Vertiefung Lineare Algebra 1

Schriftliche Unterlagen zur Vorlesung
im Wintersemester 2011/12

Franz Pauer

KAPITEL 1

Vertiefung zu Kap. 2, §3

In diesem Kapitel sei K ein Körper.

§1. Erzeugendensysteme, lineare Unabhängigkeit und Basen

In der Vorlesung „Lineare Algebra 1" wurde der Begriff *Linearkombination* eines n-Tupels von Vektoren eingeführt. Wir erweitern diesen Begriff nun auf beliebige (auch unendliche) Familien $(v_i)_{i\in I}$ von Vektoren. Die Definition im Skriptum Lineare Algebra 1 entspricht dann dem Spezialfall $I := \{1,\dots,n\}$.

Definition 1 : Sei V ein Vektorraum über K und I eine (beliebige) Menge.
Eine Familie $(c_i)_{i\in I}$ von Elementen in K heißt *Koeffizientenfamilie*, wenn $c_i \neq 0$ für nur endlich viele $i \in I$ ist.
Sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V.
Ein Vektor $w \in V$ heißt eine *Linearkombination* von $(v_i)_{i\in I}$, wenn es eine Koeffizientenfamilie $(c_i)_{i\in I}$ gibt, sodass

$$w = \sum_{i\in I, c_i\neq 0} c_i v_i$$

ist. Wir schreiben im weiteren einfach

$$\sum_{i\in I} c_i v_i \qquad \text{anstatt} \qquad \sum_{i\in I, c_i\neq 0} c_i v_i\,.$$

Die Menge aller Linearkombinationen von $(v_i)_{i\in I}$ ist ein Untervektorraum von V und enthält alle Vektoren v_i, $i \in I$. Er heißt der *von v_i, $i \in I$, erzeugte Untervektorraum von V* und wird mit

$${}_K\langle v_i \mid i \in I\rangle \qquad \text{oder} \qquad \sum_{i\in I} K v_i$$

bezeichnet.

Definition 2 : Sei $V \neq \{0\}$ ein Vektorraum über K. Eine Familie $(v_i)_{i\in I}$ von Vektoren in V heißt genau dann ein *Erzeugendensystem* von V bzw. *linear unabhängig* in V bzw. eine *Basis* von V, wenn jeder Vektor in V auf mindestens eine bzw. höchstens eine bzw. genau eine Weise als Linearkombination von $(v_i)_{i\in I}$ geschrieben werden kann.
Wir schreiben *linear abhängig* anstatt *nicht linear unabhängig*.
Die leere Menge ist eine *Basis von* $\{0\}$.

Satz 3 : *Sei $V \neq \{0\}$ ein Vektorraum über K und $(v_i)_{i\in I}$ eine Familie von Vektoren in V.*

(1) *Eine Basis von V ist ein linear unabhängiges Erzeugendensystem.*
(2) *Die Familie $(v_i)_{i\in I}$ von Vektoren ist genau dann ein Erzeugendensystem von V, wenn*
$$_K\langle v_i \mid i \in I\rangle = V$$
ist.
(3) *Die Familie $(v_i)_{i\in I}$ von Vektoren ist genau dann linear unabhängig, wenn für jede Koeffizientenfamilie $(c_i)_{i\in I}$ aus*
$$\sum_{i\in I} c_i v_i = 0$$
folgt, dass
$$c_i = 0 \quad \textit{für alle } i \in I$$
ist.

Beweis:

(1) und (2) folgen aus der Definition der Begriffe Erzeugendensystem, linear unabhängig und Basis.
(3) Wenn sich jeder Vektor aus V auf höchstens eine Weise als Linearkombination von $(v_i)_{i\in I}$ schreiben lässt, dann folgt aus
$$\sum_{i\in I} c_i v_i = 0_V = \sum_{i\in I} 0_K v_i$$
auf Grund der Eindeutigkeit $c_i = 0_K$ für $1 \leq i \leq n$.
Sei umgekehrt $(v_i)_{i\in I}$ linear unabhängig in V und $w \in V$ so, dass es eine Koeffizientenfamilie $(c_i)_{i\in I}$ mit
$$w = \sum_{i\in I} c_i v_i$$
gibt. Falls $(d_i)_{i\in I}$ eine weitere Koeffizientenfamilie mit
$$w = \sum_{i\in I} d_i v_i$$
ist, erhält man
$$0_V = w - w = \sum_{i\in I} c_i v_i - \sum_{i\in I} d_i v_i = \sum_{i\in I} (c_i - d_i) v_i \,.$$
Nach Annahme folgt $c_i - d_i = 0_K$ für alle $i \in I$, also $c_i = d_i$ für alle $i \in I$.

Beispiel 4 : Die Familie
$$(E_{k\ell})_{\substack{1\leq k\leq m\\ 1\leq \ell\leq n}} = (E_{k\ell})_{(k,\ell)\in\{1,\ldots,m\}\times\{1,\ldots,n\}}$$

der Standard-Matrizen ist eine Basis von $K^{m\times n}$ und heißt die *Standardbasis* von $K^{m\times n}$.

Für $A \in K^{m\times n}$ ist

$$A = \sum_{\substack{1\le k\le m\\ 1\le \ell\le n}} A_{k\ell}E_{k\ell}\,,$$

also ist A die Koordinatenfamilie von A bezüglich der Standardbasis $(E_{k\ell})_{\substack{1\le k\le m\\ 1\le \ell\le n}}$ (und $A_{k\ell}$ die Koordinate von A bei $E_{k\ell}$).

Beispiel 5 : Es seien I eine endliche Menge und V die Menge aller Funktionen von I in einen Körper K. Für $i \in I$ sei δ_i die durch

$$\delta_i(j) := \delta_{ij},\ j \in I,$$

definierte Funktion von I nach K. Dann ist die Familie

$$(\delta_i)_{i\in I}$$

eine K-Basis von V. Für $f \in V$ ist

$$f = \sum_{i\in I} f(i)\delta_i\,.$$

Beispiel 6 : Es sei K ein Körper und $F := \{(a_i)_{i\in\mathbb{N}} \,|\, a_i \in K, i \in \mathbb{N}\}$ die Menge aller Folgen in K. Mit den durch

$$(a_i)_{i\in\mathbb{N}} + (b_i)_{i\in\mathbb{N}} := (a_i + b_i)_{i\in\mathbb{N}} \ \text{ und } \ c\cdot(a_i)_{i\in\mathbb{N}} := (ca_i)_{i\in\mathbb{N}}$$

definierten Rechenoperationen wird F zu einem Vektorraum. Mit e_i ($i \in \mathbb{N}$) bezeichnen wir die Folge $(0,\ldots,0,1,0,\ldots,0,\ldots)$, wobei 1 an der i-ten Stelle steht. Die Familie $(e_i)_{i\in\mathbb{N}}$ ist linear unabhängig, aber kein Erzeugendensystem von F. Zum Beispiel kann die Folge $(1,1,1,\ldots,1,\ldots)$ nicht als Linearkombination der Familie $(e_i)_{i\in\mathbb{N}}$ geschrieben werden.

Der von der Familie $(e_i)_{i\in\mathbb{N}}$ erzeugte Untervektorraum ist die Menge aller Folgen, von denen nur endlich viele Folgenglieder nicht 0 sind. Er heißt *Vektorraum der endlichen Folgen in K*. Die Familie $(e_i)_{i\in\mathbb{N}}$ ist eine Basis dieses Vektorraums, daher ist er unendlich-dimensional.

§2. Rechnen mit Koordinaten

In diesem Abschnitt seien V und W endlich-dimensionale Vektorräume über K, $n \in \mathbb{N}$ die Dimension von V und $\underline{v} := (v_1,\ldots,v_n)$ eine Basis von V.

Definition 7 : Seien p und q positive ganze Zahlen und

$$\underline{u} := (u_1,\ldots,u_p) \in V^p$$

ein p-Tupel von Vektoren in V. Für $T \in K^{p\times q}$ sei

$$\underline{u}T := \big((\underline{u}T)_1, \dots, (\underline{u}T)_q\big) := \left(\sum_{i=1}^{p} T_{i1}u_i, \dots, \sum_{i=1}^{p} T_{iq}u_i\right) \in V^q.$$

Die Bezeichnung $\underline{u}T$ ist eine Merkhilfe, weil man analog zur Matrizenrechnung über K den Vektor $(\underline{u}T)_j \in V$ nach der Regel

„Zeile $\underline{u}$ mal Spalte T_{-j}"

berechnet.

Satz 8 : *Seien p, q, r positive ganze Zahlen und $\underline{u} := (u_1, \dots, u_p) \in V^p$. Dann gilt:*

(1) $\underline{u}I_p = \underline{u}$
(2) *Für $T \in K^{p\times q}$ und $U \in K^{q\times r}$ ist* $\underline{u}(TU) = (\underline{u}T)U$.
(3) *Für $T \in \mathrm{GL}_p(K)$ und $\underline{u'} := \underline{u}T$ gilt* $\underline{u} = \underline{u'}T^{-1}$.

Beweis: (1) gilt nach Definition, (2) rechnet man nach und (3) folgt aus $\underline{u} = \underline{u}I_p = \underline{u}(TT^{-1}) = (\underline{u}T)T^{-1} = \underline{u'}T^{-1}$.

Definition 9 : Seien $w \in V$ und $c_1, \dots, c_n \in K$ die Koordinaten von w bezüglich $\underline{v}$, also

$$w = \sum_{i=1}^{n} c_i v_i = \underline{v}c \in V\,,$$

Dann heißt die Spalte

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in K^{n\times 1}$$

die *Koordinatenspalte* von w bezüglich $\underline{v}$.

Satz 10 : *Sei $\underline{u} := (u_1, \dots, u_q) \in V^q$. Dann gilt:*

(1) *Es gibt genau eine Matrix $T \in K^{n\times q}$ mit*

$$\underline{u} = \underline{v}T\,.$$

Diese Matrix T heißt die Transformationsmatrix *von $\underline{v}$ zu $\underline{u}$, und die Spalten von T sind die Koordinatenspalten von $u_1, \dots, u_q$ bezüglich $\underline{v}$.*

(2) *Die Familie $\underline{u}$ ist genau dann eine Basis von V, wenn T invertierbar ist.*

Beweis:

(1) Sei $T \in K^{n\times k}$ jene Matrix, die sich durch Nebeneinanderschreiben der Koordinatenspalten von $u_1,\dots,u_k$ bezüglich $\underline{v}$ ergibt. Dann ist $u_j = \underline{v}T_{-j}$ für $1 \le j \le k$, also $\underline{u} = \underline{v}T$. Wenn umgekehrt $\underline{u} = \underline{v}S$ mit $S \in K^{n\times k}$ ist, d.h. $u_j = \underline{v}S_{-j}$ für $1 \le j \le k$, dann ist S_{-j} die Koordinatenspalte von u_j bezüglich $\underline{v}$ für $1 \le j \le k$.

(2) Wenn $T \in \mathrm{GL}_n(K)$ ist, dann ist $\underline{v} = \underline{u}T^{-1}$ nach Satz 8, also sind die Vektoren $v_1,\dots,v_n$ Linearkombinationen von $u_1,\dots,u_n$, somit ist ${}_K\langle u_1,\dots,u_n\rangle = V$ und, weil V n-dimensional ist, $(u_1,\dots,u_n)$ eine Basis von V. Wenn umgekehrt $\underline{u}$ eine Basis von V ist, dann gibt es eine Matrix $U \in K^{n\times n}$ mit $\underline{v} = \underline{u}U$. Weil $\underline{u}$ eine Basis ist, folgt aus $\underline{u}I_n = \underline{v}T = \underline{u}(UT)$, dass $UT = I_n$ ist und analog $TU = I_n$, also ist $T \in \mathrm{GL}_n(K)$.

Satz 11 : *Sei $\underline{u}$ eine Basis von V und $T \in \mathrm{GL}_n(K)$ die Transformationsmatrix von $\underline{v}$ zu $\underline{u}$. Ist c die Koordinatenspalte von $w \in V$ bezüglich $\underline{v}$, dann ist*

$$T^{-1}c$$

die Koordinatenspalte von w bezüglich $\underline{u}$, d.h. bei Basiswechsel mit der Matrix T „transformieren sich die Koordinaten“ mit der Matrix T^{-1}.

Beweis: Es ist $w = \underline{v}c = (\underline{u}T^{-1})c = \underline{u}(T^{-1}c)$.

Beispiel 12 : Wir betrachten die Ebene E nach Wahl eines Nullpunktes als Vektorraum. Wir wählen Punkte P_1, P_2 so, dass $\underline{P} := (P_1, P_2)$ eine Basis von E ist. Es sei

$$\underline{Q} := (Q_1, Q_2) := (\frac{3}{2}P_1 + 2P_2, \frac{1}{2}P_1 + \frac{1}{2}P_2).$$

Dann ist $\underline{Q} = \underline{P}T$, wobei

$$T := \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ 2 & \frac{1}{2} \end{pmatrix}$$

ist. Man prüft leicht nach, dass T invertierbar ist und dass

$$T^{-1} = \begin{pmatrix} -2 & 2 \\ 8 & -6 \end{pmatrix}$$

ist. Daher ist $\underline{P} = \underline{Q}T^{-1} = (-2Q_1 + 8Q_2, 2Q_1 - 6Q_2)$.

Der Punkt $P_1 + 2P_2$ hat bezüglich $\underline{P}$ die Koordinatenspalte $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ und bezüglich $\underline{Q}$ die Koordinatenspalte $T^{-1}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -4 \end{pmatrix}$.
Somit ist $P_1 + 2P_2 = 2Q_1 - 4Q_2$.

KAPITEL 2

Vertiefung zu Kap. 3, §2

§1. Strahlensatz

Satz 13: *(„Strahlensatz")*
Es seien Z_1, Z_2 zwei verschiedene, einander im Punkt 0 *schneidende Geraden in V, v_1, v_2 Punkte auf $Z_1 \setminus \{0\}$ und w_1, w_2 Punkte auf $Z_2 \setminus \{0\}$. Dann gibt es $c, d \in K \setminus \{0\}$ so, dass*

$$v_2 = cv_1 \quad \textit{und} \quad w_2 = dw_1$$

ist. Mit L_1 bzw. L_2 bezeichnen wir die Geraden durch die Punkte v_1 und w_1 bzw. v_2 und w_2. Dann gilt:

(1) *L_1 und L_2 sind genau dann parallel, wenn $c = d$ ist.*
(2) *Wenn L_1 und L_2 parallel sind, dann ist $v_2 - w_2 = c(v_1 - w_1)$.*

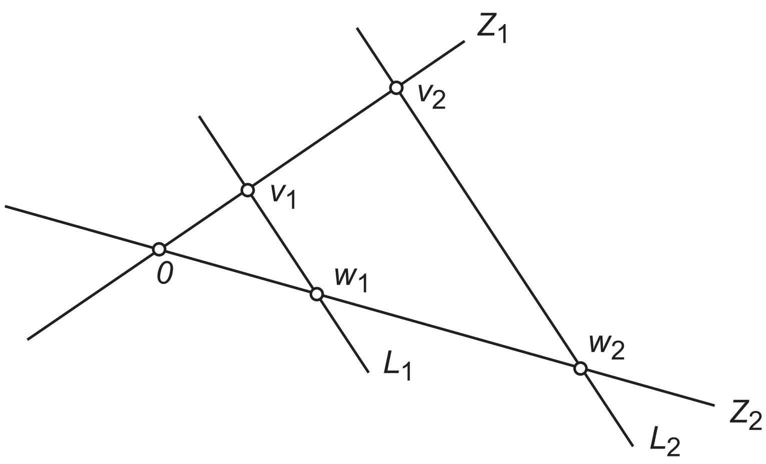


ABBILDUNG 1. Strahlensatz

Beweis:

(1) Der zu L_1 bzw. L_2 parallele Untervektorraum ist $K(v_1 - w_1)$ bzw. $K(cv_1 - dw_1)$. Weil die Geraden Z_1 und Z_2 verschieden sind, sind die Vektoren v_1 und w_1 linear unabhängig. Daher ist $K(v_1 - w_1)$ genau dann gleich $K(cv_1 - dw_1)$, wenn $c = d$ ist.

(2) Wenn L_1 und L_2 parallel sind, ist $c = d$ und

$$v_2 - w_2 = cv_1 - cw_1 = c(v_1 - w_1) \ .$$

Satz 14 : *Es seien V ein Vektorraum über einem Körper K und $Z_1 = p_1 + U_1$, $Z_2 = p_2 + U_2$ affine Unterräume von V mit Aufpunkten p_1, p_2 und parallelen Untervektorräumen U_1, U_2. Wenn Z_1 und Z_2 parallel sind, dann ist $Z_1 \subseteq Z_2$ oder $Z_2 \subseteq Z_1$ oder $Z_1 \cap Z_2 = \emptyset$.*

Beweis: Wir nehmen o.E.d.A. an, dass $U_1 \subseteq U_2$ ist. Wenn $Z_1 \cap Z_2$ nicht leer ist, dann gibt es ein $p \in Z_1 \cap Z_2$. Daher ist $Z_1 = p + U_1 \subseteq p + U_2 = Z_2$.

§2. Affine Hülle

Es sei K ein Körper und V ein Vektorraum über K.

Definition 15 : Es seien I eine endliche Menge und $(v_i)_{i \in I}$ eine Familie in V. Eine Linearkombination $\sum_{i \in I} c_i v_i$ von $(v_i)_{i \in I}$ heißt *affine Kombination* von $(v_i)_{i \in I}$, wenn $\sum_{i \in I} c_i = 1$ ist. Die Menge aller affinen Linearkombinationen von $(v_i)_{i \in I}$ heißt *affine Hülle* von $(v_i)_{i \in I}$.

Beispiel 16 : Die affine Hülle von zwei Vektoren v_1 und v_2 ist ein Punkt, wenn $v_1 = v_2$ ist, bzw. die Gerade

$$\{c_1 v_1 + c_2 v_2 \mid c_1, c_2 \in K, c_1 + c_2 = 1\} = \{v_1 + c(v_2 - v_1) \mid c \in K\},$$

wenn $v_1 \neq v_2$ ist.

Satz 17 :

(1) *Es seien M ein affiner Unterraum von V und $(v_i)_{i \in I}$ eine endliche Familie in M. Dann ist die affine Hülle von $(v_i)_{i \in I}$ in M enthalten.*
(2) *Die affine Hülle einer Familie $(v_i)_{i \in I}$ in V ist ein affiner Unterraum von V. Der dazu parallele Untervektorraum wird von $(v_i - v_j)_{i \in I,\, i \neq j}$ erzeugt, wobei $j \in I$ beliebig gewählt werden kann.*
(3) *Die affine Hülle von $(v_i)_{i \in I}$ ist der (bezüglich Inklusion) kleinste affine Unterraum, der alle v_i, $i \in I$, enthält.*

Beweis:

(1) Sei $p \in M$, U der zu M parallele Untervektorraum und $(c_i)_{i\in I}$ eine Familie in K mit $\sum_{i\in I} c_i = 1$. Dann ist

$$\sum_{i\in I} c_i v_i = \left(\sum_{i\in I} c_i\right) p - \left(\sum_{i\in I} c_i\right) p + \sum_{i\in I} c_i v_i =$$
$$= p + \sum_{i\in I} c_i (v_i - p) \in p + U = M.$$

(2) Sei $j \in I$ und

$$M := v_j + {}_K\langle v_i - v_j;\, i \in I,\, i \neq j\rangle.$$

Dann ist $(v_i)_{i\in I}$ eine Familie in M und nach (1) ist ihre affine Hülle in M enthalten.
Sei umgekehrt $(d_i)_{i\in I}$ eine Familie in K.
Dann ist

$$v_j + \sum_{\substack{i\in I\\ i\neq j}} d_i(v_i - v_j) = \sum_{\substack{i\in I\\ i\neq j}} d_i v_i + \left(1 - \sum_{\substack{i\in I\\ i\neq j}} d_i\right) v_j$$

eine affine Linearkombination von $(v_i)_{i\in I}$. Daher ist jedes Element von M in der affinen Hülle von $(v_i)_{i\in I}$ enthalten.

(3) Folgt aus (1) und (2).

Definition 18 : Affine Unterräume von V heißen *kollinear* bzw. *koplanar*, wenn sie alle in einer Geraden bzw. Ebene in V enthalten sind.

Satz 19 :

(1) *Drei Punkte v_1, v_2, $v_3 \in V$ sind genau dann kollinear, wenn die Vektoren $v_2 - v_1$ und $v_3 - v_1$ linear abhängig sind.*
(2) *Vier Punkte v_1, v_2, v_3, $v_4 \in V$ sind genau dann koplanar, wenn die Vektoren $v_2 - v_1$, $v_3 - v_1$ und $v_4 - v_1$ linear abhängig sind.*
(3) *Zwei Geraden $p_1 + Kv_1$ und $p_2 + Kv_2$ sind genau dann koplanar, wenn die Vektoren $p_1 - p_2$, v_1 und v_2 linear abhängig sind.*

Beweis: Die ersten zwei Aussagen folgen aus Satz 17, (2). Der zur affinen Hülle von $(p_1, p_2, p_1 + v_1, p_2 + v_2)$ parallele Untervektorraum wird von $p_1 - p_2$, v_1 und v_2 erzeugt.

Satz 20 : *Zwei verschiedene koplanare Geraden schneiden einander in genau einem Punkt oder sie sind parallel.*

Beweis: Seien M_1 und M_2 verschiedene koplanare Geraden und E die Ebene, die beide enthält. Wenn M_1 und M_2 nicht parallel sind, dann ist $U_1 \cap U_2 = \{0\}$ und $U_1 + U_2 = U_1 \oplus U_2$ ist der zu E parallele Untervektorraum. Wegen $p_1,\ p_2 \in E$ ist $p_1 - p_2 \in U_1 \oplus U_2$, daher gibt es eindeutig bestimmte Vektoren $u_1 \in U_1,\ u_2 \in U_2$ so, dass $p_1 - p_2 = u_1 + u_2$ ist. Somit ist $M_1 \cap M_2 = \{p_1 - u_1\} = \{p_2 + u_2\}$.

§3. Polytope und Schwerpunkte

Es seien $K = \mathbb{Q}$ oder $\mathbb{R}$ und V ein Vektorraum über K.

Definition 21 : Es seien I eine endliche Menge und $(v_i)_{i\in I}$ eine Familie in V.
Eine Linearkombination $\sum\limits_{i\in I} c_i v_i$ von $(v_i)_{i\in I}$ heißt *konvexe Linearkombination* von $(v_i)_{i\in I}$, wenn $\sum\limits_{i\in I} c_i = 1$ und $c_i \geq 0$ für alle $i \in I$ ist.
Die Menge der konvexen Linearkombinationen von $(v_i)_{i\in I}$ heißt *konvexe Hülle* von $(v_i)_{i\in I}$.
Die konvexe Hülle zweier Vektoren v_1, v_2 heißt *Strecke* zwischen v_1 und v_2.
Die konvexe Hülle dreier nicht kollinearer Punkte v_1, v_2, v_3 heißt *Dreieck* mit Eckpunkten v_1, v_2, v_3 .
Eine Teilmenge von V heißt *Polytop*, wenn sie die konvexe Hülle einer endlichen Familie in V ist.

Es sei $I := \{1,\ldots,n\}$ und $c_1,\ldots,c_n \in \mathbb{R}_{\geq 0}$. Für $c_n \neq 1$ ist

$$\sum_{i=1}^{n} c_i v_i = (1-c_n)\left(\sum_{i=1}^{n-1} \frac{c_i}{1-c_n} v_i\right) + c_n v_n = (1-c_n)w + c_n v_n\,,$$

wobei $w := \sum_{i=1}^{n-1} \frac{c_i}{1-c_n} v_i$ in der konvexen Hülle H von $(v_1,\ldots,v_{n-1})$ liegt. Daraus folgt: Für $n \geq 3$ ist die konvexe Hülle von $(v_1,\ldots,v_n)$ die Vereinigung aller Strecken zwischen v_n und den Elementen von H.

Beispiel 22 : Eine Teilmenge von $\mathbb{R}$ ist genau dann ein Polytop, wenn sie ein abgeschlossenes Intervall ist.

Satz 23 : *Es seien P die konvexe Hülle einer Familie $\left(w_j\right)_{j\in J}$ in V und $(v_i)_{i\in I}$ eine Familie in P. Dann ist die konvexe Hülle von $(v_i)_{i\in I}$ in P enthalten.*

Beweis: Für alle $i \in I$ ist der Vektor v_i eine konvexe Linearkombination $\sum\limits_{j\in J} c_{ji}w_j$ von $\left(w_j\right)_{j\in J}$.
Sei $\sum\limits_{i\in I} d_iv_i$ eine konvexe Linearkombination von $(v_i)_{i\in I}$. Dann ist

$$\sum_{i\in I} d_iv_i = \sum_{i\in I}\sum_{j\in J} d_ic_{ji}w_j = \sum_{j\in J}\left(\sum_{i\in I} d_ic_{ji}\right)w_j$$

mit $\sum\limits_{i\in I} d_ic_{ji} \geq 0,$ für alle $j \in J$, und

$$\sum_{j\in J}\left(\sum_{i\in I} d_ic_{ji}\right) = \sum_{i\in I} d_i\left(\sum_{j\in J} c_{ji}\right) = \sum_{i\in I} d_i = 1\ .$$

Daher ist $\sum\limits_{i\in I} d_iv_i \in P$.

Definition 24: Es sei $(v_i)_{i\in I}$ eine endliche Familie in V .
Der *Schwerpunkt* von $(v_i)_{i\in I}$ ist

$$\frac{1}{\#(I)}\sum_{i\in I} v_i\ .$$

Der Schwerpunkt von (v_1, v_2) heißt *Mittelpunkt der Strecke* zwischen v_1 und v_2.

Satz 25: *Es seien u, v, w drei nicht kollineare Punkte in V . Die Gerade durch u bzw. v bzw. w und den Mittelpunkt der Strecke zwischen den anderen zwei Punkten heißt* Schwerlinie *des Dreiecks mit Eckpunkten u, v, w durch u bzw. v bzw. w.*
Die drei Schwerlinien sind paarweise verschieden und schneiden einander im Schwerpunkt $\frac{1}{3}(u+v+w)$ von (u, v, w) .

Beweis: Da u, v, w nicht kollinear sind, sind nach Satz 19 die Vektoren $v-u$ und $w-u$ linear unabhängig. Also sind auch

$$v-u \quad \text{und} \quad \frac{1}{2}(v-u)+\frac{1}{2}(w-u) = \frac{1}{2}(v+w)-u$$

linear unabhängig, nach Satz 19 sind daher $u, v, \frac{1}{2}(v+w)$ nicht kollinear. Somit liegt v nicht auf der Schwerlinie durch u . Daher sind die Schwerlinien durch u und durch v verschieden und die drei Schwerlinien haben höchstens einen Schnittpunkt. Wegen

$$\frac{1}{3}(u+v+w) = \frac{1}{3}u + \frac{2}{3}\left(\frac{1}{2}(v+w)\right) = \frac{1}{3}v + \frac{2}{3}\left(\frac{1}{2}(u+w)\right) =$$
$$= \frac{1}{3}w + \frac{2}{3}\left(\frac{1}{2}(u+v)\right)$$

liegt der Schwerpunkt auf allen Schwerlinien.

§4. Affine Räume

Definition 26 : Es seien $(G,\star)$ eine Gruppe mit neutralem Element e und M eine Menge. Eine Funktion $G \times M \longrightarrow M$, $(s,m) \longmapsto s \cdot m$, ist eine *Operation der Gruppe G auf der Menge M*, wenn gilt:
für alle $m \in M$ ist $e \cdot m = m$ und
für alle $s,t \in G$ und alle $m \in M$ ist $(s \star t) \cdot m = s \cdot (t \cdot m)$.

Beispiel 27 : Die Funktion

$$S_n \times \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}, \ (\sigma,i) \mapsto \sigma(i),$$

ist eine Operation der Permutationsgruppe S_n auf der Menge $\{1,2,\ldots,n\}$.

Definition 28 : Sei V ein Vektorraum über einem Körper K, A eine Menge und

$$V \times A \to A, \ (v,a) \mapsto v \cdot a,$$

eine Operation der Gruppe $(V,+)$ auf A. (Also: Für alle $a \in A$, $v,w \in V$ ist $0 \cdot a = a$ und $(v+w) \cdot a = v \cdot (w \cdot a)$.
A zusammen mit dieser Operation ist ein *affiner Raum über* V, wenn es für alle Elemente $a,b \in A$ genau einen Vektor $v \in V$ gibt mit $v \cdot a = b$.
Die Elemente von A heißen dann *Punkte*, die Elemente von V *Vektoren* des affinen Raums.

Satz 29 : *Sei A ein affiner Raum über V und $a \in A$. Die Funktion*

$$V \longrightarrow A, \ v \longmapsto v \cdot a,$$

ist bijektiv. (Nach Wahl eines „Nullpunktes“ kann ein affiner Raum als Vektorraum betrachtet werden).

Beweis: Folgt aus der Definition.

Beispiel 30 : Sei V ein Vektorraum, $p \in V$ und U ein Untervektorraum von V. Dann ist der affine Unterraum $p+U$ mit

$$U \times (p+U) \longrightarrow p+U \;\; (v, p+u) \longmapsto p+(u+v),$$

ein affiner Raum über U. Insbesondere ist jeder Vektorraum ein affiner Raum (über sich selbst).

Beispiel 31 : Sie E die Zeichenebene oder der Anschauungsraum und $T(E)$ der Vektorraum der Translationen von E. Dann ist E mit

$$T(E) \times E \longrightarrow E, \; (t,x) \longmapsto t(x),$$

ein affiner Raum über $T(E)$.
Möchte man in der Zeichenebene keinen „Nullpunkt“ wählen, kann man sie als affinen Raum betrachten. Dann muss man zwischen Punkten ($\in E$) und Vektoren ($\in T(E)$) unterscheiden. Punkte können dann nicht addiert werden, aber Vektoren können addiert werden und auf Punkten „wirken“.

Sind P und Q Punkte von E und $P \neq Q$, dann gibt es genau eine Translation in $T(E)$, die P auf Q abbildet. Sie wird häufig mit $\vec{PQ}$ bezeichnet. Die Menge

$$\{t\vec{PQ} \,|\, t \in \mathbb{R}\} \subseteq T(E)$$

ist die Gerade durch $0_{T(E)} = id_E$ und $\vec{PQ}$ in $T(E)$. Die „Gerade durch P und Q in E“ ist dann als

$$\{(t\vec{PQ})(P) \,|\, t \in \mathbb{R}\} \subseteq E$$

definiert. Wegen $(\vec{PQ})(P) = Q$ und $(0 \cdot \vec{PQ})(P) = id_E(P) = P$ sind P und Q Punkte dieser Geraden. Die Translation $\vec{PQ}$ wird als „Richtungsvektor“ dieser Geraden bezeichnet.

KAPITEL 3

Vertiefung zu Kap. 3, §3-5

§1. Mehr über Skalarprodukte

In diesem Abschnitt sei V ein reeller Vektorraum und $\langle -,- \rangle$ ein Skalarprodukt auf V.

Definition 32 : Ein reeller Vektorraum mit einem Skalarprodukt heißt *reeller Prähilbertraum*. Ein endlich-dimensionaler reeller Prähilbertraum heißt *euklidischer Raum*.

Beispiel 33 : (Für Studierende mit Kenntnissen aus Analysis). Es seien a,b reelle Zahlen mit $a < b$ und V der Vektorraum aller stetigen Funktionen vom Intervall $[a,b]$ nach $\mathbb{R}$. Die Funktion

$$\langle -,- \rangle : V \times V \longrightarrow \mathbb{R} \ , \quad (f,g) \longmapsto \int_a^b f(x) \cdot g(x) dx \ ,$$

ist ein Skalarprodukt auf V. Die Norm von $f \in V$ ist

$$\|f\| = \sqrt{\int_a^b f(x)^2 dx} \quad .$$

Satz 34 : *(Parallelogrammgleichung) Für alle Vektoren v und w in einem reellen Prähilbertraum ist*

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

Beweis: Nachrechnen.

Definition 35 : Eine *Norm* auf V ist eine Funktion

$$N : V \longrightarrow \mathbb{R}_{\geq 0} := \{t \in \mathbb{R} \,|\, t \geq 0\} \ , \ v \longmapsto N(v) \ ,$$

mit den Eigenschaften:

(1) Es ist $N(v) = 0$ genau dann, wenn $v = 0$ ist.
(2) Für alle $c \in \mathbb{R}$, $v \in V$ ist $N(cv) = |c| N(v)$.
(3) Für alle $v,w \in V$ ist $N(v+w) \leq N(v) + N(w)$.

V zusammen mit einer Norm heißt *normierter Raum* .

Beispiel 36 : Ist $\langle -,- \rangle$ ein Skalarprodukt auf V, dann ist

$$\| \cdot \| : V \longrightarrow \mathbb{R} \quad , \quad v \longmapsto \|v\| := \sqrt{\langle v,v \rangle},$$

eine Norm.

Beispiel 37 : Die Funktion

$$N : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} \ , \ (c_1, c_2, \dots, c_n) \longmapsto \sum_{i=1}^{n} |c_i| \ ,$$

ist eine Norm (die „Summennorm") auf $\mathbb{R}^n$.
Die Parallelogrammgleichung gilt für diese Norm nicht, zum Beispiel ist für $n = 2$, $v := (2,1)$ und $w := (1,2)$:
$N(v+w)^2 + N(v-w)^2 = 36 + 4 = 40$, aber
$2 \cdot N(v)^2 + 2 \cdot N(w)^2 = 18 + 18 = 36$.

Satz 38 : *N sei eine Norm auf V.*
Dann sind die folgenden Aussagen äquivalent:

(1) *Die Parallelogrammgleichung gilt für N, dh.: Für alle $v,w \in V$ ist $N(v+w)^2 + N(v-w)^2 = 2 \cdot N(v)^2 + 2 \cdot N(w)^2$.*
(2) *Es gibt ein Skalarprodukt $\langle -,- \rangle$, das die Norm N induziert, dh.: für alle $v \in V$ ist $N(v) = \sqrt{\langle v,v \rangle}$.*

In diesem Fall ist dieses Skalarprodukt von $v,w \in V$ durch

$$\langle v,w \rangle := \frac{1}{4}(N(v+w)^2 - N(v-w)^2)$$

definiert.

Beweis: Wenn (2) gilt, dann folgt (1) aus Satz 34.
Wenn (1) gilt, dann kann damit nachgeprüft werden, dass die oben definierte Funktion $\langle -,- \rangle$ die Eigenschaften eines Skalarproduktes hat.

Definition 39 : Eine Familie $(v_i)_{i \in I}$ in V heißt *orthonormal* bezüglich $\langle -,- \rangle$, wenn für alle $i, j \in I$ gilt:

$$\langle v_i, v_j \rangle = \delta_{ij}.$$

Eine Familie $(v_i)_{i \in I}$ in V heißt *Orthonormalbasis* (kurz: *ON-Basis*) von V bezüglich $\langle -,- \rangle$, wenn sie eine Basis von V und orthonormal bezüglich $\langle -,- \rangle$ ist.

Beispiel 40 : Es sei V der Vektorraum der endlichen Folgen in $\mathbb{R}$. Für $i \in \mathbb{N}$ bezeichnen wir mit e_i die Folge $(0, \dots, 0, 1, 0, \dots, 0, \dots)$, wobei 1 an der i-ten Stelle steht. Die Familie $(e_i)_{i\in\mathbb{N}}$ ist eine Basis von V. Durch

$$\langle a,b\rangle := \sum_{i\in\mathbb{N}} a_i b_i\,, \text{für alle Folgen } a,b \text{ in } \mathbb{N}$$

wird ein Skalarprodukt auf V definiert. Die Familie $(e_i)_{i\in\mathbb{N}}$ ist eine ON-Basis bezüglich dieses Skalarproduktes.

Satz 41 : *Eine orthonormale Familie ist linear unabhängig. Insbesondere: Wenn V endlich-dimensional ist, dann ist jede orthonormale Familie mit* $\dim_K(V)$ *Elementen eine ON-Basis von V.*

Beweis: Sei $(v_i)_{i\in I}$ eine orthonormale Familie und $(c_i)_{i\in I}$ eine Koeffizienten-Familie in K. Wenn $\sum_{i\in I} c_i v_i = 0$ ist, dann ist für alle $j \in I$ auch

$$0 = \langle v_j, \sum_{i\in I} c_i v_i\rangle = \sum_{i\in I} c_i \langle v_j, v_i\rangle = c_j\ .$$

Satz 42 : *Sei $w \in V$ und $(v_i)_{i\in I}$ eine ON-Basis von V. Dann ist*

$$w = \sum_{i\in I} \langle v_i, w\rangle v_i\ .$$

(„Die Koordinate von w bei v_i ist das Skalarprodukt von v_i mit w“.)

Beweis: Sei $w = \sum_{i\in I} c_i v_i$. Dann ist

$$\langle v_j, w\rangle = \langle v_j, \sum_{i\in I} c_i v_i\rangle = \sum_{i\in I} c_i \langle v_j, v_i\rangle = \sum_{i\in I} c_i \delta_{ji} = c_j\ .$$

§2. Interpolationsaufgaben

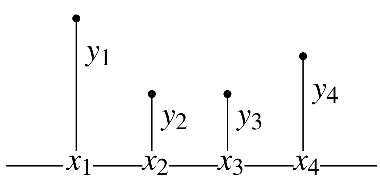
Wir betrachten die folgenden *Interpolationsaufgaben*:
Gegeben sind

- Funktionen $f_1,\dots,f_k$ von $\mathbb{R}$ nach $\mathbb{R}$,
- paarweise verschiedene reelle Zahlen $x_1,\dots,x_n \in \mathbb{R}$ und
- reelle Zahlen $y_1,\dots,y_n \in \mathbb{R}$.

Gesucht sind reelle Zahlen $c_1,\dots,c_k$ so, dass die Funktion $f := \sum_{i=1}^{k} c_i f_i$ die Bedingungen

$$f(x_1) = y_1,\ f(x_2) = y_2,\ \dots,\ f(x_n) = y_n$$

erfüllt.



Durch die Funktionen $f_1,\ldots,f_k$ wird der „Typ“ der Interpolationsaufgabe vorgegeben. Die reellen Zahlen $x_1,\ldots,x_n$ heißen *Stützstellen*, die reellen Zahlen $y_1,\ldots,y_n$ *(Funktions-)Werte* der Interpolationsaufgabe. Die gesuchte Funktion f heißt *interpolierende Funktion*.

Wir suchen also eine Funktion f des vorgegebenen Typs so, dass die Funktionswerte von f in den Stützstellen die vorgegebenen Werte der Interpolationsaufgabe sind.

Anders formuliert: Wir suchen Zahlen $c_1,\ldots,c_k$ so, dass

$$\begin{array}{rcl} f_1(x_1)c_1+f_2(x_1)c_2+\ldots+f_k(x_1)c_k &=& y_1 \\ f_1(x_2)c_1+f_2(x_2)c_2+\ldots+f_k(x_2)c_k &=& y_2 \\ \vdots & \vdots & \vdots \\ f_1(x_n)c_1+f_2(x_n)c_2+\ldots+f_k(x_n)c_k &=& y_n \end{array}$$

ist. Das ist ein System von n linearen Gleichungen mit k Unbekannten $c_1,\ldots,c_k$. In Matrizenform:

$$\begin{pmatrix} f_1(x_1) & \ldots & f_k(x_1) \\ f_1(x_2) & \ldots & f_k(x_2) \\ \vdots & \vdots & \vdots \\ f_1(x_n) & \ldots & f_k(x_n) \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} .$$

Beispiel 43: („Lineare Interpolation“).
Wenn f_1 die konstante Funktion 1 (also die Funktion, die jeder Zahl die Zahl 1 zuordnet) und f_2 die Identität (also die Funktion, die jeder Zahl sich selbst zuordnet) ist, dann suchen wir eine Funktion $f := c_1f_1 + c_2f_2$ mit

$$(f(x_i) =)\ c_1 + c_2x_i = y_i\,,\ 1 \le i \le n\,.$$

Die Aufgabe, Zahlen c_1 und c_2 mit den Eigenschaften

$$\begin{array}{rcl} c_1+c_2x_1 &=& y_1 \\ \vdots & \vdots & \vdots \\ c_1+c_2x_n &=& y_n \end{array}$$

zu finden, ist ein System von n linearen Gleichungen mit zwei Unbekannten. In Matrizenform

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} .$$

Beispiel 44: (Interpolation durch Polynomfunktionen).
Für $1 \le i \le k$ sei $f_i : \mathbb{R} \longrightarrow \mathbb{R},\ z \longmapsto z^i$, die i-te Potenzfunktion. Dann ist die gesuchte Funktion f eine Polynomfunktion vom Grad $k-1$, also
$f : \mathbb{R} \longrightarrow \mathbb{R}, z \longmapsto c_1 + c_2z + \ldots + c_kz^{k-1}$.

Wir suchen reelle Zahlen $c_1, c_2, \ldots, c_k$ mit der Eigenschaft, dass

$$\begin{array}{rcl} c_1 + x_1 c_2 + \ldots + x_1^{k-1} c_k & = & y_1 \\ \vdots & \vdots & \vdots \\ c_1 + x_n c_2 + \ldots + x_n^{k-1} c_k & = & y_n \end{array}$$

ist, müssen also ein System von n Gleichungen mit k Unbekannten lösen. In Matrizenform:

$$\begin{pmatrix} 1 & x_1 & \ldots & x_1^{k-1} \\ 1 & x_2 & \ldots & x_2^{k-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \ldots & x_n^{k-1} \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} .$$

§3. Systeme linearer Gleichungen (mit und ohne Lösung)

Das System

$$\begin{array}{rcl} f_1(x_1)c_1 + f_2(x_1)c_2 + \ldots + f_k(x_1)c_k & = & y_1 \\ f_1(x_2)c_1 + f_2(x_2)c_2 + \ldots + f_k(x_2)c_k & = & y_2 \\ \vdots & \vdots & \vdots \\ f_1(x_k)c_1 + f_2(x_k)c_2 + \ldots + f_k(x_n)c_k & = & y_n \end{array}$$

von k linearen Gleichungen mit den Unbekannten $c_1, \ldots, c_n$ kann auch in der Form

$$c_1 \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_1(x_n) \end{pmatrix} + \ldots + c_k \begin{pmatrix} f_k(x_1) \\ \vdots \\ f_k(x_n) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

oder, mit den Abkürzungen

$$\mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{und} \quad \mathbf{f_i}(\mathbf{x}) := \begin{pmatrix} f_i(x_1) \\ \vdots \\ f_i(x_n) \end{pmatrix}, 1 \leq i \leq k,$$

kurz als

$$\sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x}) = \mathbf{y}$$

angeschrieben werden.

Wir können also das System linearer Gleichungen als die folgende Aufgabe interpretieren: Schreibe die Spalte $\mathbf{y}$ als Linearkombination $\sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x})$ der Spalten $\mathbf{f_i}(\mathbf{x})$, $1 \leq i \leq k$. Das ist aber nur dann möglich, wenn $\mathbf{y}$ in dem von den Spalten $\mathbf{f_i}(\mathbf{x})$, $1 \leq i \leq k$, erzeugten Untervektorraum U von $\mathbb{R}^{n \times 1}$ enthalten ist. Wenn das nicht der Fall ist, ist dieses System linearer Gleichungen nicht lösbar.

Im Fall von Beispiel „Lineare Interpolation“ ist U die von

$$\mathbf{1} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{und} \quad \mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

erzeugte Ebene in $\mathbb{R}^{n\times 1}$.

Im Fall von Beispiel „Interpolation durch Polynomfunktionen vom Grad $\leq k$“ ist U der von

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ldots, \begin{pmatrix} x_1^{k-1} \\ \vdots \\ x_n^{k-1} \end{pmatrix}$$

erzeugte Untervektorraum von $\mathbb{R}^{n\times 1}$.

Wenn die Interpolationsaufgabe einerseits eine Situation beschreibt, von der man weiß, dass es eine Lösung gibt, andererseits die Aufgabe aber nicht lösbar ist, weil $\mathbf{y}$ mit Mess- oder Rundungsfehlern behaftet ist, liegt es nahe, dass $\mathbf{y}$ „eigentlich“ ein Element von U sein sollte. Wir erzwingen die Lösbarkeit der Aufgabe, indem wir $\mathbf{y}$ durch eine Spalte $\mathbf{y}'$ in U ersetzen!

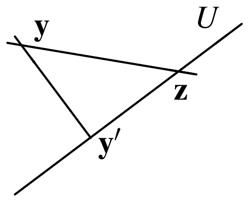
Wie sollen wir diese Spalte $\mathbf{y}'$ aber wählen? Am einfachsten ist es, $\mathbf{y}'$ in U so zu wählen, dass der Abstand zwischen $\mathbf{y}'$ und $\mathbf{y}$ möglichst klein ist. Wir suchen also ein Element des Vektorraums U so, dass der Abstand $\|\mathbf{y}'-\mathbf{y}\|$ zwischen $\mathbf{y}$ und $\mathbf{y}'$ so klein wie möglich ist. Wir müssen nun festlegen, welchen Abstand wir meinen: Wenn wir den (durch das Standardskalarprodukt auf dem $\mathbb{R}^n$ induzierten) euklidischen Abstand im $\mathbb{R}^n$ wählen, dann ist

$$\|\mathbf{y}'-\mathbf{y}\| := \sqrt{\sum_{i=1}^{n} (y_i' - y_i)^2}\,.$$

Für positive reelle Zahlen a und b ist $a \leq b$ genau dann, wenn $a^2 \leq b^2$ ist. Daher ist der $\|\mathbf{y}'-\mathbf{y}\|$ genau dann minimal, wenn die Summe $\sum_{i=1}^{n}(y_i'-y_i)^2$ der „Fehlerquadrate“ minimal ist.
Die Spalte $\mathbf{y}'$ ist der Fußpunkt des Lotes von $\mathbf{y}$ auf U.

§4. Lineare Regression

Genau dann ist der Abstand zwischen $\mathbf{y}'$ und $\mathbf{y}$ kleiner oder gleich dem Abstand zwischen $\mathbf{y}$ und jedem anderen Element $\mathbf{z}$ von U, wenn die Gerade durch $\mathbf{y}'$ und $\mathbf{y}$ normal zum Untervektorraum U steht. Das folgt leicht aus dem Satz von Pythagoras:



Das Dreieck mit den Eckpunkten $\mathbf{y}'$, $\mathbf{y}$ und $\mathbf{z}$ hat bei $\mathbf{y}'$ einen rechten Winkel. Der Abstand zwischen $\mathbf{y}$ und $\mathbf{z}$ ist die Länge der Hypotenuse, also größer oder gleich der Länge einer Kathete, also dem Abstand zwischen $\mathbf{y}$ und $\mathbf{y}'$.

Die Gerade durch $\mathbf{y}'$ und $\mathbf{y}$ steht genau dann normal zu U, wenn alle Skalarprodukte von $\mathbf{y}' - \mathbf{y}$ mit den erzeugenden Spalten von U gleich 0 sind. Für $\mathbf{y}' = \sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x}) \in U$ muss also gelten:

$$\langle \mathbf{y}' - \mathbf{y}, \mathbf{f_j}(\mathbf{x}) \rangle = 0,\ 1 \le j \le k.$$

Anders geschrieben:

$$\sum_{i=1}^{k} c_i \langle \mathbf{f_i}(\mathbf{x}), \mathbf{f_j}(\mathbf{x}) \rangle = \langle \mathbf{y}, \mathbf{f_j}(\mathbf{x}) \rangle,\ 1 \le j \le k.$$

Wenn wir dieses System von k linearen Gleichungen mit k Unbekannten $c_1, \ldots, c_k$ lösen, dann erhalten wir die „annähernd“ interpolierende Funktion $f = \sum_{i=1}^{k} c_i f_i$.

Bei linearer Interpolation ist

- $\mathbf{y}' = c_2 \mathbf{x} + c_1 \mathbf{1} \in U$ und
- die Gerade durch $\mathbf{y}$ und $\mathbf{y}'$ steht normal auf der von $\mathbf{x}$ und $\mathbf{1}$ erzeugten Ebene U.

Also ist

- $\langle c_2 \mathbf{x} + c_1 \mathbf{1} - \mathbf{y}, \mathbf{x} \rangle = 0$ und
- $\langle c_2 \mathbf{x} + c_1 \mathbf{1} - \mathbf{y}, \mathbf{1} \rangle = 0$.

Daraus erhalten wir das folgende System von zwei linearen Gleichungen mit zwei Unbekannten c_1 und c_2:

- $c_2 \langle \mathbf{x}, \mathbf{x} \rangle + c_1 \langle \mathbf{1}, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$
- $c_2 \langle \mathbf{x}, \mathbf{1} \rangle + c_1 \langle \mathbf{1}, \mathbf{1} \rangle = \langle \mathbf{1}, \mathbf{y} \rangle$

Als Lösung erhalten wir

$$c_2 = \frac{\langle \mathbf{1}, \mathbf{1} \rangle . \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{1}, \mathbf{x} \rangle . \langle \mathbf{1}, \mathbf{y} \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle . \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{1}, \mathbf{x} \rangle^2} \quad \text{und} \quad c_1 = \frac{\langle \mathbf{x}, \mathbf{x} \rangle . \langle \mathbf{1}, \mathbf{y} \rangle - \langle \mathbf{1}, \mathbf{x} \rangle . \langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle . \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{1}, \mathbf{x} \rangle^2}.$$

Wenn $\langle -,- \rangle$ das Standard-Skalarprodukt ist, dann ist $\langle \mathbf{x},\mathbf{y} \rangle = \sum_{i=1}^n x_i y_i$, $\langle \mathbf{1},\mathbf{x} \rangle = \sum_{i=1}^n x_i$, $\langle \mathbf{1},\mathbf{y} \rangle = \sum_{i=1}^n y_i$, $\langle \mathbf{1},\mathbf{1} \rangle = n$, $\langle \mathbf{x},\mathbf{x} \rangle = \sum_{i=1}^n x_i^2$ und $\langle \mathbf{y},\mathbf{y} \rangle = \sum_{i=1}^n y_i^2$, daher

$$c_2 = \frac{n\sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

und

$$c_1 = \frac{(\sum_{i=1}^n x_i^2)(\sum_{i=1}^n y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} .$$

Wir haben damit die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$, $z \longmapsto c_2 z + c_1$, so bestimmt, dass der (euklidische) Abstand vom n-Tupel der gegebenen (gemessenen oder gerundeten) ungenauen Funktionswerte $(y_1,\ldots,y_n)$ zum n-Tupel der berechneten Funktionswerte $(f(x_1),\ldots,f(x_n))$ möglichst klein ist, also $\sum_{i=1}^n (y_i - (c_2 x_i + c_1))^2$ möglichst klein ist. Der Graph dieser Funktion heißt *Regressionsgerade* oder *Trendlinie* der Punkte $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$. Man rechnet leicht nach, dass

$$f(\frac{1}{n}\sum_{i=1}^n x_i) = \frac{1}{n}\sum_{i=1}^n y_i$$

ist. Das Paar der arithmetischen Mittel von $(x_1,\ldots,x_n)$ und $(y_1,\ldots,y_n)$ liegt also immer auf der Regressionsgeraden.