

# **EINIGE BEMERKUNGEN ÜBER DIE POLARE ZERLEGUNG EINER REGULÄREN MATRIX UND DIE GEOMETRIE DER ORTHOGONALEN GRUPPE**

Autor(en): **Rummler, Hansklaus**

Objektyp: **Article**

Zeitschrift: **L'Enseignement Mathématique**

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

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

PDF erstellt am: **25.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-55087>

## **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.

# EINIGE BEMERKUNGEN ÜBER DIE POLARE ZERLEGUNG EINER REGULÄREN MATRIX UND DIE GEOMETRIE DER ORTHOGONALEN GRUPPE

von Hansklaus RUMMLER

## § 0. EINFÜHRUNG

Bekanntlich lässt sich jede reguläre  $n \times n$ -Matrix  $A$  mit reellen Koeffizienten eindeutig als Produkt  $A = US$  schreiben, wobei  $U$  orthogonal und  $S$  positiv definit symmetrisch ist (vgl. [3]). Diese sogenannte polare Zerlegung hat eine ebenso einfache wie interessante geometrische Deutung:  $U$  ist die beste orthogonale Approximation von  $A$ , wenn man den Raum  $\mathbf{R}^{n \times n}$  aller reellen  $n \times n$ -Matrizen mit der üblichen euklidischen Struktur versieht. Man kann das auch so ausdrücken: Fasst man die Spalten der Matrix  $A = (a_1, \dots, a_n)$  als Basis des  $\mathbf{R}^n$  auf, so ist  $U = (u_1, \dots, u_n)$  diejenige eindeutig bestimmte Orthonormalbasis, für die  $\sum_{j=1}^n \|a_j - u_j\|^2$  minimal ist.

Darüber hinaus zeigt sich, dass für festes  $A \in GL(n, \mathbf{R})$  die Funktion  $f_A: O(n) \rightarrow \mathbf{R}$ ,  $f_A(U) := \|A - U\|^2$  Aufschlüsse über die Geometrie von  $O(n)$  gibt: Es ist fast immer eine perfekte Morse-Funktion.

Eine Übertragung der Resultate auf komplexe Matrizen bereitet keine Schwierigkeiten, aber der Einfachheit halber beschränken wir uns hier auf den reellen Fall; aus dem gleichen Grunde verzichten wir auch darauf, singuläre Matrizen zu betrachten.

## § 1. DIE BESTE ORTHOGONALE APPROXIMATION EINER REGULÄREN MATRIX

Mit  $\mathbf{R}^{n \times n}$  bezeichnen wir den reellen Vektorraum aller reellen  $n \times n$ -Matrizen.  $1$  ist die Einheitsmatrix, und  $A^*$  die Transponierte der Matrix  $A$ . Wir versehen  $\mathbf{R}^{n \times n}$  mit dem Standard-Skalarprodukt

$$\langle A, B \rangle := \operatorname{tr}(A^*B) = \sum_{i,j=1}^n A_{ij}B_{ij}.$$

Folgende Eigenschaften dieses Skalarproduktes sind nahezu trivial:

- (1) Ist  $U$  orthogonal, so sind Links- und Rechts-Multiplikation mit  $U$  Isometrien.
- (2) Das Transponieren ist eine Isometrie.
- (3)  $\mathbf{R}^{n \times n}$  ist die orthogonale direkte Summe der beiden Unterräume der symmetrischen bzw. antisymmetrischen Matrizen:

$$\mathbf{R}^{n \times n} = \mathbf{R}_{\text{sym}}^{n \times n} \oplus \mathbf{R}_{\text{asym}}^{n \times n}.$$

Sei jetzt  $A \in GL(n, \mathbf{R})$  eine feste reguläre  $n \times n$ -Matrix. Wir suchen eine orthogonale Matrix  $U \in O(n)$ , für die der Abstand  $\|A - U\|$  minimal ist. Bezeichnet  $T_U O(n)$  den Tangentialraum an  $O(n)$  in  $U$ ,

$$T_U O(n) = \{UB; B \in \mathbf{R}_{\text{asym}}^{n \times n}\},$$

so gilt:  $\|A - U\|$  minimal  $\Rightarrow A - U \perp T_U O(n)$

$$\Leftrightarrow \langle A - U, UB \rangle = 0 \quad \text{für alle } B \in \mathbf{R}_{\text{asym}}^{n \times n}$$

$$\Leftrightarrow \langle U^*A - \mathbf{1}, B \rangle = 0 \quad \text{für alle } B \in \mathbf{R}_{\text{asym}}^{n \times n}$$

$$\Leftrightarrow U^*A - \mathbf{1} \in \mathbf{R}_{\text{sym}}^{n \times n}$$

$$\Leftrightarrow U^*A \in \mathbf{R}_{\text{sym}}^{n \times n}$$

$$\Leftrightarrow A = US \text{ mit } S \in \mathbf{R}_{\text{sym}}^{n \times n}.$$

Dabei gilt offensichtlich noch  $S^2 = A^*A$ . Ist umgekehrt  $S$  symmetrisch und gilt  $S^2 = A^*A$ , so ist  $U = AS^{-1}$  orthogonal.

Damit haben wir eine notwendige Bedingung für die Minimalität des Abstandes  $\|A - U\|$ : Es muss  $A = US$  sein mit einer symmetrischen Matrix  $S$ , die die Gleichung  $S^2 = A^*A$  erfüllt.

Als nächstes bestimmen wir alle diese Matrizen  $S$ : Da  $A$  regulär ist, ist die symmetrische Matrix  $A^*A$  positiv definit, und wir können ihre verschiedenen Eigenwerte in der Form  $\lambda_1^2, \dots, \lambda_r^2$  schreiben mit  $\lambda_1, \dots, \lambda_r > 0$ .  $\mu_1, \dots, \mu_r$  seien die entsprechenden Multiplizitäten, und für  $i = 1, \dots, r$  sei

$$\Lambda_i = \begin{pmatrix} \lambda_i & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_i \end{pmatrix} \in \mathbf{R}^{\mu_i \times \mu_i}.$$

Dann können wir  $A^*A$  in der Form

$$A^*A = C \begin{pmatrix} \Lambda_1^2 & & 0 \\ & \ddots & \\ 0 & & \Lambda_r^2 \end{pmatrix} C^*$$

schreiben, wo  $C$  eine orthogonale Matrix ist, deren Spalten Eigenvektoren von  $A^*A$  sind. Folglich ist

$$S = C \begin{pmatrix} S_1 & & 0 \\ & \ddots & \\ 0 & & S_r \end{pmatrix} C^*$$

mit  $S_i^2 = \Lambda_i^2$ , d.h.  $S_i = \lambda_i U_i$ , wobei  $U_i$  orthogonal und symmetrisch ist. Für  $A = US$  ist in diesem Falle

$$\begin{aligned} \|A - U\|^2 &= \|A\|^2 - 2\langle A, U \rangle + \|U\|^2 \\ &= \|A\|^2 - 2\operatorname{tr}(A^*U) + n \\ &= \|A\|^2 + n - 2\operatorname{tr}S \\ &= \|A\|^2 + n - 2 \sum_{i=1}^r \lambda_i \operatorname{tr}U_i. \end{aligned}$$

Nun ist aber  $\operatorname{tr}U_i \leq \mu_i$ , wobei Gleichheit genau für  $U_i = \mathbf{1}$  gilt. Damit ist  $\|A - U\|^2 \geq \|A\|^2 + n - 2 \sum_{i=1}^r \lambda_i \mu_i$ , und der kleinste Wert wird genau für  $U_1 = \mathbf{1}, \dots, U_r = \mathbf{1}$  erreicht, d.h. für

$$S = C \begin{pmatrix} \Lambda_1 & & \\ & \ddots & \\ & & \Lambda_r \end{pmatrix} C^*,$$

die einzige positiv definite symmetrische Lösung der Gleichung  $S^2 = A^*A$ . Damit haben wir gezeigt:

**SATZ 1.** Zu  $A \in GL(n, \mathbf{R})$  gibt es genau eine orthogonale Matrix  $U \in O(n)$ , für die der Abstand  $\|A - U\|$  minimal ist. Sie ist eindeutig bestimmt durch die polare Zerlegung  $A = US$  mit  $U$  orthogonal und  $S$  positiv definit symmetrisch.

Wie wir bereits in der Einführung bemerkt haben, kann man  $U$  als diejenige Orthonormalbasis von  $\mathbf{R}^n$  interpretieren, die  $A$  am nächsten liegt.

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  $f_A: O(n) \rightarrow \mathbf{R}$ ,  $f_A(U) := \|A - U\|^2$  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

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

wenn  $A^*A = C \begin{pmatrix} \lambda_1^2 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n^2 \end{pmatrix} C^*$  ist, wobei die Spalten der orthogonalen

Matrix  $C$  die entsprechenden Eigenvektoren von  $A^*A$  sind. Zur Vereinfachung verwenden wir folgende Bezeichnungen:

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

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

$S = S_\varepsilon = C[\varepsilon][\lambda]C^*$  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  $(U_\varepsilon \exp(tB))_{t \in \mathbf{R}}$ ,  $B \in \mathbf{R}_{\text{asym}}^{n \times n}$ . Eine einfache Rechnung ergibt

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

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

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

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

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

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.  $(B_{ij})_{1 \leq i < j \leq n}$  ist eine Orthogonalbasis von  $\mathbf{R}_{\text{asym}}^{n \times n}$ . Diese Basis diagonalisiert die Form  $Q$ :

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

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 nummeriert haben, dass  $0 < \lambda_1 < \dots < \lambda_n$  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  $\varepsilon = (\pm 1, \dots, \pm 1)$  mit  $\varepsilon_1 \cdot \dots \cdot \varepsilon_n = +1$ . Zur Kennzeichnung dieser Punkte verwenden wir die Potenzmenge  $P\{1, \dots, n-1\}$ , indem wir für  $\alpha \subset \{1, \dots, n-1\}$   $U_\alpha := U_{\varepsilon(\alpha)}$  setzen, wobei  $\varepsilon(\alpha) := (\varepsilon_1, \dots, \varepsilon_n)$  sei

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

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

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

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

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  $f_A: SO(n) \rightarrow \mathbf{R}$  perfekt. (Vgl. [1].)*

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



$SO(n)$  zu untersuchen und zu zeigen, dass diese Kurve die kürzeste Verbindung zwischen  $\mathbf{1}$  und  $\mathbf{1} + 2B_{ij}$  darstellt. Sei dazu  $V$  eine beste orthogonale Approximation von  $M_{ij}$ . ( $V$  ist nicht eindeutig, da  $M_{ij}$  nicht regulär ist.) Dann besitzt  $M_{ij}$  die polare Zerlegung  $M_{ij} = VT$ , wobei  $T$  symmetrisch ist und  $T^2 = M_{ij}^* M_{ij} = M_{ij}$  gilt, also

$$\begin{aligned} \|M_{ij} - V\|^2 &= \|T - \mathbf{1}\|^2 = \|T\|^2 - 2\operatorname{tr}T + \|\mathbf{1}\|^2 \\ &= \|M_{ij}\|^2 - 2\operatorname{tr}T + \|\mathbf{1}\|^2 = n - 2 - 2\operatorname{tr}T + n \\ &= 2n - 2 - 2\operatorname{tr}T. \end{aligned}$$

Numerieren wir die Standardbasis des  $\mathbf{R}^n$  in geeigneter Weise um, so ist

$$T^2 = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & 0 \end{pmatrix}, \text{ also } T = \begin{pmatrix} T_1 & & 0 \\ & 0 & \\ 0 & & 0 \end{pmatrix}$$

mit  $T_1 \in O(n-2)$  symmetrisch, und damit  $\operatorname{tr}T \leq n-2$ , wobei Gleichheit genau für  $T_1 = \mathbf{1}$  gilt. Folglich ist  $\|M_{ij} - V\|^2 \geq 2n - 2 - 2n + 4 = 2$ , und jeder Weg in  $SO(n)$  zwischen  $\mathbf{1}$  und  $\mathbf{1} + 2B_{ij}^2$  verläuft ausserhalb der offenen Kugel vom Radius  $\sqrt{2}$  um  $M_{ij}$  in  $\mathbf{R}^{n \times n}$  und ist damit mindestens so lang wie der Kreisbogen  $(\exp(tB_{ij}))_{0 \leq t \leq \pi}$ .

#### LITERATUR

- [1] BOTT, R. Lectures on Morse Theory, old and new. *Bull. Amer. Math. Soc.* 7 (1982), 331-358.
- [2] EHRESMANN, C. Sur la topologie des groupes simples clos. *C. R. Acad. Sci. Paris* 208 (1939), 1263-1265.
- [3] HALMOS, P. R. *Finite-dimensional Vector Spaces*, 2nd ed., Princeton, N. J.: Van Nostrand (1958).
- [4] JEGER, M. *Einführung in die Kombinatorik*, Bd. 2, Stuttgart: Klett (1976).

(Reçu le 5 août 1985)

Hansklus Rummler

Institut de Mathématiques  
de l'Université de Fribourg  
CH-1700 Fribourg