# §2. Die Abstandsfunktion $f_A(U): = \| A-U \|^2$

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

Band (Jahr): **32 (1986)**

Heft 1-2: **L'ENSEIGNEMENT MATHÉMATIQUE**

PDF erstellt am: **11.08.2024**

**Nutzungsbedingungen**
Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

**Haftungsausschluss**
Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

**http://www.e-periodica.ch**

Im Gegensatz zu der durch das Gram-Schmidt-Verfahren konstruierten hängt sie nicht von der Ordnung der Basis $A$ ab; genauer gilt:

SATZ 2. *$A$ und $U$ seien wie in Satz 1, $P$ eine Permutationsmatrix. Dann ist $\tilde{U} = UP$ die beste orthogonale Approximation von $\tilde{A} = AP$.*

*Beweis.* $\| \tilde{A} - \tilde{U} \| = \| AP - UP \| = \| A - U \|$, da $P$ ja als Permutationsmatrix orthogonal ist. Gäbe es ein $\tilde{\tilde{U}} \in O(n)$ mit $\| \tilde{A} - \tilde{\tilde{U}} \| < \| A - U \|$, so wäre $\tilde{\tilde{U}}P^*$ eine bessere Approximation von $A$ als $U$.

Ebenso einfach sind die Beweise der folgenden Eigenschaften dieser besten orthogonalen Approximation:

SATZ 3.

(1) *Die beste orthogonale Approximation einer positiv definiten symmetrischen Matrix ist die Einheitsmatrix.*

(2) *Ist $U$ die beste orthogonale Approximation von $A$, so ist $U^*$ diejenige von $A^*$.*

## § 2. DIE ABSTANDSFUNKTION $f_A(U) := \| A - U \|^2$

Sei $A \in GL(n, \mathbf{R})$ fest. Um das Minimum der Abstandsfunktion $f_A : O(n) \to \mathbf{R}$, $f_A(U) := \| A - U \|^2$ zu bestimmen, haben wir oben alle kritischen Punkte dieser Funktion bestimmt: Es sind genau diejenigen Matrizen $U \in O(n)$, für die $A = US$ ist mit symmetrischem $S$, so dass $S^2 = A^*A$ ist. Diese Gleichung hat genau dann endlich viele — und zwar $2^n$ — symmetrische Lösungen $S$, wenn $A^*A$ $n$ verschiedene Eigenwerte $0 < \lambda_1^2 < \dots < \lambda_n^2$ hat: Es sind die Matrizen

$$S = C \begin{pmatrix} \pm\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \pm\lambda_m \end{pmatrix} C^*,$$

wenn $A^*A = C \begin{pmatrix} \lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix} C^*$ ist, wobei die Spalten der orthogonalen Matrix $C$ die entsprechenden Eigenvektoren von $A^*A$ sind. Zur Vereinfachung verwenden wir folgende Bezeichnungen:

$$[\lambda]: = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \qquad \text{mit } 0 < \lambda_1 < \dots < \lambda_n,$$

$$[\varepsilon]: = \begin{pmatrix} \varepsilon_1 & & 0 \\ & \ddots & \\ 0 & & \varepsilon_n \end{pmatrix} \qquad \text{mit } \varepsilon_i = \pm 1, i = 1, \dots, n,$$

$S = S_\varepsilon = C[\varepsilon]\,[\lambda]C^*$ und $U = U_\varepsilon = AS^{-1}$. Dann gilt:

SATZ 4. *Hat $A^*A$ $n$ verschiedene Eigenwerte, so ist $f_A$ eine Morse-Funktion auf $O(n)$, und zwar hat der kritische Punkt $U_\varepsilon$ den Index $(i_1-1) + \dots + (i_k-1)$, wenn $\varepsilon_{i_1} = \dots = \varepsilon_{i_k} = -1$ ist und die restlichen $\varepsilon_i = +1$.*

*Beweis.* Zur Untersuchung der Funktion $f_A$ dürfen wir wegen Eigenschaft (1) des Skalarproduktes auf $\mathbf{R}^{n \times n}$ und wegen Satz 1 o.B.d.A. annehmen, dass $A = S$ positiv definit symmetrisch ist und ausserdem bereits in Diagonalform vorliegt, also $A = [\lambda]$ mit den oben eingeführten Bezeichnungen. Die kritischen Punkte von $f_A$ sind dann gerade die Matrizen $U_\varepsilon = [\varepsilon]$, wobei $f_A$ in $\mathbf{1}$ das Minimum annimmt. Ferner ist $S_\varepsilon = [\varepsilon]\,[\lambda]$.

Um $f_A$ in der Nähe des kritischen Punktes $U_\varepsilon$ zu untersuchen, beschränken wir die Funktion auf die Kurven $(U_\varepsilon \exp(tB))_{t \in \mathbf{R}}$, $B \in \mathbf{R}^{n \times n}_{\text{asym}}$. Eine einfache Rechnung ergibt

$$\| A - U_\varepsilon \exp(tB) \|^2 = \| A - [\varepsilon] \|^2 - t^2 \operatorname{tr}(S_\varepsilon B^2) + o(t^2),$$

so dass wir also die quadratische Form $Q: R^{n \times n}_{\text{asym}} \to \mathbf{R}$,

$$Q(B) = -\operatorname{tr}(S_\varepsilon B^2) = \operatorname{tr}(B^* S_\varepsilon B)$$

bzw. die zugehörige symmetrische Bilinearform untersuchen müssen. Dazu führen wir in $\mathbf{R}^{n \times n}_{\text{asym}}$ eine geeignete Basis ein:

$$B_{ij}: = E_{ij} - E_{ji}, 1 \leqslant i < j \leqslant n,$$

wobei $E_{ij}$ diejenige Matrix ist, die genau eine Eins in der $i$-ten Zeile an der $j$-ten Stelle hat und sonst lauter Nullen. Dann ist $\langle B_{ij}, B_{kl} \rangle = 2\delta_{ik}\delta_{jl}$, d.h. $(B_{ij})_{1 \leqslant i < j \leqslant n}$ ist eine Orthogonalbasis von $\mathbf{R}^{n \times n}_{\text{asym}}$. Diese Basis diagonalisiert die Form $Q$:

$$\begin{aligned} Q(B_{ij}, B_{kl}) &= -\mathrm{tr}(S_\varepsilon B_{ij} B_{kl}) = -\mathrm{tr}([\varepsilon]\,[\lambda] B_{ij} B_{kl}) \\ &= \delta_{ik}\delta_{jl}(\varepsilon_i\lambda_i + \varepsilon_j\lambda_j) - \delta_{il}\delta_{jk}(\varepsilon_i\lambda_i + \varepsilon_j\lambda_j) \\ &= \delta_{ik}\delta_{jl}(\varepsilon_i\lambda_i + \varepsilon_j\lambda_j)\,, \end{aligned}$$

da wegen $i < j$ und $k < l$ stets $\delta_{il}\delta_{jk} = 0$ ist.

Insbesondere ist also $Q(B_{ij}, B_{kl}) = 0$ für $(i, j) \neq (k, l)$, und $Q(B_{ij}, B_{ij}) = \varepsilon_i\lambda_i + \varepsilon_j\lambda_j \neq 0$ wegen unserer Voraussetzung $\lambda_i \neq \lambda_j$. Dabei ist $Q(B_{ij}, B_{ij}) < 0$ genau dann, wenn $\varepsilon_j = -1$ ist, da wir ja die Werte $\lambda_i$ so numeriert haben, dass $0 < \lambda_1 < \ldots < \lambda_n$ ist. Daraus ergibt sich sofort die angegebene Formel für den Index.

Als nächstes wollen wir untersuchen, welche Indizes bei den kritischen Punkten der Funktion $f_A$ auftreten. Wir nehmen dazu natürlich wieder an, dass $A^*A$ $n$ verschiedene Eigenwerte hat und beschränken uns ausserdem auf den Fall $\det A > 0$ und die Untersuchung von $f_A|_{SO(n)}$. Dort hat $f_A$ dann die $2^{n-1}$ kritischen Punkte $U_\varepsilon$ für $\varepsilon = (\pm 1, \ldots, \pm 1)$ mit $\varepsilon_1 \cdot \ldots \cdot \varepsilon_n = +1$. Zur Kennzeichnung dieser Punkte verwenden wir die Potenzmenge $P\{1, \ldots, n-1\}$, indem wir für $\alpha \subset \{1, \ldots, n-1\}$ $U_\alpha := U_{\varepsilon(\alpha)}$ setzen, wobei $\varepsilon(\alpha) := (\varepsilon_1, \ldots, \varepsilon_n)$ sei

$$\text{mit } \varepsilon_1 := (-1)^{\mathrm{card}(\alpha)}$$

$$\text{und } \varepsilon_i := \begin{cases} -1\,, & \text{falls } i-1 \in \alpha \\ +1\,, & \text{falls } i-1 \notin \alpha \end{cases} \qquad \text{für } i = 2, \ldots, n\,.$$

Bezeichnet $\nu(\alpha)$ den Index von $f_A$ im kritischen Punkt $U_\alpha$, so ist nach Satz 4 gerade $\nu(\alpha) = \sum_{i \in \alpha} i$. Insbesondere treten also alle Werte von $0 = \nu(\emptyset)$ bis $\frac{n(n-1)}{2} = \nu(\{1, \ldots, n-1\})$ auf, und zwar mit der Häufigkeit $\mu(\nu)$, die durch die erzeugende Funktion

$$P(x) = \sum_{\nu=0}^{\infty} \mu(\nu) x^\nu = \prod_{i=1}^{n-1} (1 + x^i)$$

beschrieben wird (vgl. [4]). $P(x)$ ist aber auch das Poincaré-Polynom von $SO(n)$ über $\mathbf{Z}/2$ (vgl. [2]). Damit haben wir das folgende Ergebnis:

Satz 5. *Hat $A \in GL(n, \mathbf{R})$ positive Determinante und hat $A^*A$ $n$ verschiedene Eigenwerte, so ist die Morse-Funktion $f_A: SO(n) \to \mathbf{R}$ perfekt.* (Vgl. [1].)

Insbesondere gibt es also auf $SO(n)$ keine Morse-Funktion mit weniger kritischen Punkten.