, welche Eigenschaften die Abbildung $\mathbf{x} \mapsto \mathbf{A} * \mathbf{x}$ hat, die ja von $\mathbb{R}^n$ nach $\mathbb{R}^m$ geht. Natürlich (man verifiziere das jede[r] für sich gilt $\mathbf{A} * (\mathbf{x} + \mathbf{x}') = \mathbf{A} * \mathbf{x} + \mathbf{A} * \mathbf{x}'$ und weiteres $\mathbf{A} * (\lambda \mathbf{x}) = \lambda \mathbf{A} * \mathbf{x}$ für jeden Skalar λ . Es folgt daraus einfach (mit Hilfe der Methode der vollständigen Induktion) dass jede endliche Linearkombination (beliebiger Länge) von der Abbildung respektiert wird, d. h. es gilt

$$\mathbf{A} * \left(\sum_i c_i \mathbf{x}_i \right) = \sum_i c_i (\mathbf{A} * \mathbf{x}_i).$$

Daraus folgt natürlich auch: Es $\mathbf{v}^i$ ein Erzeugendensystem für $\mathbb{R}^n$, und kennt man die Bilder dieser Vektoren unter $\mathbf{A}$, so kann man daraus auch das Bild eines beliebigen Vektors in $\mathbb{R}^n$ bestimmen: Beweis. Bezeichnen wir die Bilder der $\mathbf{v}^i$ mit $\mathbf{u}^i$, also $\mathbf{u}^i := \mathbf{A} * \mathbf{v}^i$ so ist für ein beliebiges Element $\mathbf{z}$ jedenfalls eine Darstellung von in der Form $\mathbf{z} = \sum_i c_i \mathbf{v}^i$ möglich (nach Definition des Begriffes Erzeugendensystem), also muss wegen der Respektierung von Linear-Kombinationen unter der Wirkung einer Matrix gelten:

$$\mathbf{A} * \left(\sum_i c_i \mathbf{v}^i \right) = \sum_i c_i (\mathbf{A} * \mathbf{v}^i) = \sum_i c_i \mathbf{u}^i.$$

Man stellt also den Bildvektor als Linearkombination der $\mathbf{u}^i$ dar.

Es ist eine schöne Aufgabe, die Bahn eines Vektors von Länge 1 mit beliebiger Richtung unter einer allgemeinen 2×2 -Matrix festzustellen! (was kommt da heraus? und welche Spekulationen schließen sich daran an, wenn man aus der geometrischen Situation Beobachtungen gewonnen hat?)

Weiterer Begriff:

Hat man einfach eine Menge M von Vektoren (im $\mathbb{R}^m$) dann gibt es einen eindeutig bestimmten kleinsten Raum, oft mit $[M]$ bezeichnet, der bzgl. Addition und skalarer Multiplikation abgeschlossen ist. Dieser Raum, oft auch als lineares Erzeugnis von M bezeichnet, besteht *genau* aus den Linear-Kombinationen der Elemente von M (mit beliebig vielen Elementen und beliebigen Skalaren).

KURZE ZWISCHENZUSAMMENFASSUNG:

Die Matrix Vektor-Multiplikation (eine $m \times n$ -Matrix wird auf einen Spaltenvektor der Länge n angewendet, und ergibt einen Vektor im $\mathbb{R}^m$) ist so gestaltet, dass das Ergebnis eine Linearkombination der Spalten von $\mathbf{A}$ ist. Der Bildraum von $\mathbf{x} \mapsto \mathbf{A} * \mathbf{x}$ ist also genau der Spaltenraum

²Warum diskutiert man nicht den rechteckigen Fall?

1. MATRIXMULTIPLIKATION UND LINEARE GLEICHUNGSSYSTEME

Matrixmultiplikation ist assoziativ.

Daraus kann man leicht folgern, dass die Existenz von Links- und Rechtsinversen Matrizen dazu führt, dass es eine Matrix $\mathbf{B}$ gibt, mit $\mathbf{B} * \mathbf{A} = I_n = \mathbf{A} * \mathbf{B}$. Man schreibt für $\mathbf{B}$ $\mathbf{A}^{-1}$. Es gilt $(\mathbf{A}^n)^{-1} = (\mathbf{A}^{-1})^n$, für $n \in \mathbb{N}$.

DEFINITION: Lineare Gleichungssysteme sind “genau solche” die sich durch Matrix-Multiplikation in der kompakten Form $\mathbf{A} * \mathbf{x} = \mathbf{b}$ schreiben lassen. Wenn $\mathbf{b} = \mathbf{0}$ gilt, dann heißt das Gleichungssystem *homogen*, ansonsten spricht man von einem *inhomogenen Gleichungssystem*.

Die Lösungsmenge eines homogenen linearen Gleichungssystems ist ein (linearer) Teilraum von $\mathbb{R}^n$, d.h. es ist eine Teilmenge von $\mathbb{R}^n$, die bzgl. Addition und skalarer Multiplikation abgeschlossen ist, also auch gegenüber der Bildung von Linearkombinationen.

Es gibt drei Regeln, die man zur Lösung von linearen Gleichungssystemen mit m Gleichungen in n Variablen verwenden kann:

- (1) Multiplizieren einer Gleichung mit einer Zahl ungleich Null;
- (2) Addieren eines Vielfachen einer (festen) Zeile von anderen Zeilen;
- (3) Vertauschen von zwei Zeilen.

Alle drei Schritte haben zwei Eigenschaften gemeinsam: Sie sind können wieder rückgängig gemacht werden, somit ist im Laufe der Umformungen eines lin. Gleichungssystems im Zuge der Gauss Elimination **lösungsäquivalent**, und weiters lassen sich die Schritte durch Multiplikation von links mit Hilfe einer *Elementarmatrix* darstellen. Die entsprechende Elementarmatrix entsteht, indem man die entsprechenden Schritte auf die Einheitsmatrix I (Einsen in der Hauptdiagonale) von Format $n \times n$ anwendet. Die Möglichkeit, die elem. Transformationen wieder rückgängig zu machen, kann rückübersetzt werden in die Eigenschaft der Elementarmatrizen, invertierbar zu sein!!

WEITERE BEGRIFFE: 4.4.2006: Vorbereitung f. 6.4.2006 zur Erinnerung:

Definition 1. Eine (quadratische, d.h. eine $n \times n$ Matrix $\mathbf{A}$ heisst invertierbar, wenn es dazu eine andere (eindeutig bestimmte) Matrix $\mathbf{B}$, mit $\mathbf{A} * \mathbf{B} = I = \mathbf{B} * \mathbf{A}$ gilt. Man schreibt dann naheliegenderweise $\mathbf{A}^{-1}$ für $\mathbf{B}$.

Übung: $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ aufgrund der Symmetrie der Definition.

Übung: $(\mathbf{B} * \mathbf{A})^{-1} = \mathbf{A}^{-1} * \mathbf{B}^{-1}$

Es folgt unmittelbar aus der Definition dass aufgrund der Gleichung $\mathbf{A} * \mathbf{A}^{-1} = I$ die k -te Spalte $\mathbf{b}^k$ von $\mathbf{B} = \mathbf{A}^{-1}$ genau die Lösungen der Gleichung des Gleichungssystems $\mathbf{A} * \mathbf{b}^k = \mathbf{e}^k$ ist, für $1 \leq k \leq n$.

Lemma 1. Eine Matrix $\mathbf{A}$ ist genau dann invertierbar, wenn $\mathbf{c} \mapsto \mathbf{A} * \mathbf{c}$ ein bijektive Abbildung ist. Die Umkehrabbildung ist dann genau durch die Matrix-Multiplikation mit $\mathbf{A}^{-1}$ gegeben.

Proof. Wenn die Abbildung bijektiv ist, dann sind auch alle Einheitsvektoren im Bildbereich der Abbildung $\mathbf{c} \mapsto \mathbf{A} * \mathbf{c}$, also gibt es für jedes k eine eindeutig bestimmte Lösung des Problems $\mathbf{A} * \mathbf{b}^k = \mathbf{e}^k$, also gibt es eine inverse Matrix. Umgekehrt, wenn es eine inverse Matrix, so ist natürlich zu jeder Gleichung $\mathbf{A} * \mathbf{x} = \mathbf{y}$ die Lösung $\mathbf{x} = \mathbf{A}^{-1} \mathbf{y}$ die eindeutig bestimmte Lösung (Details fehlen hier noch). $\square$

Aussagen, die in der Vorlesung am 6.4. vorgetragen wurden (samt den entsprechenden Definitionen):

Theorem 1. *Der Nullraum (oder Kern) einer Matrix, d.h. $N(\mathbf{A}) := \{\mathbf{x}, \mathbf{A} * \mathbf{x} = \mathbf{0}\}$ ist ein (nichtleerer) linearer Teilraum von $\mathbb{R}^n$, d.h. eine Menge die gegenüber Addition und skalarer Multiplikation abgeschlossen ist, also auch gegenüber Linearkombinationsbildungen.*

Theorem 2. *Die allgemeine Lösungsmenge eines linearen Gleichungssystems (also eines Gleichungssystems das in der Form $\mathbf{A} * \mathbf{x} = \mathbf{b}$ geschrieben werden kann) ist dadurch beschreibbar, dass man zu einer speziellen Lösung alle Lösungen des entsprechenden homogenen Gleichungssystems addiert (später werden wir von einem affinen Teilraum von $\mathbb{R}^n$ sprechen).*

Kommentare zur Vorlesung am 24.04.2006:

Wir haben gesehen, dass die Gauss-Elimination im “gluecklichen Fall” (n Gleichungen mit n Variablen, und so dass man mittels Gauss-Jordan Schritten zur Einheitsmatrix kommt) eine Matrix $\mathbf{B}$ erzeugt wird, mit $\mathbf{B} * \mathbf{A} = I_n$. Somit wird aus dem linearen Problem (Glsys.) $\mathbf{A} * \mathbf{x} = \mathbf{b}$ ein *lösungsäquivalents* System der Form

$$\mathbf{x} = (\mathbf{B} * \mathbf{A}) * \mathbf{x} = \mathbf{B} * (\mathbf{A} * \mathbf{x}) = \mathbf{B} * \mathbf{b},$$

welches ja offensichtlich für *jede* rechte Seite $\mathbf{b}$ eine *eindeutig bestimmte* (weil ja “berechenbare”) Lösung $\mathbf{x} = \mathbf{B} * \mathbf{x}$ liefert. Insbesondere besagt das, dass in diesem Fall die Abbildung $T_A : \mathbf{x} \mapsto \mathbf{A} * \mathbf{x}$ ein *bijektive* Abbildung ist.

Wir wollen nun den Zusammenhang zwischen Matrizen und Matrix-Multiplikation noch ein wenig genauer studieren, um die Existenz und Nützlichkeit von inversen Matrizen besser zu verstehen (hoffe ich jedenfalls). Dazu eine kleine Definition:

Definition 2. Eine Abbildung T von $\mathbb{R}^n$ nach $\mathbb{R}^m$ heisst *linear* wenn sie mit Addition und skalarer Multiplikation verträglich ist, d.h. wenn $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ und $T(\lambda \mathbf{x}) = \lambda T(\mathbf{x})$ gilt, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$.

Harmlose Übung: ein Abbildung ist genau dann linear, wenn sie mit Linearkombinationen verträglich ist (wie üblich, Induktionsbeweis nach der Anzahl der Terme):

lin-equ

$$(2) \quad T\left(\sum_{k=1}^l \lambda_k \mathbf{z}_k\right) = \sum_{k=1}^l \lambda_k T(\mathbf{z}_k).$$

Theorem 3. Jede lineare Abbildung T von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ist durch eine Matrix gegeben. Ist T (“abstrakt”) gegeben, so kann man die zugehörige Matrix $\mathbf{A}$ gewinnen, indem man $\mathbf{a}^k = T(\mathbf{e}^k)$ setzt.

Proof. Wir wissen ja schon, dass die Abbildungen $T_A : \mathbf{x} \mapsto \mathbf{A} * \mathbf{x}$ diese Eigenschaften haben! Umgekehrt setzt man einfach eine Matrix $\mathbf{A}$ zusammen, indem man die Bilder der Einheitsvektoren als Spalten einer (neue Matrix) zusammensetzt. Diese hat das richtige Format: $m \times n$, und offenbar gilt $T_A \mathbf{e}^k = \mathbf{A} * \mathbf{e}^k = T(\mathbf{e}^k)$, d.h. die beiden (!) linearen Abbildungen, das gegebene T und das neu gebaute T_A stimmen auf den Einheitsvektoren überein. Benützt man nun die Linearitätsformel [2], ^{lin-equ} kommt man zu:

$$\mathbf{A} * \mathbf{x} = \mathbf{A} \left(\sum_{k=1}^l x_k \mathbf{e}^k \right) = \sum_{k=1}^l x_k \mathbf{A} * \mathbf{e}^k = \sum_{k=1}^l x_k T(\mathbf{e}^k) = T \sum_{k=1}^l x_k \mathbf{e}^k = T(\mathbf{x}).$$

□

Ein kleiner Hilfssatz ist nützlich für den kommenden (wichtigen) Satz:

Lemma 2. Die Umkehrabbildung einer linearen und invertierbaren Abbildung T ist linear.

Proof. Bedenke, dass eine invertierbare Abbildung T eine inverse Abbildung T^{-1} hat, und dass die beiden Aussagen $T\mathbf{x} = \mathbf{y}$ und $T^{-1}\mathbf{y} = \mathbf{x}$ vollkommen gleichwertig sind.

Zeige z.B. die Verträglichkeit mit der Addition (bzw. skal. Mult.):

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

□

Was nun folgt, ist das wichtige Bindeglied zwischen der Theorie der linearen Abbildungen und dem Lösen von linearen Gleichungssystemen bzw. der Matrizenrechnung:

Theorem 4. *i) Die Komposition von zwei linearen Abbildungen ist wieder linear, und die zugehörigen Matrizen entstehen durch Matrix-Multiplikation.*

ii) Eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$ ist genau dann invertierbar wenn die zugehörige Matrix als Matrix invertierbar ist, und die Inversionsbegriffe für (lineare) Abbildungen bzw. für Matrizen sind verträglich.

Proof. Zuerst einmal ist klar, dass die Komposition von linearen Abbildungen wieder linear ist (einfache Übung!). Hat man also zwei Abbildungen, die (nach dem Domino-Prinzip f.d. entsprechenden Dimension) zusammensetzbar sind, so ist $T_{\mathbf{B}} \circ T_{\mathbf{A}}$ natürlich eine lineare Abbildung, und $\mathbf{C} := \mathbf{B} * \mathbf{A}$ ein Matrix von passendem Format, es geht also nur darum zu zeigen, dass die von $\mathbf{C}$ vermittelte lineare Abbildung (systematischer Name $T_{\mathbf{C}}$ gerade $T_{\mathbf{B}} \circ T_{\mathbf{A}}$ ist. Dazu muss man sich aber offenbar die Bilder der Einheitsvektoren unter $T_{\mathbf{B}} \circ T_{\mathbf{A}}$ ansehen:

$$(T_{\mathbf{B}} \circ T_{\mathbf{A}})(\mathbf{e}^k) = T_{\mathbf{B}}(T_{\mathbf{A}}(\mathbf{e}^k)) = T_{\mathbf{B}}(\mathbf{a}^k) = (\mathbf{B} * \mathbf{a}^k) = (\mathbf{c}^k)$$

wobei man zuerst verwendet, wie die zusammengesetzte Abbildung wirkt, dann die Tatsache dass die Spalten von $\mathbf{A}$ genau die Bilder der Einheitsvektoren sind, und zuletzt, dass die Matrix-Multiplikation von $\mathbf{B}$ (von links) spaltenweise auf die Spalten von $\mathbf{A}$ wirkt (um zu den Spalten der Produktmatrix zu kommen, die wir oben $\mathbf{C}$ genannt haben).

Der zweite Teil folgt nun daraus, dass die zu $T = T_{\mathbf{A}}$ inverse Abbildung T^{-1} natürlich dadurch charakterisiert ist, dass man durch zusammensetzen die "identische" Abbildung $\mathbf{x} \mapsto \mathbf{x} = I_n * \mathbf{x}$ bekommt. $\square$

Bisher wissen wir: Invertierbare lineare Abbildungen gehören zu invertierbaren Matrizen, und - wenn man Glück hat, kann man die Gauss-Elimination verwenden, um die inverse Matrix zu bestimmen (de facto sagt die Theorie dass es immer geht!).

Wir haben aber noch nicht ausreichende Information darüber, dass die mittels Gauss-Elimination gefundene LINKS-Inverse Matrix auch eine Rechtsinverse ist (oder dass es eine Rechts-Inverse gibt), bzw. die Behauptung gezeigt (die richtig ist), dass die reine Existenz der inversen Matrix auch schon garantiert, dass die Inverse Matrix mit dieser konkreten Methode (warum nur oder genau diese drei Schritte, etc.) berechnet werden kann.

Ende der Stunde am 5.5.2006: Der erste "konkrete/abstrakte" Vektorraum (der nicht nur aus n -Tupeln im $\mathbb{R}^n$ besteht):

Die Menge aller *Polynomfunktionen* $p(z)$ auf irgendeinem (nicht-degenerierten) Teilgebiet des $\mathbb{C}$, z.B. auf $\mathbb{R}$, oder auf der ganzen komplexen Ebene, oder auf $I = [a, b]$, z.B. auf $[0, 1]$

Woche: 8/11.Mai

Definition: Allgemeine Vektorräume, Erzeugendensystem, Basis, linear unabhängige Mengen

Exercise: Das Bild einer linear abh. Menge unter einer linearen Abbildung ist wieder linear abhängig (natürlich mit denselben Koeffizienten).

11.Mai:

Feststellung, dass die Anzahl der Basiselemente in einem Vektorraum mit endlichem Erzeugendensystem *unabhängig* von der konkreten Basis ist: also ist $\dim(V)$ ein wohldefinierter Begriff.

Klarerweise ist $\dim(\mathbb{R}^n) = n$.

Korollar: Hat man k lin. unabh. Elemente in einem n -dimensionalen Vektorraum, so gilt $k \leq n$ (anders ausgedrückt: hat man mehr als $\dim(V)$ Vektorraum, so sind sie notwendigerweise linear abhängig).

Zeilenrang = Spaltenrang: der **Rang** einer Matrix **A**.

Die Polynomfunktionen vom (maximalen) Grad r , genannt $\mathcal{P}_r(\mathbb{R})$, bilden einen $r + 1$ -dim. Vektorraum. Die "Standard-Basis" ist $\mathcal{M}_r(\mathbb{R}) = \{1, x, x^2, \dots, x^r\}$. Die Menge $\mathcal{P}(\mathbb{R}) = \bigcup_{r \geq 0} \mathcal{P}_r(\mathbb{R})$ ist ein Vektorraum, der nicht endlichdim. ist, d.h. keine endliche, lin. unabh. Menge erzeugt einen Teilraum, der nicht alle Polynomfunktionen enthält.

Der Taylor'sche Lehrsatz ist eine Form der Rekonstruktion der Koeffizienten eines Polynoms mit Hilfe der Ableitungsfunktion.

Die Verschiebung $p(x) \mapsto p(x - 1)$ (Rechtshift des Graphen von $p(x)$) ist eine invertierbare Abbildung aus der Menge $\mathcal{P}_r(\mathbb{R})$ für jedes $r \geq 0$. Diese linearen Abbildungen haben eine Matrixdarstellung bzgl. der Basis $\mathcal{M}_r(\mathbb{R})$. Die Potenzen dieser Matrix stellen die Verschiebung um $k \in \mathbb{N}$ bzw. die negativen Potenzen $-k \in \mathbb{N}$.

Die Ableitungsabbildung $p(x) \mapsto p'(x)$ ist nilpotent, d.h. eine Potenz davon ist gleich Null. Konkret: die r -te Ableitung einer jeden Polynomfunktion $p(x) \in \mathcal{P}_r(\mathbb{R})$ ist identisch gleich Null.

Eine bijektive und lineare Abbildung T von einem Vektorraum in einen anderen Vektorraum wird ein Isomorphismus genannt. Kompositionen (und inverse Abbildungen) zu Isomorphismen sind wieder Isomorphismen (wir haben früher schon gezeigt, dass die Umkehrabb. einer linearen Abb. wieder linear ist, d.h. wenn eine $n \times n$ -Matrix eine bijektive Abbildung erzeugt, dann ist auch die Umkehrabbildung durch eine Matrix dargestellt, und natürlich durch die (eindeutig bestimmte) inverse Matrix.

SATZ: Jede Basis mit n Elementen eines Vektorraums V definiert einen Isomorphismus zwischen V und $\mathbb{R}^n$, und umgekehrt.

Andere Beispiele: $C(I)$, der Vektorraum aller stetigen reellwertigen Funktionen auf einem Intervall I , z.B. $I = [a, b]$ oder speziell $I = [0, 1]$. Auch dieser Raum ist "unendlichdimensional".

Die Dimensionsfunktion ist monoton: $V_1 \subseteq V_2$ impliziert $\dim(V_1) \leq \dim(V_2)$, mit Gleichheit der Dimensionen nur für den Fall der Gleichheit von V_1 und V_2 .

noch zu diskutieren: $\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2)$ genau dann wenn die Summe eine direkte ist, d.h. genau dann wenn $V_1 \cap V_2 = \{0\}$.

Vorschau: 15.Mai 2006:

Alternative Basen für $\mathcal{P}_r(\mathbb{R})$: z.B. $\{1, (x-1), \dots, (x-1)^r\}$

Weitere passende Themen wären jetzt: Matrixdarstellung bzgl. *beliebiger* Basen, Basiswechsel, auch als Vorbereitung für die Diagonalisierungsfrage.

Linearkombinationen von Linearkombinationen sind wieder Linearkombinationen der ursprünglichen Elemente, wobei das auch durch Matrix-Multiplikation ausgedrückt werden kann >>> details missing so far!

DIESE DINGE sollen etwas später erst wieder aufgegriffen werden, dafür sollen nun als nächstes die geomtrisch anschaulichen Begriffe von Winkel, Orthonormalsystemen etc., beide basierend auf dem Begriff des Skalarproduktes von Vektoren (im $\mathbb{R}^n$) eingeführt werden. Dieses Hilfsmittel wird uns einerseits die *Orthonormalbasen* als die idealen Basen ausweisen, für die das Problem der Koeffizientenbestimmung ganz einfach ist, andererseits auch einen neuen Blick auf das Gauss'schen Eliminationsverfahren eröffnen!

Definition 3. Sei V ein Vektorraum. Eine bilineare Abbildung von $V \times V$ (als eine Abb., die bei Festhalten des einen Arguments eine lineare Funktion in dem anderen Argument ist) nach dem zugrundeliegenden Körper, also (für uns) $\mathbb{R}$ oder $\mathbb{C}$, üblicherweise geschrieben als $(v, w) \mapsto \langle v, w \rangle$ heisst Skalarprodukt, wenn es positiv definit, wenn es die Eigenschaft hat, dass $\langle v, v \rangle \geq 0$ für alle $v \in V$ gilt, und außerdem: $\langle v, v \rangle = 0$ genau dann wenn $v = 0$ gilt (bzw. einfach "nur dann wenn").

Mit jedem Skalarprodukt ist auch eine Längenmessung von "Vektoren" möglich: Man definierte die eines Vektors $v \in V$ als

$$\|v\| := \|\mathbf{x}\|_2 := \sqrt{\langle v, v \rangle}.$$

Auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist das (übliche, kanonische, Euklidische) Skalarproduct auf folgende Weise gegeben ³

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{k=1}^n x_k \overline{y_k}.$$

Ich gebe gleich die allgemeine Definition für $\mathbb{C}^n$, denn es ist leicht die Konjugation (Querbalken über den Koordinaten von $\mathbf{y}$ zu ignorieren, wenn $\mathbf{y} \in \mathbb{R}^n$ ist, also alle Koordinaten reell sind), als umgekehrt später noch komplizierte Zusatzregeln für den komplexen Fall dazuzulernen!

Die dem Standard-Skalarprodukt zugeordnete Längenmessung (in Kürze werden wir dies eine Norm nennen), also $\|\mathbf{x}\| := \|\mathbf{x}\|_2 := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ wird auch oft als *Euklidische Länge* betrachtet.

³Wir werden später noch eine ganze Reihe von anderen Skalarprodukten kennenlernen!

Proposition 1. *Hat man einen beliebigen endlichdimensionalen Vektorraum, und eine Basis $\mathcal{B} = (b^1, \dots, b^n)$ in V , so kann man darauf leicht ein Skalarprodukt definieren, indem man das “gewöhnliche, Euklidische Skalarprodukt von den $\mathcal{B}$ -Koordinaten von zwei Vektoren v_1, v_2 nimmt.*

Es ist tatsächlich ganz leicht nachzuprüfen, dass auf diese Weise ein Skalarprodukt definiert wird (!! alle Eigenschaften bitte selbst durchgehen!). Insbesondere ist

$$(\mathbf{x}, \mathbf{y}) \mapsto \langle \mathbf{B} * \mathbf{x}, \mathbf{B} * \mathbf{y} \rangle$$

für jede invertierbare(!) $n \times n$ -Matrix ein neues Skalarprodukt ⁴

Definition 4. Eine Familie von Vektoren in einem Vektorraum mit innerem Produkt heißt *Orthogonalsystem* wenn je zwei verschiedene Vektoren aus dieser Familie zueinander orthogonal sind (und normalerweise setzt man voraus dass keiner der Vektoren der Nullvektor in V ist).

Eine orthogonale Familie heißt *Orthonormalsystem* wenn die einzelnen Vektoren zusätzlich noch normiert sind, d.h. wenn sie alle Länge 1 (im Sinne der Norm die zu dem Skalarprodukt gehört) haben. Eine Basis, die ein Orthonormalsystem bildet, wird demgemäß als *Orthonormalbasis* bezeichnet

Definition 5. Eine Matrix, deren Spalten ein *Orthonormalsystem* bilden, wird als *unitär* bezeichnet ⁵, bzw. im Reellen eine *orthogonale* Matrix genannt. ⁶ genannt.

Die unitären (orthog.) Matrizen sind dadurch gekennzeichnet, dass ihre inverse Matrix von der Form $U = U'$ ist, wobei $U' = \text{conj}(U^t)$ definiert wird, d.h. die inverse Matrix zu einer solchen Matrix U ist die Matrix V , deren Eintragungen durch $v_{k,l} = \text{conj}(u_{l,k})$ gegeben ist, für $1 \leq k, l \leq n$.

Nun zeigen wir, dass Orthogonalsysteme linear unabhängig sind, und dass vor allem die Berechnung von Koeffizienten bzgl. eines Orthonormalsystems geringen Aufwand bedeutet.

Theorem 5. *Die unitären $n \times n$ - Matrizen bilden eine Gruppe. Die reellen (orthogonalen) Matrizen bilden eine Untergruppe derselben. Das inverse Element einer jeden unitären Matrix ist.*

Man kann diesen Satz leicht dazu verwenden um festzustellen, dass es zu je zwei Orthonormalbasen im $\mathbb{R}^n$ eine unitäre $n \times n$ -Matrix $\mathbf{U}$ gibt, die die Basisvektoren ineinander überführt.

Beweis(!): Schreiben wir die beiden Basen als Spalten in eine Matrix, so sind diese Matrizen, nennen wir sie $\mathbf{U}$ und $\mathbf{V}$ jeweils unitär (nach Def.). Offenbar ist die Matrix $\mathbf{V}$ gleichzeitig das Bild der Standard-Einheitsvektoren unter der unitären Transformation

⁴wir werden später noch sehen, dass alle möglichen Skalarprodukte auf diese Weise interpretiert werden können, für passendes $\mathbf{B}$, wobei das $\mathbf{B}$ nicht eindeutig bestimmt ist.

⁵Dieser Sprachgebrauch wird vor allem für komplexe Matrizen verwendet, ich werde in der Folge vor allem von unitären Matrizen sprechen, auch wenn sie reell sind.

⁶Die Begründung für diese scheinbar “wirre” Terminologie besteht darin, dass diese Matrizen orthogonale Familien wieder in orthogonale Familien überführen, also sollte diese Matrizen richtigerweise “orthogonalitätserhaltende Matrizen” genannt werden, und “orthogonale Matrizen” ist eine Kurzform davon.

$\mathbf{x} \mapsto \mathbf{V} * \mathbf{x}$. Analoges gilt für $\mathbf{U}$, also ist $\mathbf{U}' = \mathbf{U}^{-1}$ die Matrix, die die Spalten von $\mathbf{U}$ in die Einheitsbasis zurückführt. Somit ist die Zusammensetzung von $\mathbf{U}'$, gefolgt (im Sinne von Anwendung) von $\mathbf{V}$, also die Matrix $\mathbf{W} = \mathbf{V} * \mathbf{U}'$ (die ja selbst wieder eine unitäre Matrix ist, siehe oben) eine (und im Endeffekt die einzige) unitäre Matrix welche die Spalten von $\mathbf{U}$ genau auf die entsprechenden Spalten von $\mathbf{V}$ abbildet!

Das ganze nochmals als eigener Satz formuliert:

Theorem 6. *Gegeben eine beliebige ONB des $\mathbb{C}^n$ bzw. $\mathbb{R}^n$, und eine unitäre (bzw. orthog.) $n \times n$ -Matrix, dann ist das Bild wieder eine ONB von $\mathbb{C}^n$ bzw. $\mathbb{R}^n$.*

*Umgekehrt kann man sagen: Sind $\mathcal{B}$ und $\mathcal{C}$ zwei ONBs für $\mathbb{C}^n$ bzw. $\mathbb{R}^n$, so gibt es eine unitäre Matrix $\mathbf{U}$ sodass $\mathbf{c}^k = \mathbf{U} * \mathbf{b}^k$, für $1 \leq k \leq n$ gilt.*

Proof. Argument f.d. zweiten Teil: Man nehme $\mathbf{U} = \mathbf{C} * \mathbf{B}^{-1}$, wobei $\mathbf{B}$ und $\mathbf{C}$ die entsprechenden Matrizen sind, deren Spalten die beiden ONBs ausmachen. Die Aussage kann aber auch (mit etwas allgemeineren Hilfsmitteln) ohne zu-Hilfenahme dieser Matrixdarstellung bewiesen werden (aber noch nicht hier). $\square$

Theorem 7. *Die unitären Matrizen bewahren das Skalarprodukt, d.h. für jede unitäre Matrix $\mathbf{U}$ gilt:*

$$\langle \mathbf{U} * \mathbf{x}, \mathbf{U} * \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Insbesondere bewahren unitäre (orthog.) Matrizen sowohl Längen ($\mathbf{y} = \mathbf{x}$ in der obigen Gleichung) und in der Folge (siehe später) auch Winkel zwischen Vektoren.

Zusammenfassung:

Beobachtung (1): ganz vom Anfang: das Bilden von Linearkombinationen von Vektoren kann auf folgende Weise realisiert werden: Schreibt man die (reellen oder komplexen) Vektor als Spalten in eine Matrix $\mathbf{A}$, so kann man eine Linearkombination dieser Vektoren (üblicherweise gedacht als Elemente von $\mathbb{C}^m$, mit Koeffizienten $c_1, \dots, c_n$ dadurch bekommen, indem man $\mathbf{A} * \mathbf{c}$ bildet.

Man kann aber auch eine Matrix von Spaltenvektoren zu einer Matrix von Zeilenvektoren umbauen, indem man sie transponiert: $\mathbf{B} = \mathbf{A}^t$ hat als Eintragungen die Werte $b(k, l) = a(l, k)$ (d.h. die Rolle von Zeilen bzw. Spaltennummer werden vertauscht).

Einfache Übungsaufgabe:

$$\mathbf{A}^t * \mathbf{B}^t = (\mathbf{B} * \mathbf{A})^t.$$

Für komplexe Matrizen ist die Matrix $\mathbf{A}' := \text{conj}(\mathbf{A}^t)$ von größerer Bedeutung. Es gilt wieder $\mathbf{A}' * \mathbf{B}' = (\mathbf{B} * \mathbf{A})'$.

Die Anwendung einer Matrix $\mathbf{B}'$ auf einen Vektor $\mathbf{y}$ entspricht der Bildung von Skalarprodukten der Form $\langle \mathbf{y}, \mathbf{b}^k \rangle$, wobei die $\mathbf{b}^k$ -s die Spalten der ursprünglichen Matrix $\mathbf{B}$ durchlaufen.

In Kürze: Warum ist die Matrix so wichtig: weil die (charakteristische) Gleichung gilt:

$$\langle \mathbf{A} * \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{A}' * \mathbf{y} \rangle$$

VERBALE BESCHREIBUNG der Situation: Die Matrix $\mathbf{A}'$ ist diejenige Matrix, die man *auf der anderen Seite anschreiben muß*, wenn man die Wirkung der Matrix innerhalb des Skalarproduktes verschieben muss. Man beachte, dass schon bei Skalaren

(interpretiert als skalaren Vielfachen der entsprechenden Einheitsmatrix) im komplexen Fall die Konjugation auftritt, während man im Falle von reellen Matrizen einfach die transponierten Matrizen zu nehmen hat.

Wenn man die Vektoren $\mathbf{x}$ und $\mathbf{y}$ oben überhaupt durch Kollektionen von Spaltenvektoren ersetzt, so erkennt man, dass es letztendlich auf folgende Rechnung hinausläuft:

$$\mathbf{Y}' * (\mathbf{A} * \mathbf{X}) = (\mathbf{Y}' * \mathbf{A}'') * \mathbf{X} = (\mathbf{A} * \mathbf{Y})' * \mathbf{X}.$$

Die Koordinaten eines Vektors bzgl. einer Orthonormalbasis sind nicht nur leicht zu berechnen, sondern können aufgrund der obigen Formel auch leicht zur Bestimmung der Länge eines Vektors bestimmt werden. Insbesondere gilt (auch wenn das System nur ein Orthonormalsystem ist (und keine Basis), sagen wir 2 orthonormale Vektoren im $\mathbb{R}^3$ die eine Ebene aufspannen): Allgemein gilt für jedes ONS von r Vektoren im $\mathbb{R}^m$: $\|\sum_{k=1}^r c_k \mathbf{u}_k\| = \sqrt{\sum_{k=1}^r |c_k|^2}$. Die Abbildung $\mathbf{c} \mapsto \mathbf{x} = \mathbf{U} * \mathbf{c} = \sum_{k=1}^r c_k \mathbf{u}_k$ ist also *isometrisch* (d.h. größenhaltend, weil die euklid. Länge von $\mathbf{c}$ und $\mathbf{x} \in \mathbb{R}^m$ gleich sind. Auch die Winkel bleiben erhalten. Ohne Beweis behaupten wir dass davon auch die Umkehrung gilt.

Lemma 3. *Sie T eine lineare Abbildung von $\mathbb{R}^r$ nach $\mathbb{R}^m$ (beide mit dem Standard Skalarprodukt bzw. der Euklidischen Längenmessung versehen) welche isometrisch ist, dann gibt es passende Orthonormalbasen von $\mathbb{R}^r$ bzw. $\mathbb{R}^m$, sodass die Abbildung beschrieben werden kann als Abbildung des Koeffizienten Vektors $\mathbf{c} \in \mathbb{R}^r$ (Koeffizienten bzgl. der ONB in $\mathbb{R}^r$) auf die entsprechende Lin.Komb. in $\mathbb{R}^m$.*

BEISPIELE von einfachen orthogonalen Matrizen:

Drehung in der Ebene (im Raum ein weniger aufwändiger):

MATLAB Code zur Erstellung einer Drehmatrix in der Ebene:

```
function RM = rotang(alpha);
alpha = pi*alpha/180; % conversion to degrees
RM = [ cos(alpha), sin(alpha); -sin(alpha), cos(alpha)];
```

Warum? Antwort: Man muss nur feststellen, was die Bilder der Einheitsvektoren unter einer solchen (? ist es wirklich eine lineare Abbildung?) Drehung ist.

Spiegelung an einer Ebene im Raum: Vorerst an der (x, y) -Ebene in $\mathbb{R}^3$: $(x, y, z) = (x, y, -z)$. Das kann als Matrix natuerlich in Diagonalform geschrieben werden, mit den Eintragungen $(1, 1, -1)$. Alternativ betrachtet man diese Diagonalmatrix als eine Differenz zwischen der Einheitsmatrix und dem Doppelten der Projektion auf die z -Achse, also im Sinne der entsprechenden Diagonalmatrizen $(1, 1, -1) = (1, 1, 1) - 2 * (0, 0, 1)$. Das lässt sich leicht Verallgemeinern: Ist eine Ebene durch einen normierten Richtungsvektor $\mathbf{n}$ gegeben, so ersetze man den dritten Einheitsvektor $(0, 0, 1)$ durch $\mathbf{n}$, anders ausgedrückt, die Spiegelung an der durch den Normalvektor $\mathbf{n}$ gegebenen Ebene hat wird durch die Matrix $I_3 - \mathbf{n} * \mathbf{n}'$ dargestellt.

Drehung um eine Achse im Raum: Man muss offenbar eine ONB des $\mathbb{R}^3$ so wählen, dass die dritte (invariante) Richtung der Drehachse entspricht, und andererseits in den beiden anderen Koordinaten die Drehungen wie oben beschrieben aussehen, d.h. es ergibt sich in dieser Darstellung (bzgl. dieser Basis) eine Darstellung in der in der letzten Zeile bzw. Spalte nur genau eine von Null verschiedene Eintragung (und zwar ein Einser) auftreten. Daraus ergibt sich auch gleich wieder eine Frage: wie sieht die Matrix einer solchen Abbildung in der Standardbasis oder in irgendeiner anderen Basis aus? Diese Frage wird uns später noch in aller Allgemeinheit beschäftigen.

Charakterisierung von Rang r Matrizen

Wie immer gelten die Aussagen sowohl über dem Körper der reellen Zahlen wie auch über dem Körper der komplexen Zahlen.

Theorem 8. *Es sei T eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$, mit r -dimensionalen Bildraum, oder anders ausgedrückt: Der Rang der Matrix, welche T darstellt (durch die Gleichung $T(\mathbf{z}) = \mathbf{A} * \mathbf{z}$ für alle $\mathbf{z} \in \mathbb{R}^n$) sei r . Dann lässt sich T als eine Summe von Abbildungen vom Rang 1 schreiben, oder was dasselbe ist, es gilt*

$$T(\mathbf{z}) = \sum_{k=1}^r \langle \mathbf{z}, \mathbf{x}^k \rangle \mathbf{y}^k$$

Proof. Man kann de facto die Vektoren $\mathbf{y}^k$ beliebig als Basis des Bildraumes, also des Spaltenraumes der Matrix vorgeben. Z.B. kann man mit Hilfe der Gauss-Elimination die Spalten-Nummern der Pivot Elemente eruieren, und die entsprechenden Spalten der ursprünglichen Matrix nehmen (in meinem Standardbeispiel würde man also die 1-te, 3-te und 4-te Spalte als $\mathbf{y}^1 \dots \mathbf{y}^r$ nehmen). Die Eindeutigkeit der Koeffizienten der Bildvektoren $T(\mathbf{z})$ durch die Koordinaten bzgl. dieser Basis garantiert, dass für jedes k , $1 \leq k \leq r$ gilt: die Abbildung $\mathbf{z} \mapsto c^{(k)}(\mathbf{z})$ in der Darstellung $T(\mathbf{z}) = \sum_{k=1}^r c^{(k)}(\mathbf{z}) \mathbf{y}^k$ sind alle linear. Wir können also die Diskussion (durch Fixieren von k) auf einen solchen (Rang 1) Operator $\mathbf{z} \mapsto c^{(k)}(\mathbf{z}) \mathbf{y}^k$ reduzieren, und schreiben nur mehr $c(\mathbf{z})$ bzw. $\mathbf{y}$ (unter Weglassung des Index k). Da es sich um eine lin. Abb. von $\mathbb{R}^n$ nach $\mathbb{R}$ handelt, gibt es eine $1 \times n$ -Matrix, also einen Zeilenvektor $\mathbf{h}$ geben, sodass $c(\mathbf{z}) = \mathbf{h} * \mathbf{z}$ gilt. Es ist also nur mehr notwendig (um den komplexen Fall mitzudiskutieren) den gesuchten Vektor $\mathbf{x}$ als $\mathbf{h}'$ zu nehmen (das ist nun ebenfalls ein Spaltenvektor im $\mathbb{R}^n$), sodass dann $\mathbf{h} * \mathbf{z} = \langle \mathbf{z}, \mathbf{x} \rangle$, und wir erhalten insgesamt $c(\mathbf{z}) \mathbf{y} = (\mathbf{y} * \mathbf{x}') * \mathbf{z}$ (unter Nutzung der Assoziativität der Matrix-Multiplikation). Indem man diese Überlegung sukzessive auf die Vektoren $\mathbf{y}^1 \dots \mathbf{y}^r$ anwendet, kommt man zu dem gewünschten Ergebnis. $\square$

Besonderes Interesse erweckt der Fall, dass $\mathbf{x}^k = \mathbf{y}^k = \mathbf{o}^k$ ist, wobei die Folge $(\mathbf{o}^k)$ ein Orthonormalsystem sei, das einen Teilraum W von $\mathbb{R}^n$ erzeugen möge (man nennt ein solches ONS auch eine "Orthonormalbasis" für W). Dann kann man die Abbildung $\mathbf{z} \mapsto P(\mathbf{z}) = \sum_k \langle \mathbf{z}, \mathbf{o}^k \rangle \mathbf{o}^k$ betrachten.

VORSCHAU auf die unmittelbar bevorstehende Periode: der Vorteil von *Orthonormalsystemen*, die Herstellung von Orthonormalsystemen nach dem Verfahren von *Gram-Schmidt*.

Anmerkung f. später: Führt u.a. zur Interpretation eines lin. Gleichungssystems als einen Schnitt von Teilräumen (z.B. ist ein 3×3 -Gleichungssystem als Schnitt von drei affinen Ebenen interpretierbar.)

Definition 6. Das *Kreuzprodukt* ist eine spezielle Konstruktion, die es im $\mathbb{R}^3$ erlaubt, zu einem Paar von Vektoren einen dazu orthonormalen Vektor zu erzeugen ⁷ FORMEL! (siehe Bücher)

Gegeben zwei Vektoren $\mathbf{a}$ und $\mathbf{b}$ in $\mathbb{R}^3$ schreiben wir $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ für den Vektor, mit Koordinaten (! Prinzip der zyklischen Vertauschung, berechnet modulo 3): $c_k = a_{k+2}b_{k+1} - a_{k+1}b_{k+2}$.

WICHTIGE BEISPIELE für Abbildungen, die von orthogonalen Matrizen beschreiben werden. Geometrisch wohl klar: Man gibt die Abbildung so vor, dass eine (bel.) ONB $\mathbf{B}$ auf eine anderen ONB $\mathbf{C}$ des $\mathbb{R}^n$ abgebildet wird. Das leistet gerade die Matrix $\mathbf{U} = \mathbf{C} * \mathbf{B}^{-1}$. Dann ist klar, dass $\mathbf{U}$ unitär (orthog.) ist, und das gewünschte liefert. Für jeden Vektor $\mathbf{x}$ gilt dann also $\mathbf{x} \mapsto \mathbf{U} * \mathbf{x} = \mathbf{C} * (\mathbf{B} * \mathbf{x})$

Hinweis: Rechenaufwand: In MATLAB, welches eine *interpretierende* Sprache ist (es muss also nicht jedes Programm kompiliert werden, bevor es lauffähig ist), würde die Eingabe $\mathbf{x} \mapsto \mathbf{A} * \mathbf{B}\mathbf{x}$ als $\mathbf{x} \mapsto (\mathbf{A} * \mathbf{B}) * \mathbf{x}$ interpretiert, was $n^3 + n^2$ Multiplikationen kostet (?warum), im Gegensatz zur effizienteren Klammersetzung (in diesem Falle) $\mathbf{x} \mapsto \mathbf{A} * (\mathbf{B} * \mathbf{x})$, welche MATLAB dazu zwingt, zuerst $\mathbf{B} * \mathbf{x}$ zu realisieren (Aufwand n^2 Multiplikationen) und dann nochmals eine Matrix-Vektormultiplikation zu realisieren (wieder n^2 Mult.), also insgesamt $2n^2$ Multiplikationen). Natürlich ist der Unterschied erst für großes n spürbar!

Zurück zur Beschreibung der unitären Abbildung von oben. Wegen $\mathbf{B}^{-1} = \mathbf{B}$ können wir sie als $\mathbf{x} \mapsto \mathbf{U} * \mathbf{x} = \mathbf{C} * (\mathbf{B}' * \mathbf{x})$ beschreiben, bzw. in der üblichen Schreibweise mit Skalarprodukten

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

Definition 7. Es sei M eine Teilmenge eines (abstrakten oder konkreten) Vektorraumes V mit innerem Produkt $\langle \cdot, \cdot \rangle$. Der zu M orthogonale Teilraum $M^\perp$ gegeben durch

$$M^\perp := \{w \in V, \langle m, w \rangle = 0 \text{ für alle } m \in M\}.$$

Der Name ist jedenfalls gerechtfertigt, weil man leicht sieht, dass die so definierte Menge $M^\perp$ einen Unterraum von V bildet (abgeschlossen gegenüber Linearkombinationen). Weiters ist klar, dass $M^{\perp\perp} := (M^\perp)^\perp$ ebenfalls wieder ein Teilraum von V ist, und dass $M \subseteq M^{\perp\perp}$. Schliesslich müssen auch Linearkombinationen der Elemente von M in diesen Raum $M^{\perp\perp}$ liegen. In Kürze werden wir sehen, dass Gleichheit zwischen diesen Räumen besteht. Es gilt also: $M^{\perp\perp}$ ist der kleinste, M enthaltende Teilraum von V (man spricht auch vom "linearen Erzeugnis" von M).

Eine sehr wichtige Beobachtung ist nun aber auch noch die Tatsache, dass die orthogonalen Komplemente von endlichen Mengen M (mit m Elementen, $m \in \mathbb{N}$) gerade

⁷Jedenfalls dann, wenn diese beiden Vektoren eine Ebene aufspannen, also wenn sie lin. unabh. sind.

die Nullräume von homogenen Gleichungssystemen sind. Um diese Behauptung zu verifizieren genügt es, diese Vektoren⁸, man könnte also auch das Rezept angeben: Man schreibe diese m Vektoren als Spalten einer Matrix $\mathbf{B}$ an, und nehme dann $\mathbf{B}'$ als Systemmatrix eines linearen, *homogenen* Gleichungssystems. Dann ist die Übereinstimmung der beiden Begriffe: Nullraum des erstellten Gleichungssystems und Definition des zu M orthog. Teilraumes $M^\perp$ sofort ersichtlich. Wenn man noch bedenkt, dass im Zuge der Gaußelimination keine Veränderung der Lösungsmenge geschieht, erkennt man auch leicht das Prinzip: die Menge $M^\perp$ bleibt unverändert wenn man von einem Erzeugendensystem (z.B. die Zeilen von $\mathbf{B}'$) zu einer Basis (z.B. das System $\mathbf{B}'$ in Zeilenstufenform gebracht).

Die sogenannte *orthogonale Zerlegung* eines Raumes mit Skalarprodukt V nach einem Teilraum W und seinem orthog. Komplement $W^\perp$ erzeugt eine *direkte Summendarstellung* des Raumes, d.h. jedes $v \in V$ hat eine eindeutige Darstellung als $v = w + w^\perp$, mit $w \in W$ und $w^\perp \in W^\perp$. Aufgrund der Orthogonalität der beiden Teilräume gilt ausserdem⁹

$$\|v\|^2 = \|w\|^2 + \|w^\perp\|^2.$$

Wir wir später noch sehen werden gibt es zu jedem Teilraum $W \subseteq V$ eine Orthonormalbasis (mit r Elementen) für W , und genauso eine Orthonormalbasis für $W^\perp$ (mit $n - r$ Elementen, wenn $n = \dim(V)$ ist). Folglich kann man eine Orthonormalbasis für den ganzen Raum V erhalten, indem man die diese beiden Familien vereinigt, nennen wir diese Basis $\mathcal{B}$. (! Übung: Man verifiziere diese Aussage in allen technischen Details). Die Darstellung der Identität als

$$v = Id_V(x) = \sum_{k=1}^n \langle v, b^k \rangle b^k$$

zerfällt dann also in die Projektion P_W auf W , in der Form $P_W v = \sum_{k=1}^r \langle v, b^k \rangle b^k$ und $I - P_W$, welche klarerweise die Summe über die restlichen Terme ($k = r + 1 \dots n$) bilden.

Diese Situation ist *typisch* für die Beschreibung von orthogonalen Projektionen auf Teilräumen, denn natürlich (das ist die einfachere Richtung) kann man für jede Orthonormalbasis von V und jede Teilmenge dieser Basis (durch Umnummerieren der Elemente kann man diese Teilmenge immer in der Form $J = \{k : 1 \leq k \leq r\}$ bringen) ist natürlich das lineare Erzeugnis¹⁰ ein Raum $W \subseteq V$ wie oben beschrieben. Kurz gesagt, die Menge aller Projektionsoperatoren vom Rang r sind genau diejenigen, die auf folgende Weise entstehen (Details folgen noch):

```
A = rand(n); % erzeugt eine Zufallsmatrix, n x n;
B = orth(A); % baut diese BASIS in eine ONB fuer Rn um;
BW = B(:,1:r); % waehlt die ersten r Spalten davon aus;
P = BW * BW'; % bestimmt die Projektion auf W; 11
```

⁸Im Komplexen muss man korrekterweise die konjugierten dieser endlich vielen Vektoren als Zeilen in eine Matrix schreiben.

⁹Der sog. Pythagoräische Satz ist leicht aus den Eigenschaften des Skalarproduktes auch in der abstrakten Form herleitbar!

¹⁰Also die Linearkombinationen dieser r Basis Elemente, oder der kleinste von ihnen aufgespannte Teilraum.

Daraus ergibt sich auch die folgende wichtige Dimensionsformel für einen beliebigen Teilraum W eines n -dim. Vektorraums V mit Skalarprodukt $\langle \cdot, \cdot \rangle$:

ecomp-dims

$$(3) \quad \dim(W) + \dim(W^\perp) = n$$

Eine kleine nützliche Beobachtung sei hier eingefügt (auch wenn sie systematisch anderswohin gehört):

Lemma 4. *Die Einschränkung einer linearen Abbildung zwischen zwei Vektorräumen mit Skalarprodukt, also $T : \mathbf{V}_1 \mapsto \mathbf{V}_2$ ist injektiv, wenn man den Definitionsbereich auf das orthog. Komplement des Nullraumes von T einschränkt.*

Später wollen wir nochmals auf diese Thema zurückkommen, um folgende Situation genauer zu diskutieren: Wendet man dasselbe Argument dann auch noch auf die Matrix $\mathbf{A}'$ an, so findet man, dass der Nullraum der Matrix $\mathbf{A}'$ gerade das orthog. Komplement des Zeilenraumes von $(\mathbf{A}')$, also genau des Spaltenraums von $\mathbf{A}$ ist. Wenn man sich also eine allgemeine $m \times n$ Matrix $\mathbf{A}$ vorstellt, so hat man auf diese Weise sowohl im $\mathbb{R}^m$ als auch im $\mathbb{R}^n$ jeweils einen Teilraum (Zeilen bzw. Spaltenraum der Matrix) und sein orthog. Komplement, den Nullraum von $\mathbf{A}$ bzw. $\mathbf{A}'$. Dies sind die (wichtigen) vier Teilräume. Wir wissen bereits, dass Zeilenrang = Spaltenrang gilt.

Also bekommen wir die zweite Dimensionsformel. Wir betrachten die Situation, die im vorletzten Paragraphen angesprochen ist, und nehmen an, dass $\mathbf{V}_1$ ein n -dimensionaler Vektorraum ist.

im-formel2

$$(4) \quad \dim(\text{Null}(T)) + \dim(\text{range}(T)) = n$$

Kommentare zur orthogonalen Projektion auf einen Teilraum.

Theorem 9. *Es sei $(\mathbf{q}^j)_{j=1}^r$ ein Orthonormalsystem in $\mathbb{R}^n$, welches einen (offenbar) r -dimensionalen Teilraum $W \subseteq \mathbb{R}^n$ erzeugt.¹² Dann ist die Abbildung*

$$\mathbf{z} \mapsto P(\mathbf{z}) := \sum_{j=1}^r \langle \mathbf{z}, \mathbf{q}^j \rangle \mathbf{q}^j$$

die (eindeutig bestimmte) orthogonale Projektion von $\mathbb{R}^n$ auf W , welche durch folgende beiden Eigenschaften charakterisiert ist:

- (1) $P = P^2$;
- (2) $P = P'$;

Proof. Erst einmal ist klar, dass $P(\mathbb{R}^n) \subseteq W$ gilt, weil die Vektoren $\mathbf{q}^j$ in dem Teilraum W liegen. Andererseits gilt für $w \in W$ (und daher insbesondere für $w = P(\mathbf{z})$): $P(w) = w$, sodass wir sehen dass der Bildbereich von P genau gleich W ist, und dass $P^2 = P$ gilt.

¹¹Diese Methode ist für viele Zwecke durchaus OK, obwohl es kritische zu hinterfragen ist, ob auf diese Weise nicht nur mithilfe von stochastischen Methoden gewisse Teilräume erzeugt werden, bzw. ob die "Verteilung" richtig ist. Die genaue Diskussion dieser Frage soll hier ausgespart werden.

¹²Wir werden bald sehen, dass jeder r -dim. Teilraum W eines Vektorraumes V mit Skalarprodukt ein solches erzeugendes Orthonormalsystem besitzt.

Die Matrix für die Rang 1 Abbildung $\mathbf{z} \mapsto \langle \mathbf{z}, \mathbf{q}^j \rangle \mathbf{q}^j$ ist durch das (fast triviale) Matrix Produkt $\mathbf{q}^j * (\mathbf{q}^j)'$ gegeben, welche offenbar symmetrisch¹³ Als endliche Summe solcher hermitescher Matrizen hat auch P diese Eigenschaft.

Wir folgern nun, dass aus den beiden Eigenschaften (1) und (2) auch folgt, dass mit P auch $I - P$ eine Projektion ist mit $(I - P)^2 = P$ und $(I - P) = P$ (einfache Übungsaufgabe), die also ebenfalls eine orthogonale Projektion ist.

Anhand eines Beispiels, eines von 2 lin. unabh. Vektoren $\mathbf{a}^1$ und $\mathbf{a}^2$ aufgespannten, zweidim. Teilraumes V (d.h. die Elemente von V sind genau die Linearkombinationen dieser beiden Vektoren, oder in allgemeiner Schreibweise $V = \mathbf{A} * \mathbb{R}^2$, wobei $\mathbf{A}$ die durch die beiden Spaltenvektoren $\mathbf{a}^1$ und $\mathbf{a}^2$ gegebenen 3×2 -Matrix ist. Aber wie kann man nun ein **Biorthogonalsystem** $\mathbf{B}$ finden, sodass man jedes $v \in V$ schreiben kann als

$$\mathbf{v} = \sum_k \langle \mathbf{v}, \mathbf{b}^k \rangle \mathbf{a}^k.$$

Da die gesuchten 2 Vektoren in V gesucht werden, kann man ihre Matrix $\mathbf{B}$ als $\mathbf{B} = \mathbf{A} * \mathbf{X}$ schreiben, wobei die (beiden) Spalten der (2×2) -Matrix $\mathbf{X}$ die unbekannten Koeffizienten enthalten möge, die beschreiben, wie jedes $\mathbf{b}^j$ aus den Spalten von $\mathbf{A}$ durch Lin. Komb. hervorgeht.

Die Biorthogonalitätsrelation $\langle \mathbf{a}^k, \mathbf{b}^k \rangle = \delta_{j,k}$ (Kronecker Delta) ist in Matrix Schreibweise nichts anderes als $\mathbf{B}' * \mathbf{A} = \mathbf{Id}$. Es muss also gelten:

$$\mathbf{B}' * \mathbf{A} = \mathbf{X}' * (\mathbf{A}' * \mathbf{A}) = \mathbf{Id},$$

was auf die Forderung $\mathbf{X}' = \text{inv}(\mathbf{A}' * \mathbf{A})$ hinausläuft. Wenn wir uns erinnern dass $\mathbf{G} := \mathbf{A}' * \mathbf{A}$ die Eigenschaft $\mathbf{G}' = \mathbf{G}$ erfüllt, also hat seine inverse $\mathbf{X}$ dieselbe Eigenschaft (man nennt solche Matrizen *hermitesch*, im reellen Fall sind das einfach die symmetrischen Matrizen, d.h. diejenigen Matrizen, für die $a_{k,l} = a_{l,k}$ gilt). Wir kommen also zu dem Schluss dass $\mathbf{B} = \mathbf{A} * \mathbf{G}^{-1}$. Um also das k -te Element von $\mathbf{B}$ zu finden, genügt es die k -te Spalte der inversen Gram-Matrix zu nehmen und die dort befindlichen Eintragungen als Koeffizienten für die notwendige Linearkombination der Spalten von $\mathbf{A}$ zu nehmen. Wir haben also gefunden: das zu $\mathbf{A}$ biorthogonale System in V ist also $\mathbf{B} = \mathbf{A} * (\mathbf{A}' * \mathbf{A})^{-1}$.

Um die Invertierbarkeit der Gram-Matrix zu garantieren, muss man sich nur überlegen, dass der Nullraum von $\mathbf{A}$ und der Nullraum von $\mathbf{G} = \mathbf{A}' * \mathbf{A}$ gleich sind. Da die Spalten von $\mathbf{A}$ als linear unabh. angenommen wurden, ist dieser Nullraum trivial (d.h. er reduziert sich auf die einelementige Menge $\{0\}$). Da aber $\mathbf{G}$ eine quadratische $n \times n$ Matrix ist, die injektiv ist, muss wegen der Rang-Formel die Dimension des Bildes ebenfalls n sein (nämlich Raumdimension minus Defekt von $\mathbf{A}$, was def. gemäß die Dimension des Nullraums ist, und die ist in unserem Fall gleich Null), somit ist $\mathbf{x} \mapsto \mathbf{A} * \mathbf{x}$ auch eine surjektive Abbildung, und somit invertierbar!

Beweis der Gleichheit der Kerne von $\mathbf{G}$ und $\mathbf{A}$. Klar ist, dass jedes Element $\mathbf{x}$ im Kern von $\mathbf{A}$ wegen $\mathbf{G} * \mathbf{x} = \mathbf{A}' * (\mathbf{A} * \mathbf{x}) = \mathbf{A}' * \mathbf{0} = \mathbf{0}$ auch im Kern von $\mathbf{G}$ liegt. Umgekehrt, sei $\mathbf{x}$ im Kern von $\mathbf{G}$, es gelte also $\mathbf{G} * \mathbf{x} = \mathbf{A}' * \mathbf{A} * \mathbf{x} = \mathbf{0}$. Daraus folge aber sofort die Gleichheit

¹³Im Komplexen solle man sie **hermitesch** nennen, um zum Ausdruck zu bringen, dass sie die Eigenschaft $\mathbf{A}' = \mathbf{A}$ erfüllt.

$$\langle \mathbf{A} * \mathbf{x}, \mathbf{A} * \mathbf{x} \rangle = \langle \mathbf{A}' * \mathbf{A} * \mathbf{x}, \mathbf{x} \rangle = 0,$$

bzw. $\|\mathbf{A} * \mathbf{x}\|^2 = 0$, also $\mathbf{A} * \mathbf{x} = \mathbf{0}$, und somit liegt $\mathbf{x}$ im Kern von $\mathbf{A}$.

Viel interessanter noch ist die Aussage, dass $(I - P)$ die Projektion auf das orthog. Komplement von W definiert. Dazu zeigen wir zunächst einmal, dass der Bildbereich von $(I - P)$ in $W^\perp$ liegt, indem wir zeigen, dass für beliebige Paare $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt

$$\langle P\mathbf{x}, (I - P)\mathbf{y} \rangle = \langle P\mathbf{x}, \mathbf{y} - P\mathbf{y} \rangle = \langle P^2\mathbf{x}, \mathbf{y} \rangle - \langle P\mathbf{x}, P\mathbf{y} \rangle = \langle P\mathbf{x}, P'\mathbf{y} \rangle - \langle P\mathbf{x}, P\mathbf{y} \rangle = 0.$$

□

Die Argumentation oben kann auch leicht so interpretiert werden. Mit jeder Orthonormalbasis des $\mathbb{R}^n$ ergeben sich diverse Paare von zueinander komplementären und orthogonalen Teilräumen, so wie das Paar W und $W^\perp$ im obigen Satz.

Zwei wichtige Formeln: Parallelogrammregel

parallelogram

(5) $\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$

sowie die *Polarisierungsidentität* (im Komplexen):

ization-id

(6) $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=0}^3 i^k \|\mathbf{x} + i^k \mathbf{y}\|^2.$

Für den Beweis der letzteren Gleichung erweist es sich als nützlich, vorher zu zeigen, dass (bis auf den Faktor $\sqrt{4} = 2$) die 4×4 Vandermonde Matrix $\mathbf{F}_4$ unitär ist: Sie entsteht durch Wahl der vierten Einheitswurzeln in $\mathbf{C}$ in der Folge $1, -i, -1, i$ (als im Uhrzeigersinn, beginnend bei $1 = i^0$, wobei man beachten möge, dass $\pi/2 = (2\pi)/4$ ist, oder $e^{2\pi i x} = i$ für $x = \pi/2$).

TOPICS für Übungen

- Verschiedene Darstellungsformen von (affinen) Teilräumen (vgl. Schule: Punkt-Vektorform, Parameterdarstellung der Ebenen, bzw. als Lösungsmengen von Gleichungssystemen)
- (Vorbereitung für die FFT): Die DFT (discrete Fourier transform matrix) $\mathbf{F}$ ist die Vandermonde-Matrix für die Menge der n -ten Einheitswurzeln im Komplexen, typischerweise von $1 = \omega^0$ beginnend Uhrzeigersinn durchlaufend, wobei $\omega_n = \omega = e^{2\pi i/n}$. Die Eintragungen der ersten Spalte sind also lauter Einser, die in der zweiten Spalte der Reihe nach, von oben nach unten: $1, \overline{\omega}, \omega^2, \dots, \overline{\omega^{n-1}}$. Die weiteren Spalten sind (punktweise) die höheren Potenzen dieser zweiten Spalte, bis zur letzten (wo die $n - 1$ -ten Potenzen all dieser Werte stehen). Es ist auch bemerkenswert, dass die so entstehende Matrix auch symmetrisch ist, d.h. dass $a_{j,k} = a_{k,j}$ gilt, sie ist allerdings nicht hermitisch. Vielmehr bilden die Spalten ein Orthogonalsystem (jeder Vektor hat die Länge $\sqrt{n}$, sodass die inverse Matrix von der Form $(1/n)\text{conj}(\mathbf{F})$).

Affine Teilräume von $\mathbb{R}^n$

Definition 8. Ein *affiner Teilraum* von $\mathbb{R}^n$ ist eine Menge von der Form $\mathbf{x}_0 + V$, wobei V ein linearer Teilraum von $\mathbb{R}^n$ ist.

Theorem 10. Die affinen Teilräume des $\mathbb{R}^n$ sind genau die Lösungsmengen von inhomogenen linearen Gleichungssystemen.

Üblicherweise wird diese Aussage in folgender Form gemacht: Die allgemeine Lösung eines inhomogenen lin. Gleichungssystems besteht aus (irgend) einer (festen) Lösung des inhomog. Problems plus sämtliche Lösungen des homogenen linearen Gleichungssystems.

Proof. Aufgrund der Linearität kann man sofort feststellen, dass ich die Differenz von zwei Lösungen eines inhomog. Gleichungssystems $\mathbf{A} * \mathbf{x} = \mathbf{b}$ im Lösungsraum des homog. Glssys. befindet, wegen $\mathbf{A} * (\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0}$. $\square$

Definition 9. Eine Abbildung zwischen zwei Vektorräumen, welche von der Form: $A : \mathbf{x} \mapsto \mathbf{c} + T(\mathbf{x})$ ist, heißt eine affine Abbildung.

Klarerweise ist eine affine Abbildung genau dann linear, wenn der konstante Term $\mathbf{c} = \mathbf{0}$ erfüllt (die Darstellung ist ja eindeutig, und es gilt $A(\mathbf{0}) = \mathbf{c}$). Weiters haben affine Abbildungen eine Darstellung mit Hilfe von Matrizen:

Lemma 5. Es sei A eine affine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^m$. Dann gibt es eine eindeutig bestimmte Matrix $\mathbf{A}$ (Format $m \times n$) und einen Vektor $\mathbf{c} \in \mathbb{R}^m$, sodass gilt:

$$A(\mathbf{x}) = \mathbf{c} + \mathbf{A} * \mathbf{x}$$

EXERCISE: viele *geometrische Aufgaben* sind affin invariante Fragestellungen, wie etwa das Bestimmen des Mittelpunktes einer Strecke (oder allgemeinere Teilungsverhältnisse), die Bestimmung eines Schwerpunktes von mehreren Punkten im $\mathbb{R}^n$ und ähnliche Aufgaben. In der üblichen Geometrie wird aufgrund von "anschaulichen" Argumenten sehr oft von affinen Invarianten ausgegangen ¹⁴

Lemma 6. Die Bestapproximation auf einen affinen Teilraum ist eine affine Projektion, d.h. eine idempotente Abbildung von der Form ...

Was steht sonst noch auf dem Wunsch/Speisezettel

Grundbegriffe der *Determinantentheorie* ¹⁵ Die Lösbarkeit von 2×2 bzw. 3×3 Gleichungssystemen kann leicht durch die Berechnung einer "Zahl" realisiert werden.

Stoff der von Dr. Holger Rauhut gemacht wurde (Kurzbericht):

Konkret habe ich aus den Eigenschaften D1-D3 von Determinanten in dem Buch "Algorithmische Linear Algebra" die weiteren Eigenschaften D4-D11 nachgewiesen (D12 habe ich weggelassen).

Ausserdem habe ich die Formel für die Determinante von 2×2 Matrizen und 3×3 Matrizen gemacht, den Determinantenproduktsatz, Determinante der Transponierten und

¹⁴Man unterscheidet später noch zwischen der in der Schule verwendeten *affinen Geometrie* und der *projektiven Geometrie*.

¹⁵Wird normalerweise viel früher gemacht, weil es in der Praxis viel verwendet wird, und auch in den Büchern meist als wichtig dargestellt wird.

Determinante einer oberen Dreiecksmatrix (und erzählt, dass die Determinante bei Elementarschritten beim Lsen von linearen Gleichungssystemen erhalten bleibt.)

Bei den Eigenschaften D4-D11 war insbesondere dabei, dass die Determinante alternierend und linear in jeder Spalte ist, und dass die Determinante verschwindet genau dann, wenn die Spalten linear abhängig sind.

Ausserdem habe ich sogar die allgemeine Formel mit Permutationen und Signum von Permutationen gemacht (das liess sich dann doch nicht vermeiden, da ich sonst relativ weit am Anfang nicht mehr weitergekommen wäre). Das Signum einer Transposition habe ich allerdings nicht sehr ausführlich gemacht, insbesondere nichts bewiesen.

Allerdings hatte ich keine Zeit zu beweisen, dass die Determinante wirklich existiert, sondern nur erzählt, dass man anhand der allgemeinen Formel nachweisen muss, dass die Determinante wirklich die geforderten Eigenschaften besitzt.

Lemma 7. *Der (Absolut-)Betrag der Determinante beschreibt die Volumensänderung unter linearen Abbildungen. In der Tat gilt folgendes Prinzip: Eine lineare Abbildung (auch eine affine Abbildung, welche von der Form $\mathbf{x} \mapsto \mathbf{y} + \mathbf{A} * \mathbf{x}$ ist, führt Parallelepipede wieder in Parallel-epipede über. Die Zahl $|\det(\mathbf{A})|$ beschreibt den Modifikationsfaktor, soweit es affine lineare Abb. betrifft, mit einer quadr. Matrix $\mathbf{A}$.*

Darauf folgt auch die Produkt-Formel für Matrizen:

prod-det1

$$(7) \quad \det(\mathbf{A} * \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}).$$

Es gibt viele Methoden, um die (Existenz und die Eigenschaften) der Matrix als Funktion der (quadr.) Matrizen darzustellen.

Stoff, der von Andreas Klotz gemacht wurde (Kurzfassung): Information: Ich habe gestern das Gram-Schmidt Orthonormalisierungsverfahren behandelt (als Bp Legendre Polynome, dort musste ich kurz darauf hinweisen, wie das Skalarprodukt fuer stetige funktionen auf dem Intervall erklärt ist – Luecke bei positiver Definitheit). Anschliessend Cauchy-Schwarz (geometrische Motivation via Projektion eines Vektors auf den vom zweiten Vektor aufgespannten Raum, daraus dann der Beweis); Bpe dazu: Orthogonalprojektion von auf 1-dim UR hat Norm $\leq$ Norm(u); die CSU als Integralabschätzung, analog zur MWS der Integralrechnung (bzw L1-Linf-Hoelder); Winkel in R-VRen mit Skalarprodukt, Orthogonale Abbildungen erhalten Skalarprodukte (Bew. wurde nachgeliefert), damit dann auch Längen und Winkel.

WICHTIGE FORMELN zum LERNEN:

Cauchy-Schwarz Ungleichung:

$$(8) \quad |\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

Sie wird verwendet um zu zeigen, dass die vom Skalarprodukt abgeleitete Norm (auch bei “abstrakten Skalarprodukten”) die sog. Dreiecksungleichung erfüllt:

$$(9) \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$$

Parallelogrammregel

parallelogram

$$(10) \quad \|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$$

sowie die *Polarisierungsidentität* (im Komplexen):

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=0}^3 i^k \|\mathbf{x} + i^k \mathbf{y}\|^2. \quad (11)$$

Themen für das kommende Semester

Eigenvektoren, Eigenwerte, charakteristisches Polynom, Pseudoinverse, Singulärwertzerlegung, Regressionsgerade bzw. Regressionskurve, gewisse (einfache) numerische Aspekte (Konditionszahl), quadratische Formen, positive definite Matrizen, deren Wurzeln, symmetrische Orthonormalisierung

SCHNELLE TESTFRAGEN (im Aufbau)

- *Sämtliche Definitionen* (von linearer Unabhängigkeit bis zur unitären Matrix)
- Worauf ruht der *Dimensionsbegriff*? (ähnliche Frage: Wie kann man die Größe von Vektorräumen beschreiben).

Analogie zu endlichen Mengen: Je zwei endliche Mengen gleicher Kardinalität lassen sich bijektiv aufeinander abbilden.

- Wie bestimmt man eine Basis für den Zeilenraum bzw. den Spaltenraum einer Matrix; Wie zeigt man, dass diese beiden dieselbe Dimension haben (anders gefragt: wie kann man den Rang einer Matrix definieren?)
- Was ist die Verbindung zwischen Linearer Algebra und Matrix Kalkül (wie kann man die lineare Abbildung durch eine Matrix beschreiben, wie ist der Zusammenhang zwischen Komposition von linearen Abbildungen und der Matrix Multiplikation), Umkehrung einer linearen Abbildung durch Inversion einer Matrix (wobei es keine Rolle spielt, welche Basis man dafür wählt);
- Welche Schritte sind im Rahmen der Gauss Elimination zulässig?? (drei). Wie lassen sich diese als (Links-) Multiplikation mit Elementar-Matrizen (wie sehen diese aus) interpretieren. Was bedeutet es (geometrisch), dass die Umformungen im Rahmen der Gauss Elimination "lösungsäquivalent" sind?

Zusatzfrage: Was passiert mit dem Zeilen bzw. dem Spaltenraum im Zuge der Gauss-Elimination? Wie kann man die Invertierbarkeit einer linearen Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$ (warum nicht $\mathbb{R}^m$?) mithilfe der Gauss-Elimination feststellen?

- Man beschreibe den Zusammenhang zwischen linearen Gleichungssystemen und affinen Teilräumen und deren Darstellung (wie von der Schule bekannt). Z.B. gehört dazu die Frage: Gegeben eine lin. inhomog. Gleichungssystem, wie kann man die Lösungsmenge in Parameterform schreiben? Und umgekehrt, gegeben ein Teilraum in Parameterform, wie bekommt man ein (es gibt meist viele Möglichkeiten) passendes lin. Gleichungssystem. Wann ist dieses Gleichungssystem homogen (genau dann wenn der affine Teilraum den Nullvektor enthält, warum).
- *Skalarprodukte*: Welche Eigenschaften fordert man von einem Skalarprodukt (! *Sesquilinearität*); Wie ist aus einem (bel.) Skalarprodukt eine Größenmessung

zu machen? Beweis des Satzes von Pythagoras (Dreiecksungleichung und sog. Cauchy-Schwartz Ungleichung wird ein wenig später behandelt!).

- Orthonormalsysteme, orthog. Projektion auf einen Teilraum (charakterisiert durch die Eigenschaft der Projektion $\mathbf{P}$ mit $\mathbf{P} * \mathbf{P} = \mathbf{P}$ und $\mathbf{P}' = \mathbf{P}$).
- Darstellung der identischen Abbildung durch Orthonormalbasen
- Winkel und Längenerhaltung unter unitären Transformationen
- Basen für endlichdim. Teilräume eines Raumes mit Skalarprodukt, Einführung des Begriffes des Biorthogonalsystems (mit Hilfe der sog. *Gram-Matrix* $\mathbf{A}' * \mathbf{A}$).
- 4 Räume (Zeilen und Spaltenraum) und die zwei Nullräume von $\mathbf{A}$ bzw. $\mathbf{A}'$, die jeweils eine orthog. Zerlegung von $\mathbb{R}^n$ bzw. $\mathbb{R}^m$ liefern.
- Erläuterung der MNLSQ (minimal norm least squares solution) eines (möglicherweise) nicht im strengen Sinne lösbaren inhom. lin. Gleichungssystems ($\text{pinv}(\mathbf{A})$).
- Ausgleichsrechnung: best Fit durch z.B. kubische Polynomfunktionen, Vandermondematrix (und im exakten Fall: Lagrange Interpolationspolynoms).
- Vektorräume: Aufzählung von möglichst vielen Beispielen, z.B. normale n -tupel von reellen oder komplexen Zahlen, Matrizen, stetige Funktionen auf einem Intervall (oder auf ganz $\mathbb{R}$), die symmetrischen 5×5 -Matrizen (als Vektorraum über dem Körper der reellen Zahlen), die Polynomfunktionen (von festem Grad oder von bel. Grad), etc....
- Umwandlung der (affinen) Ebenen im Raum ($\mathbb{R}^3$) von Punkt-Parameterform und umgekehrt (um zu sehen, dass die affinen Teilräume genau die Lösungsmengen für inhomogene lineare Gleichungssysteme sind).
- Was sind hermitesche, unitäre, symmetrische, invertierbare Matrizen (und dergleichen mehr)? Welche Verbindungen gibt es. Ist jede symmetrische Matrix invertierbar? Ist jede orthogonale Matrix invertierbar? (jeweils mit Begründung). Kann es sein, dass die Spalten einer Matrix ein Orthogonalsystem bilden, ohne dass die Matrix orthogonal genannt werden kann??