

§1. Die beste orthogonale Approximation einer regulären Matrix

Objekttyp: **Chapter**

Zeitschrift: **L'Enseignement Mathématique**

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

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

PDF erstellt am: **19.09.2024**

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.

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 $\lambda_1, \dots, \lambda_r > 0$. μ_1, \dots, μ_r seien die entsprechenden Multiplizitäten, und für $i = 1, \dots, r$ sei

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

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: