Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Basiswechsel, Rang und Normalform für lineare Abbildungen

- Basiswechsel
- Normalform für lineare Abbildungen
- Rang einer Matrix
- Dimensionsformel für Untervektorräume

In der letzten Vorlesung hatten wir gesehen, wie eine linear Abbildung $f\colon V \to W$ durch eine Matrix

$$A = M_{\mathcal{B}}^{\mathcal{A}}(f)$$

beschrieben werden kann, wobei $\mathcal{A} = \{v_1, \dots, v_n\}$ eine Basis von V und $\mathcal{B} = \{w_1, \dots, w_n\}$ eine Basis von W ist. Ziel dieser Vorlesung ist es zu verstehen, wie sich Basiswechsel auswirken.

Basiswechsel, 1

Satz und Definition. *Es seien* V, W *zwei endlich-dimensionale* K*-Vektorräume mit Basen*

$$\mathcal{A} = \{v_1, \ldots, v_n\},\ \mathcal{B} = \{w_1, \ldots, w_m\}$$

und $f\colon V \to W$ *eine lineare Abbildung.* $A = M^{\mathcal{A}}_{\mathcal{B}}(f)$ *sei die Matrixdarstellung von* f *bezüglich dieser Basen.*
Sind $\mathcal{A}' = \{v'_1, \ldots, v'_n\}$, $\mathcal{B}' = \{w'_1, \ldots, w'_m\}$, *dann ergibt sich die Matrixdarstellung* $B = M^{\mathcal{A}'}_{\mathcal{B}'}(f)$ *in den neuen Basen wie folgt:*

$$B = TAS^{-1},$$

wobei $T = M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)$, $S = M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)$ *die sogenannten* **Basiswechselmatrizen** *sind.*

Hierbei bezeichnen $\mathrm{id}_V : V \to V$ und $\mathrm{id}_W : W \to W$ jeweils die identischen Abbildungen.

Basiswechsel, 2

Mit anderen Worten: Das folgende Diagramm kommutiert:

$$\begin{tikzcd}
K^n \arrow[r, "B"] \arrow[d, "\varphi_{\mathcal{A}'}"] & K^m \arrow[d, "\varphi_{\mathcal{B}'}"'] \\
V \arrow[r, "f"] & W \\
K^n \arrow[r, "A"] \arrow[u, "\varphi_{\mathcal{A}}"'] \arrow[uu, bend left, "M^{\mathcal{A}}_{\mathcal{A}'}(\mathrm{id}_V)=S"] & K^m. \arrow[u, "\varphi_{\mathcal{B}}"] \arrow[uu, bend right, "T=M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)"']
\end{tikzcd}$$

Insbesondere gilt $B = T \cdot A \cdot S^{-1}$.

Beweis. Klar nach Definition der Matrixdarstellung. Beispielsweise kommutiert

$$\begin{tikzcd}
W & K^m \arrow[l, "\varphi_{\mathcal{B}'}"'] \\
W \arrow[u, "\mathrm{id}_W"] & K^m. \arrow[l, "\varphi_{\mathcal{B}}"] \arrow[u, "T=M^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{id}_W)"']
\end{tikzcd}$$

□

Klassifikationsatz von linearen Abbildungen

Satz. *Sei $f : K^n \to K^m$ die durch die Matrix $A \in K^{m \times n}$ definierte lineare Abbildung. Dann existieren $S \in \mathrm{GL}(n, K)$, $T \in \mathrm{GL}(m, K)$, so dass:*

$$TAS^{-1} = \begin{array}{c} \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ \hline & 0 & & 0 \end{array} \right) \\ \underbrace{\hphantom{1 \ddots 1}}_{r} \underbrace{\hphantom{0}}_{n-r} \end{array} \quad \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\ \ddots\\ 1\end{array}}\right\} r \\ \left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\} m-r \end{array}$$

für ein $r \leq \min(n, m)$ gilt.

Ende Teil 1

Beweis. Wir wählen Basen von K^n und K^m geschickt: Wir betrachten dazu den Kern

$$\ker A = \{x \in \mathbb{R}^n \mid Ax = 0\}.$$

Ist $d = \dim(\ker A)$, so setzen wir $r = n - d$ (also $r \leq n$). Als erstes wählen wir eine Basis von $\ker A \subset K^n$, die wir mit $v_{r+1}, \dots, v_n$ durchnummerieren. Anschließend ergänzen wir diese durch Vektoren $v_1, \dots, v_r \in K^n$ zu einer Basis $\mathcal{A} = \{v_1, \dots, v_n\}$ von K^n. Seien $w_i = f(v_i), i = 1, \dots, r$, die Bilder der ersten r Vektoren. Dann sind $w_1, \dots, w_r \in K^m$ linear unabhängig: Wären sie nämlich abhängig, etwa

$$\lambda_1 w_1 + \cdots + \lambda_r w_r = 0,$$

so wäre

$$\lambda_1 v_1 + \cdots + \lambda_r v_r \in \ker A = \langle v_{r+1}, \dots, v_n \rangle,$$

das heißt, $\{v_1, \dots, v_n\}$ wäre keine Basis, außer $\lambda_1 = \cdots = \lambda_r = 0$. Dies zeigt:

$$r \leq m \; (= \dim K^m).$$

Wir ergänzen nun $w_1, \ldots, w_r$ zu einer Basis $\mathcal{B} = \{w_1, \ldots, w_r, w_{r+1}, \ldots, w_m\}$ des K^m. Bezüglich der Basen $\mathcal{A}$ und $\mathcal{B}$ hat f die Gestalt:

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \begin{array}{cl} \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & 0 \\ & & 1 & \\ \hline & & & \\ & 0 & & 0 \\ & & & \end{array}\right) & \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} r \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} m - r \end{array} \\ \underbrace{\hphantom{1 \quad \ddots \quad 1}}_{r} \; \underbrace{\hphantom{\quad 0 \quad}}_{n-r} & \end{array}$$

Dies folgt sofort aus $f(v_i) = w_i, i = 1, \ldots, r$ und $f(v_j) = 0, i = r+1, \ldots, n$.

Wenn also $S = M^{\mathcal{E}}_{\mathcal{A}}(\mathrm{id}_{K^n})$ und $T = M^{\mathcal{E}}_{\mathcal{B}}(\mathrm{id}_{K^m})$ die Basiswechselmatrizen sind, so folgt, dass das Diagramm

$$\begin{tikzcd} K^n \arrow[r, "M^{\mathcal{A}}_{\mathcal{B}}(f)"] & K^m \\ K^n \arrow[u, "S"] \arrow[r, "A"] & K^m \arrow[u, "T"'] \end{tikzcd}$$

kommutiert. $\square$

Bezüglich geeigneter Basen kann jede lineare Abbildung zwischen endlich–dimensionalen Vektorräumen, also durch eine sehr einfache Matrix beschrieben werden. Der Wechsel zur passenden Basis im Definitions– bzw. Zielvektorraum wird jeweils von einer invertierbaren Matrix realisiert.

Die Dimensionsformel für lineare Abbildungen

Korollar (aus dem Beweis). *Sei $f\colon V \to W$ eine lineare Abbildung zwischen zwei K-Vektorräumen und* $\dim V < \infty$. *Dann gilt:*

$$\dim \operatorname{Bild}(f) + \dim \ker(f) = \dim V.$$

Beweis. Ist $d = \dim \ker f$ und $n = \dim V$, dann ist mit der Notation aus dem Beweis $\operatorname{Bild}(f) = \langle w_1, \ldots, w_r \rangle$, also $\dim \operatorname{Bild}(f) = r$, wobei $r = n - d$. $\square$

In geeigneten Basen ist jede lineare Abbildung $f\colon V \to W$ zwischen endlich–dimensionalen K–Vektorräumen eine Parallelprojektion:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_r \\ v_{r+1} \\ \vdots \\ v_n \end{pmatrix} \mapsto \begin{pmatrix} v_1 \\ \vdots \\ v_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Rang einer linearen Abbildung

Definition. Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlich–dimensionalen Vektorräumen. Dann ist der Rang von f durch

$$\operatorname{rang}(f) := \dim \operatorname{Bild} f$$

definiert. Der Rang einer $m \times n$-Matrix $A \in K^{m\times n}$ ist der Rang der zugehörigen lineare Abbildung $K^n \to K^m$, $x \mapsto Ax$:

$$\operatorname{rang} A := \operatorname{rang}(\ K^n \xrightarrow{A} K^m\ ).$$

Satz. *Sei $A \in K^{m\times n}$ eine Matrix und $\tilde{A}$ die Matrix in Zeilenstufenform die der Gaußalgorithmus zurückgibt. Hat $\tilde{A}$ genau r Stufen, dann gilt*

$$r = \operatorname{rang} A.$$

Beweis. Die Transformation in Zeilenstufenform entspricht einem Basiswechsel im K^m. An der Dimension des Bildes ändert das nichts. Also $\operatorname{rang} A = \operatorname{rang} \tilde{A} = r$. □

Ende Teil 2

Die transponierte Matrix

Sei $A = (a_{ij}) \in K^{m \times n}$ eine $m \times n$-Matrix. Dann ist die **transponierte Matrix** $A^t = (b_{ij}) \in K^{n \times m}$ durch

$$b_{ij} := a_{ji}$$

definiert.

Beispiel.
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow A^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Bemerkung. $(A \cdot B)^t = B^t \cdot A^t$.

Satz. *Eine Matrix und ihre Transponierte haben den gleichen Rang:*

$$\operatorname{rang} A = \operatorname{rang} A^t.$$

Beweis. Es gilt $\operatorname{rang} A = r$ genau dann, wenn es invertierbare Matrizen $S \in \mathrm{GL}(n, K)$ und $T \in \mathrm{GL}(n, K)$ gibt, so dass TAS^{-1} in der Normalform mit r Einträgen 1 auf der Diagonalen und sonst Nullen ist. Die Matrix $(TAS^{-1})^t = (S^{-1})^t A^t T^t$ ist auch in Normalform, und $(S^{-1})^t$ und $(T^t)^{-1}$ sind ebenfalls invertierbar.

Anwendung auf lineare Gleichungssyteme

Sei $Ax = b$ mit $A \in K^{m \times n}$, $b \in K^m$, ein Gleichungssystem und

$$Ax = 0$$

das **zugehörige homogene Gleichungssystem**. Dann ist die Lösungsmenge des homogenen Gleichungssystems der Untervektorraum

$$\ker A = \{x \in K^n \mid Ax = 0\}.$$

Für inhomogene Gleichungssysteme betrachtet man zu $y \in K^n$ die Parallelräume

$$y + \ker A := \{y + x \in K^n \mid x \in \ker A\}.$$

Satz. *Ist $\tilde{x} \in K^n$* **eine** *Lösung des inhomogenen Gleichungssystems*

$$Ax = b,$$

dann ist dessen ganze Lösungsmenge $L_b = \{x \in K^n \mid Ax = b\}$ der Parallelraum

$$L_b = \tilde{x} + \ker A.$$

Beweis. Für $x \in \ker A$ und $\tilde{x}$ mit $A\tilde{x} = b$ gilt

$$A(\tilde{x} + x) = A\tilde{x} + Ax = A\tilde{x} = b,$$

also $\tilde{x} + \ker A \subset L_b$. Umgekehrt gilt $L_b \subset \tilde{x} + \ker A$:

$$\begin{aligned} x' \in L_b \quad &\Rightarrow \quad A(x' - \tilde{x}) = Ax' - A\tilde{x} = b - b = 0 \\ &\Rightarrow \quad x' - \tilde{x} \in \ker A \\ &\Rightarrow \quad x' \in \tilde{x} + \ker A. \end{aligned}$$

Ist $b \notin \mathrm{Bild}(A)$, dann existiert kein $\tilde{x} \in K^n$ mit $A\tilde{x} = b$ und $L_b = \emptyset$. $\square$

Quadratische Gleichungssyteme

Wir betrachten nun den wichtigen Spezialfall von Gleichungssystemen mit genauso vielen Gleichungen wie Unbestimmten, d.h., $A \in K^{n \times n}$.

Satz. *Sei $A \in K^{n \times n}$ und $b \in K^n$; mit f bezeichnen wir die zugehörige lineare Abbildung. Dann sind äquivalent:*

1. *A ist invertierbar, d.h., $A \in GL(n, K)$,*
2. $\ker A = 0$, *d.h., f ist ein Monomorphismus,*
3. $\operatorname{Bild} A = K^n$, *d.h., f ist ein Epimorphismus,*
4. *$Ax = b$ hat genau eine Lösung.*

Beweis.

1\. $\Rightarrow$ 2.: A ist invertierbar, d.h. $\exists\, A^{-1} \in K^{n \times n}$, so dass $A^{-1} \cdot A = E$

$$\Rightarrow \ker A \subset \ker A^{-1}A = \ker E = 0.$$

2. $\Leftrightarrow$ 3.: Die Dimensionsformel sagt:

$$\dim \ker A + \dim \operatorname{Bild} A = \dim K^n = n.$$

Also:

$$\ker\ A = 0 \Leftrightarrow \dim \operatorname{Bild} A = n \Leftrightarrow \operatorname{Bild} A = K^n.$$

2. $\wedge$ 3. $\Rightarrow$ 4.: $\operatorname{Bild} A = K^n$ besagt: Für jedes b hat die Gleichung $Ax = b$ eine Lösung. Und $\ker A = 0$ besagt: Wenn es eine Lösung gibt, ist diese eindeutig bestimmt.

4. $\Rightarrow$ 1.: Wir betrachten die Lösungen $v_j \in K^n$ der Gleichungen

$$Av_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, Av_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

und bilden die Matrix

$$B = (v_1, \dots, v_b) \in K^{n \times n}.$$

Dann gilt: $A \cdot B = E$ ($\Rightarrow B \cdot A = E$) $\Rightarrow B = A^{-1}$, also: A ist invertierbar. $\square$

Bemerkungen.

1. Häufig will man das Gleichungssystem
$$Ax = b$$
für eine Matrix $A \in K^{n \times n}$ und viele verschiedene b berechnen. Dann lohnt es sich, die Inverse A^{-1}, etwa mit Gauß zu berechnen.
2. Für $A \in GL(n, K)$ ist der Aufwand, A^{-1} mit Hilfe des Gaußalgorithmus zu berechnen, von der Größenordnung $O(n^3)$ Körperoperationen.
3. Eine Matrixmultiplikation
$$A \cdot B$$
auszurechnen für $A, B \in K^{n \times n}$ hat mit der Formel aus der Definition ebenfalls den Aufwand $O(n^3)$, denn es gibt n^2 Einträge von $A \cdot B$, und
$$c_{ik} = \sum_{j=1}^{n} a_{ik} b_{kj}$$
besteht aus n Termen. Es geht auch mit weniger Aufwand:

4. Für $n = 2$ braucht man mit der klassischen Formel 8 Multiplikationen. Eine Entdeckung von Strassen (1969) besagt, dass man mit 7 Multiplikationen auskommt. Dies liefert für allgemeines n einen niedrigeren Aufwand $O(n^{log_2 7}) \approx O(n^{2.7})$.
5. Es ist offen, ob eine asymptotische Laufzeit von $O(n^2)$ für die Matrixmultiplikation möglich ist, was optimal wäre, da die Ausgabe aus n^2 Einträgen besteht.
   Ist dies der Fall, dann lässt sich auch A^{-1} in $O(n^2)$ berechnen! Dazu eventuell später mehr.

Summen von Vektorräumen

Definition. Seien V ein K-Vektorraum und $U, W \subset V$ zwei Untervektorräume. Dann bezeichnet

$$U + W = \{v \in V \mid \exists\, u \in U, \exists\, w \in W : v = u + w\}$$

die Summe der Untervektorräume.
Die **äußere** bzw. **direkte Summe** von U und W ist

$$U \oplus W := U \times W = \{(u, w) \mid u \in U,\ \ w \in W\}.$$

Wir haben eine kanonische lineare Abbildung

$$f\colon U \oplus W \to V, \quad (u, w) \mapsto u + w.$$

Häufig wird $U \oplus W$ auch nur als Notation von $U + W$ verwendet, wenn $U \cap W = 0$.

Dimensionsformel für die Summe von Untervektorräumen

Satz. *Seien V ein K-Vektorraum, $U, W \subset V$ zwei Untervektorräume und*

$$f : U \oplus W \to V, \; (u, w) \mapsto u + w$$

die kanonische Abbildung. Dann gilt

$$\mathrm{Bild}(f) = U + W$$

und

$$Ker \; f \cong U \cap W$$

vermöge

$$g \colon U \cap W \to U \oplus W, \quad x \mapsto (x, -x).$$

Korollar. *$U, W \subset V$ seien Untervektorräume. Dann gilt:*

$$\dim U + \dim W = \dim(U \cap W) + \dim(U + W)$$

Beweis des Korollars. Es gilt $\dim(U \oplus W) = \dim U + \dim W$. Die Formel folgt aus der Dimensionsformel für lineare Abbildungen. $\square$

Beweis des Satzes. Bild $f = U + W$ ist klar.

Bild $g \subset \ker f$ ebenso, da $f(x, -x) = x - x = 0$.

Umgekehrt: Sind $(u, w) \in \ker\ f \subset U \oplus W$, dann gilt: $u + w = 0 \Rightarrow w = -u \in U \cap W$, also: $(u, w) = g(u)$ und deshalb $\ker f \subset \mathrm{Bild}\, g$, also:

$$\mathrm{Bild}\, g = \ker f.$$

Die Abbildung g ist also eine Epimorphismus

$$U \cap W \to \ker f.$$

Dieser ist ein Isomorphismus, da g auch injektiv ist. □

Beispiel. $V = \mathbb{R}^3$. Wir betrachten die beiden Untervektorräume:

$$U_t = \left\langle \begin{pmatrix} \cos t \\ \sin t \\ 0 \end{pmatrix} \right\rangle, \ W = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle.$$

Welche Dimensionen können für $U_t \cap W$ und $U_t + W$ auftreten?

Es gilt: $\dim U_t = 1,\ \dim W = 2\ \forall t$. Außerdem ist

$$\dim(U_t \cap W) = 0,\ \dim(U_t + W) = 3,$$

falls

$$\left\{ \begin{pmatrix} \cos t \\ \sin t \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

eine Basis des $\mathbb{R}^3$ ist, und dies passt zu den Formeln aus dem Korollar. Ist dies nicht der Fall, dann gilt:

$$\begin{pmatrix} \cos t \\ \sin t \\ 0 \end{pmatrix} \in \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\rangle = W \quad (U_t \subset W),$$

also: $\dim U_t = 1,\ \dim W = 2,\ \dim(U_t \cap W) = 1$ und $\dim(U_t + W) = 2$. Dies liegt vor, falls:

$$\begin{pmatrix} \cos t \\ \sin t \\ 0 \end{pmatrix} \in \left\langle \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \right\rangle \implies t = \frac{\pi}{2} \text{ bzw. } t = \frac{\pi}{2} + k\pi,\ k \in \mathbb{Z}.$$