Lernbuch Lineare Algebra und Analytische Geometrie, 2. Auflage 2012

Korrekturen

85^{15}	*statt*	$y \in M$ *lies* $y \in N$
100_{12}	*statt*	$m' + n' = m + (n') = m + (n')$ *lies* $m' + n' = m + (n')'$
102_5	*statt*	$\#\mathcal{P}(M)$ *lies* $\#\mathcal{P}_1(M)$
104_{17}	*statt*	Beispiel c) *lies* Beispiel i)
104_{14}	*statt*	M' *lies* M
109_6	*statt*	$n' + m'$ *lies* $n' + m''$
115_7	*ergänze*	und er heißt ***nullteilerfrei***,
		wenn aus $a \cdot b = 0$ folgt, dass $a = 0$ oder $b = 0$.
117_2	*statt*	n_1 *lies* n_2, *statt* n_2 *lies* n_1
122_{15}	*statt*	$a, b \in R$ *lies* $a, b \in R'$
136_5	*statt*	1) und 2) *lies* a) und b)
	statt	3) *lies* c)
137_3	*statt*	$\alpha - a_n$ *lies* $\alpha - a_N$
147^9	*statt*	$(\mathbf{i}t)$ *lies* $(\mathbf{i}t)^n$
147^{10}	*statt*	$\mathbf{i}\frac{t^4}{4!}$ *lies* $\frac{t^4}{4!}$
209^2	*statt*	$\sum_{i=1}^{l}$ *lies* $\sum_{i=1}^{k}$
209^4	*statt*	$\sum_{i=1}^{n}$ *lies* $\sum_{i=1}^{r}$
230^{13}	*statt*	$a_2 = 2$ *lies* $a_2 = -2$
255_{10}	*statt*	2.3.5 *lies* 2.3.4
$313^{3,5}$	*statt*	A_{ij} *lies* A'_{ij}
359_8	*statt*	$\mu^2 - \omega^2$ *lies* $\mu^2 - \omega^2 > 0$
427	*tausche*	die Seite aus durch 427 neu (sh. unten)
438^1	*statt*	Satz 1 *lies* Satz 2
447	*Einfügung*	5.3.8 neu (sh. unten)
459	*statt*	w *lies* ω (in der Abbildung)

$$ {}^tAAx^* = {}^tAb. \qquad (*) $$

Die Gleichungen des Systems $(*)$ werden ***Normalgleichungen*** zu den Gleichungen $(-)$ genannt. Ihre Eigenschaften ergeben sich aus dem folgenden

Lemma *Für jedes $A \in \mathrm{M}(m \times n, \mathbb{R})$ ist ${}^tA \cdot A \in \mathrm{M}(n \times n)$ symmetrisch und*

$$ \operatorname{rang}({}^tA \cdot A) = \operatorname{rang} A. $$

Vorsicht! Für beliebige Körper K ist die Aussage über den Rang nicht gültig. Man betrachte etwa $K = \mathbb{F}_2$ (siehe 1.3.7) mit

$$ {}^tA = (1,1) \quad \text{und} \quad {}^tA \cdot A = (0). $$

Beweis Die Symmetrie folgt sofort aus der Rechenregel für die Transposition in 2.4.4. Zur Berechnung des Ranges betrachten wir die linearen Abbildungen

$$ \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \xrightarrow{{}^tA} \mathbb{R}^n. $$

Nach der Dimensionsformel aus 2.3.5 genügt es zu zeigen, dass $\operatorname{Im}(A) \cap \operatorname{Ker}({}^tA) = \{0\} \subset \mathbb{R}^m$. Sei also

$$ w = {}^t(x_1, ..., x_m) \in \operatorname{Im}(A) \cap \operatorname{Ker}({}^tA). $$

Dann gibt es ein $v \in \mathbb{R}^n$ mit $w = Av$ und ${}^tAw = 0 \in \mathbb{R}^n$. Daraus folgt

$$ 0 = {}^tv({}^tAw) = {}^tv({}^tAAv) = {}^tww = x_1^2 + ... + x_m^2 \quad \text{und somit} \quad w = 0. $$

■

Nach Voraussetzung ist $\operatorname{rang} A = n$, also folgt aus dem Lemma ${}^tAA \in \mathrm{GL}(n; \mathbb{R})$. Somit ist das System $(*)$ eindeutig lösbar durch

$$ x^* = ({}^tAA)^{-1} \cdot {}^tAb. $$

Für die Norm des minimalen Fehlers ergibt sich unter Benutzung von $(*)$

$$ \|v(x^*)\|^2 = {}^t(Ax^* - b)(Ax^* - b) = {}^tbb - {}^tbAx^*. $$

Damit ist das oben formulierte Problem für die Theorie gelöst: Man multipliziert die im allgemeinen unlösbare Bedingung $Ax = b$ von links mit ${}^tA$ und erhält eine eindeutige beste Approximation.

5.3.8 Die Singulärwert-Zerlegung*

In 2.5.3 hatten wir für jede Matrix $A \in \mathrm{M}(m \times n; K)$ vom Rang r über einem beliebigen Körper K eine Normalform

$$D = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}(m \times n; K)$$

konstruiert. Das bedeutet, dass es Matrizen $S \in \mathrm{GL}(m; K)$ und $T \in \mathrm{GL}(n; K)$ gibt derart, dass

$$D = S \cdot A \cdot T^{-1} \qquad \text{d.h.} \qquad A = S^{-1} \cdot D \cdot T.$$

Die zweite Gleichung beschreibt eine Zerlegung von A. Betrachtet man die Matrizen als lineare Abbildungen, so ergibt das ein kommutatives Diagramm

$$\begin{tikzcd} K^n \arrow[r, "A"] \arrow[d, "T"'] & K^m \arrow[d, "S"] \\ K^n \arrow[r, "D"] & K^m. \end{tikzcd}$$

Im Spezialfall $K = \mathbb{R}$ sollte eine Normalform von A auch Informationen über die Längenveränderungen bei Anwendung der Abbildung A enthalten. Das ist dann möglich, wenn die Transformationsmatrizen S und T orthogonal gewählt werden. In D entsteht dann statt E_r eine Diagonalmatrix mit positiven Einträgen, den sogenannten „Singulärwerten" von A.

Nach diesen Vorbereitungen kommen wir zum

Satz über die Singulärwertzerlegung *Zu jeder Maxtrix $A \in \mathrm{M}(m \times n; \mathbb{R})$ vom Rang r gibt es orthogonale Matrizen $S \in \mathrm{O}(m)$ und $T \in \mathrm{O}(n)$, sowie ein*

$$D = \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} \in M(m \times n; \mathbb{R}) \quad \textit{mit} \quad \Sigma_r = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{pmatrix} \in \mathrm{M}(r \times r; \mathbb{R})$$

und $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r > 0$ derart, dass $A = {}^tS \cdot D \cdot T$.

*Die positiven Zahlen $\sigma_1, \ldots, \sigma_r$ sind dabei bis auf die Reihenfolge eindeutig bestimmt, sie heißen **Singulärwerte** von A.*

Beweis Der Kniff ist, die symmetrische Matrix

$${}^tA \cdot A \in \mathrm{M}(n \times n; \mathbb{R})$$

zu benutzen. Nach dem Spektralsatz aus 5.3.7 gibt es dazu eine Orthonormalbasis $(v_1,\dots,v_n)$ von $\mathbb{R}^n$ bestehend aus Eigenvektoren von ${}^tA \cdot A$ zu reellen Eigenwerten $\lambda_1,\dots,\lambda_n$. Es gilt für $j = 1,\dots,n$

$$\lambda_j = \langle v_j, \lambda_j v_j\rangle = {}^tv_j \cdot ({}^tAA)v_j = {}^t(Av_j)\cdot(Av_j) \geqslant 0. \tag{$*$}$$

Nach dem Lemma aus 5.3.4 ist $\operatorname{rang}({}^tAA) = \operatorname{rang} A = r$. Daher kann man die Eigenwerte so anordnen, dass

$$\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_r > 0 \quad \text{und} \quad \lambda_{r+1} = \dots = \lambda_n = 0.$$

Setzen wir $\sigma_j := \sqrt{\lambda_j} \geqslant 0$ für $j = 1,\dots,n$, so ist nach $(*)$

$$||Av_j||^2 = \lambda_j \quad \text{also} \quad ||Av_j|| = \sigma_j.$$

Für $j = 1,\dots,r$ setzen wir

$$w_j := \frac{1}{\sigma_j}Av_j, \quad \text{also ist} \quad ||w_j|| = 1 \quad \text{und} \quad Av_j = \sigma_j w_j.$$

Die Vektoren $w_1,\dots,w_r \in \mathbb{R}^m$ sind auch orthogonal, denn für $i \neq j$ ist

$$\sigma_i\sigma_j\langle w_i, w_j\rangle = {}^t(Av_i)(Av_j) = \langle v_i, \lambda_j v_j\rangle = \lambda_j\langle v_i, v_j\rangle = 0.$$

Nach dem Verfahren von GRAM-SCHMIDT wird die orthonormale Familie $(w_1,\dots,w_r)$ zu einer Orthonormalbasis $(w_1,\dots,w_m)$ von $\mathbb{R}^m$ ergänzt. Seien nun

$S \in \mathrm{O}(m)$ die Matrix mit den Vektoren $w_1,\dots,w_m$ als Zeilen, und $T \in \mathrm{O}(n)$ die Matrix mit den Vektoren $v_1,\dots,v_n$ als Zeilen.

Diese Matrizen ergeben ein kommutatives Diagramm linearer Abbildungen

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] \arrow[d, "T"'] & \mathbb{R}^m \arrow[d, "S"] \\ \mathbb{R}^n \arrow[r, "D"] & \mathbb{R}^m \end{tikzcd} \quad \text{wobei} \quad \begin{tikzcd} v_j \arrow[r, mapsto] \arrow[d, mapsto] & \sigma_j w_j \arrow[d, mapsto] \\ e_j \arrow[r, mapsto] & \sigma_j e_j \end{tikzcd} \quad \text{für } j = 1,..r.$$

Für $j = r+1,\dots,n$ ist $Av_j = 0$ und $De_j = 0$. Damit ist die Existenz von S und T bewiesen.

Die Eindeutigkeit der Singulärwerte ist leicht zu sehen: Ist $A = {}^tS \cdot D \cdot T$, so folgt

$${}^tA \cdot A = ({}^tT \cdot {}^tD \cdot S)({}^tS \cdot D \cdot T) = {}^tT({}^tD \cdot D)T.$$

Also sind die σ_j^2 die Eigenwerte von ${}^tA \cdot A$. ■

Beispiel Um den Rechenaufwand möglichst gering zu halten, betrachten wir eine sehr einfache Abbildung $A : \mathbb{R}^3 \to \mathbb{R}^2$, nämlich

$$A := \begin{pmatrix} 0 & 3 & 4 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{mit} \quad {}^tA \cdot A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 9 & 12 \\ 0 & 12 & 16 \end{pmatrix}.$$

Das charakteristische Polynom von ${}^tA \cdot A$ ist gleich

$$X^3 - 26X^2 + 25X = (X-25)(X-1)X,$$

Also ist $\lambda_1 = 25$, $\lambda_2 = 1$ und $\lambda_3 = 0$. Die zugehörige Orthonormalbasis des $\mathbb{R}^3$ ist gegeben durch

$$v_1 = \frac{1}{5}\begin{pmatrix} 0 \\ 3 \\ 4 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_3 = \frac{1}{5}\begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix}.$$

Mit Hilfe von $\sigma_1 = 5$ und $\sigma_2 = 1$ ergibt sich

$$w_1 = \frac{1}{\sigma_1} A v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad w_2 = \frac{1}{\sigma_2} A v_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Damit erhält man die Transformationsmatrizen

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad T = \frac{1}{5}\begin{pmatrix} 0 & 3 & 4 \\ 5 & 0 & 0 \\ 0 & 4 & -3 \end{pmatrix}, \quad \text{sowie} \quad D = S \cdot A \cdot {}^tT = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Zur Geometrie der Abbildung: $Av_1 = 5w_1$ und $Av_2 = w_2$ bedeutet, dass in Richtung von v_1 stark verändert wird, in Richtung von v_2 gar nicht. Der Vektor v_3 erzeugt den Kern von A.

Als erste Anwendung der Singulärwert-Zerlegung erhält man aus den Werten der σ_j präzise Informationen über die metrischen Eigenschaften der durch A definierten linearen Abbildung. Matrizen dienen aber auch als Speicher von Daten verschiedenster Art. Dabei können oft relativ kleine Singulärwerte $\sigma_{k+1} \geqslant ... \geqslant \sigma_r > 0$ zur Vereinfachung ohne großen Verlust von Informationen vernachlässigt werden. Dazu setzt man

$$\Sigma_k = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{pmatrix} \in \mathrm{M}(k \times k; \mathbb{R}) \quad \text{und} \quad D_k := \begin{pmatrix} \Sigma_k & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}(m \times n; \mathbb{R}).$$

Die Matrix $A_k := {}^tS \cdot D_k \cdot T$ hat dann nur noch den Rang k und vereinfacht A.

Setzen wir in obigem Beispiel $k = 1$, so wird

$$A_1 = {}^tS \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot T = \begin{pmatrix} 0 & 3 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

Mehr dazu findet man etwa bei [L-M, Kap. 19].