Mathematik für InformatikerInnen 2

Frank-Olaf Schreyer

Universität des Saarlandes, SS 2020

Positive definite Matrizen und Normalformen von orthogonalen und unitären Matrizen

Die Themen heute sind

- Hauptachsentransformation für hermitesche Matrizen
- Normalformen von unitären und orthogonalen Matrizen
- Positiv definite Matrizen
- Hurwitz-Kriterium für positiv definit

Die Hauptachsentransformation funktioniert auch für hermitesche Matrizen. Wir können damit die Frage beantworten, welche hermitesche Matrizen ein Skalarprodukt definieren.

Hauptachsentransformation

Satz. *Sei $B \in \mathbb{C}^{n\times n}$ eine hermitesche Matrix. Dann existiert eine unitäre Matrix $T \in \mathrm{SU}(n)$, so dass*

$$TB\overline{T}^t = TBT^{-1} = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix mit reellen Einträgen ist.

Beweis. Wir betrachten wieder einen Eigenwert λ_1 von B und einen zugehörigen normierten Eigenvektor $v_1 \in \mathbb{C}^n$. Das orthogonale Komplement

$$H = v_1^\perp$$

ist ein Untervektorraum, der von B in sich abgebildet wird. Rekursion greift: Es existiert eine Orthonomalbasis $\{v_2, \ldots, v_n\}$ von H aus Eigenvektoren von B.

$$\overline{T}^t = (v_1\, v_2 \ldots v_n) \in \mathrm{U}(n)$$

ist dann eine unitäre Matrix, die $TB\overline{T}^t = D$ erfüllt. Um eine Matrix in $\mathrm{SU}(n)$ zu bekommen, können wir v_1 durch $\overline{\mu}v_1$ ersetzen, wobei $\mu = \det \overline{T}^t$ ist. □

Um unitäre Transformationen geometrisch zu verstehen, betrachten wir deren Jordansche Normalform.

Satz. *Sei $A \in U(n)$ eine unitäre Matrix. Die Eigenwerte λ_j von A haben alle den Betrag $|\lambda_j| = 1$, und es existiert eine unitäre Transformationsmatrix $T \in \mathrm{SU}(n)$, so dass*

$$TA\overline{T}^t = TAT^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

gilt.

Beweis. Die Beweisidee ist dieselbe wie bei der Hauptachsentransformation. Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A \in U(n) \subset \mathrm{GL}(n, \mathbb{C})$ und v ein normierter Eigenvektor. Da A eine Isometrie von $\mathbb{C}^n$ beschreibt, gilt

$$1 = \langle v, v \rangle = \langle Av, Av \rangle = \langle \lambda v, \lambda v \rangle = \overline{\lambda}\lambda \langle v, v \rangle = \overline{\lambda}\lambda.$$

λ hat also den Betrag $|\lambda| = 1$. Wegen

$$\langle v, w \rangle = \langle Av, Aw \rangle = \langle \lambda v, Aw \rangle = \lambda \langle v, Aw \rangle$$

bildet A das orthogonale Komplement

$$H = v^{\perp} = \{ w \in \mathbb{C}^n \mid \langle v, w \rangle = 0 \}$$

in sich ab, und Induktion greift. Um am Ende eine Transformationsmatrix $T \in \mathrm{SU}(n)$ statt nur in $U(n)$ zu bekommen, kann man einen der Eigenvektoren, die die Spalten von $\overline{T}^t$ bilden mit einem geeigneten $\lambda \in \mathrm{U}(1)$ multiplizieren.

Ende Teil 1

Normalformen von orthogonalen Transformationen

Satz. *Sei* $A \in \mathrm{O}(n) \subset \mathrm{GL}(n, \mathbb{R})$. *Dann existiert eine Transformation* $S \in \mathrm{SO}(n)$, *so dass* $SAS^t = SAS^{-1}$ *die Gestalt*

$$\begin{pmatrix} \cos\alpha_1 & -\sin\alpha_1 & & & & & & 0 \\ \sin\alpha_1 & \cos\alpha_1 & & & & & & \\ & & \ddots & & & & & \\ & & & \cos\alpha_r & -\sin\alpha_r & & & \\ & & & \sin\alpha_r & \cos\alpha_r & & & \\ & & & & & \pm 1 & & \\ & & & & & & \ddots & \\ 0 & & & & & & & \pm 1 \end{pmatrix}$$

hat. Dabei ist r *die Anzahl der Paare von komplex konjugierten Eigenwerten von* A *und* $n - 2r$ *die Anzahl der reellen Eigenwerte.*

Beweis. Orthogonale Matrizen sind auch unitär, da
$E = S\overline{S}^t = S\overline{S}^t$ gilt. ($\Rightarrow \mathrm{O}(n) \subset \mathrm{U}(n)$.)
Folglich haben sämtliche komplexe Eigenwerte von A den Betrag 1.
Die reellen Eigenwerte können also nur ± 1 sein. Andererseits hat das charakteristische Polynom $\chi_A(t)$ reelle Koeffizienten. Die komplexen Nullstellen treten daher in Paaren $\lambda, \overline{\lambda}$ von zueinander konjugiert komplexen Zahlen auf. Da die Beträge 1 sind, gilt

$$\lambda = \cos\alpha + i\sin\alpha, \overline{\lambda} = \cos\alpha - i\sin\alpha$$

für gewisse Winkel $\alpha \in]0, 2\pi[$ mit $\alpha \neq \pi$.
Sei $v \in \mathbb{C}^n$ ein Eigenvektor zu λ mit $\|v\| = \sqrt{2}$. Wir zerlegen

$$v = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + i \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \operatorname{Re} v + i \operatorname{Im} v$$

komponentenweise in Real- und Imaginärteil. Dann ist
$\overline{v} = \operatorname{Re} v - i \operatorname{Im} v$ ein Eigenvektor von A zum Eigenwert $\overline{\lambda}$:

$$(A - \lambda E)v = 0 \Leftrightarrow (A - \overline{\lambda}E)\overline{v} = 0.$$

$\operatorname{Re} v$ und $\operatorname{Im} v \in \mathbb{R}^n$ sind $\mathbb{R}$-linear unabhängig, da v und $\overline{v} \in \mathbb{C}^n$ als Eigenvektoren von zueinander verschiedenen Eigenwerten $\mathbb{C}$-linear unabhängig sind. $U = \operatorname{Spann}(\operatorname{Re} v, \operatorname{Im} v) \subset \mathbb{R}^n$ ist also ein 2-dimensionaler Untervektorraum, der von A in sich abgebildet wird:

$$\begin{aligned}
A \operatorname{Re} v &= A\frac{1}{2}(v + \overline{v}) = \frac{1}{2}(\lambda v + \overline{\lambda}\overline{v}) \\
&= \frac{1}{2}((\cos\alpha + i\sin\alpha)(\operatorname{Re} v + i \operatorname{Im} v) \\
&\qquad +(\cos\alpha - i\sin\alpha)(\operatorname{Re} v - i \operatorname{Im} v)) \\
&= \cos\alpha \operatorname{Re} v - \sin\alpha \operatorname{Im} v
\end{aligned}$$

$$\begin{aligned}
A \operatorname{Im} v &= A\frac{1}{2i}(v - \overline{v}) = \frac{1}{2i}(\lambda v - \overline{\lambda}\overline{v}) \\
&= \frac{1}{2i}((\cos\alpha + i\sin\alpha)(\operatorname{Re} v + i \operatorname{Im} v) \\
&\qquad -(\cos\alpha - i\sin\alpha)(\operatorname{Re} v - i \operatorname{Im} v)) \\
&= \sin\alpha \operatorname{Re} v + \cos\alpha \operatorname{Im} v
\end{aligned}$$

Bzgl. der Basis $\{\operatorname{Re} v, \operatorname{Im} v\}$ wird also A eingeschränkt auf U durch die Matrix

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

dargestellt. Wir zeigen, dass $\{\operatorname{Re} v, \operatorname{Im} v\}$ eine Orthonormalbasis von $U \subset \mathbb{R}^n$ ist. Dies folgt aus $\langle \overline{v}, v\rangle_{\mathbb{C}} = 0$ und $\langle v, v\rangle_{\mathbb{C}} = 2$. Die erste Gleichung besagt

$$0 = \sum_{j=1}^{n} \overline{\overline{z}}_j z_j = \sum_{j=1}^{n} z_j^2 = \sum_{j=1}^{n} x_j^2 - \sum_{j=1}^{n} y_j^2 + 2i \sum_{j=1}^{n} x_j y_j$$

$$\Rightarrow \langle \operatorname{Re} v, \operatorname{Re} v\rangle_{\mathbb{R}} = \langle \operatorname{Im} v, \operatorname{Im} v\rangle_{\mathbb{R}} \text{ und } \langle \operatorname{Re} v, \operatorname{Im} v\rangle_{\mathbb{R}} = 0.$$

Die zweite besagt

$$2 = \langle \operatorname{Re} v, \operatorname{Re} v\rangle_{\mathbb{R}} + \langle \operatorname{Im} v, \operatorname{Im} v\rangle_{\mathbb{R}}.$$

Insgesamt folgt

$$\|\operatorname{Re} v\| = 1 = \|\operatorname{Im} v\| \text{ und } \operatorname{Re} v \perp \operatorname{Im} v.$$

Um nun den Beweis abzuschließen, verwenden wir wieder Rekursion. Sei

$$W = U^{\perp} = \{w \in \mathbb{R}^n \mid \langle w, u \rangle = 0 \; \forall u \in U\}.$$

Auch W wird von A in sich abgebildet:

$$w \in W \implies \langle Aw, \tilde{u} \rangle = \langle Aw, Au \rangle = \langle w, u \rangle = 0 \quad \forall u \in U \text{ und } \tilde{u} = Au.$$

Da $U \xrightarrow{A} U$ ein Isomorphismus ist, folgt $Aw \in U^{\perp} = W$. Auf $W \xrightarrow{A} W$ können wir eine neue Matrixdarstellung berechnen, indem wir im ersten Schritt mit Hilfe des Gram-Schmidt-Verfahrens eine Orthonormalbasis $\mathcal{B}'$ von $W \cong \mathbb{R}^{n-2}$ bestimmen. In der Basis $\mathcal{B} = \{\mathrm{Re}, \mathrm{Im}\, v\} \cup \mathcal{B}'$ hat A die Gestalt

$$M_{\mathcal{B}}^{\mathcal{B}}(A) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & A' \end{pmatrix},$$

wobei $A' \in \mathrm{O}(n-2)$ eine orthogonale $(n-2) \times (n-2)$ Matrix ist.

Es ist uns also gelungen, einen 2×2 Rotationsblock abzuspalten. Wir können dies solange wiederholen bis die verbleibende orthogonale $(n-2r) \times (n-2r)$ Matrix $A'' \in \mathrm{O}(n-2r)$ keine komplexen Eigenwerte, genauer keine Eigenwerte in $\mathbb{C} \setminus \mathbb{R}$, hat. Schließlich bestimmen wir mit Gram-Schmidt noch Orthonormalbasen von

$$\mathrm{Eig}(A, 1) \text{ und } \mathrm{Eig}(A, -1).$$

Um am Ende eine Transformationsmatrix $S \in \mathrm{SO}(n)$ zu bekommen, können wir, wenn nötig, einen Eigenvektor zu ± 1 mit -1 multiplizieren, oder bei einem komplexen Paar von Eigenwerten λ und $\overline{\lambda}$ vertauschen, was $\operatorname{Im} v$ durch $-\operatorname{Im} v$ ersetzt. $\square$

Bemerkung. Eine orthogonale Matrix $A \in \mathrm{O}(n)$ liegt in $\mathrm{SO}(n)$ genau dann, wenn $\dim \mathrm{Eig}(A, -1)$ gerade ist. Wir können nämlich

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos \pi & -\sin \pi \\ \sin \pi & \cos \pi \end{pmatrix}$$

als eine Drehung um 180^o auffassen.

Ende Teil 2

Bemerkung. Die Berechnung der Normalform einer orthogonalen Matrix A ist mit einem erheblichen Aufwand verbunden. Ein schwieriger Punkt ist dabei, die Nullstellen des charakteristischen Polynoms $\chi_A(t)$ explizit zu bestimmen. Wir haben dazu eigentlich nichts gesagt.
Bei einer Drehung um 45^o hat die Drehmatrix die Einträge

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix},$$

bei 60^o haben wir

$$\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{3} & -\frac{\sqrt{3}}{2} \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} \end{pmatrix}.$$

Der Drehung um 15^o entspricht daher die Matrix

$$\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{\sqrt{2}+\sqrt{6}}{4} & -\frac{\sqrt{6}-\sqrt{2}}{4} \\ \frac{\sqrt{6}-\sqrt{2}}{4} & \frac{\sqrt{2}+\sqrt{6}}{4} \end{pmatrix}.$$

Oben haben wir $\mathrm{SO}(n) \subset \mathsf{SU}(n)$ verwendet. Es gibt auch eine Inklusion

$$\mathsf{SU}(n) \subset \mathrm{SO}(2n).$$

Diese beruht auf einer anderen Betrachtung des Körper der komplexen Zahlen.

$$(\mathbb{C}, +, \cdot) \cong (\{\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \mid x, y \in \mathbb{R}\}, +, \cdot),$$

$$z = x + iy \mapsto \begin{pmatrix} x & -y \\ y & x \end{pmatrix}.$$

Die Abbildung

$$\mathsf{SU}(n) \hookrightarrow \mathrm{SO}(2n)$$

definieren wir, indem jeden Eintrag z_{kj} von $T = (z_{kj})$ durch die entsprechende reelle 2×2 Matrix ersetzen.

Positiv definite Matrizen

Satz. *Sei $A \in \mathbb{C}^{n \times n}$ eine hermitesche Matrix. Die folgenden Bedingungen sind äquivalent:*

1) *Alle Eigenwerte von A sind strikt positiv.*
2) $\overline{z}^t Az > 0 \quad \forall z \in \mathbb{C}^n \backslash \{0\}$.
3) *Durch $\langle z, w \rangle_A = \overline{z}^t Aw$ wird ein Skalarprodukt auf $\mathbb{C}^n$ definiert.*

Definition. Eine hermitesche Matrix A, die die äquivalenten Bedingungen des Satzes erfüllt, nennt man **positiv definit**, in Zeichen $A > 0$. Eine symmetrische reelle Matrix $A \in \mathbb{R}^{n \times n}$ ist positiv definit, wenn sie aufgefasst als hermitesche Matrix positiv definit ist.

Beweis. Wir müssen die Äquivalenz der drei Aussagen zeigen.

1) $\Rightarrow$ 2) Nach dem Satz über die Hauptachsentransformation existiert eine Orthonormalbasis $\{v_1, \dots, v_n\}$ aus Eigenvektoren von A. Stellen wir $z = c_1 v_1 + \dots c_n v_n$ in dieser Basis da, so folgt

$$
\begin{aligned}
\overline{z}^t A z &= \Big(\sum_{j=1}^n \overline{c_j v_j}^t\Big) A \Big(\sum_{k=1}^n c_k v_k\Big) \\
&= \sum_{j,k=1}^n \overline{c_j} c_k \overline{v_j}^t \lambda_k v_k \\
&= \sum_{j,k=1}^n \overline{c_j} c_k \lambda_k \delta_{jk}, \qquad \text{da } \overline{v_j}^t \cdot v_k = \delta_{jk} \\
&= \sum_{k=1}^n |c_k|^2 \lambda_k > 0
\end{aligned}
$$

da alle $\lambda_k > 0$ und wenigstens ein $c_k \neq 0$.

2) $\Leftrightarrow$ 3) Wir wissen bereits, dass durch $b(z,w) = \overline{z}^t Aw$ eine Sesquilinearform definiert wird, die $\overline{b(w,z)} = b(z,w)$ erfüllt, da $\overline{A}^t = A$ gilt. Die einzige fehlende Bedingung, damit b ein Skalarprodukt definiert, ist die Bedingung S4, welche mit 2) übereinstimmt.

2) $\Rightarrow$ 1) Sei λ ein Eigenwert von A und $z \in \mathbb{C}^n$ ein zugehöriger Eigenvektor. Dann gilt nach 2)

$$0 < \overline{z}^t Az = \overline{z}^t (\lambda z) = \lambda \|z\|^2,$$

also $\lambda > 0$, da $\|z\|^2 > 0$ ist. □

Eine hermitesche Matrix bzw. eine reelle symmetrische Matrix A heißt **negativ definit**, wenn $-A > 0$. A heißt positiv semi-definit, wenn alle Eigenwerte nicht negativ sind. Gibt es sowohl negative wie auch positive Eigenwerte, nennt man A **indefinit**.

Ende Teil 3

Hurwitz-Kriterium

Satz. *Eine hermitesche Matrix* $A = (a_{kj} \in \mathbb{C}^{n\times n}$ *ist positiv definit genau dann, wenn*

$$\det \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0 \qquad \forall\ k = 1, 2, \dots, n.$$

gilt.

Für

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & & a_{1n} \\ a_{21} & a_{22} & a_{13} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & & a_{3n} \\ & \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

müssen die Determinanten der eingezeichneten Untermatrizen bestimmt werden.

Beispiel. Es gilt

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} > 0$$

nach dem Kriterium, da $2 > 0, \quad 4 - 1 = 3 > 0,$
$8 - 2 - 2 = 4 > 0$ und (Entwicklung nach der ersten Zeile)
$2 \cdot 4 - (-1)(-1)3 = 5 > 0.$

Der Satz von Gerschgorin gibt für diese Matrix nur, dass sie positiv semi-definit ist. Alle Eigenwerte liegen in dem Kreis

$$\{z \in \mathbb{C} \mid |z - 2| \leqslant 2\}.$$

Da sie reell sind, liegen sie also in dem Intervall $[0, 4]$.

Beweis. Es sei $A > 0$. Dann sind nach Definition auch alle Untermatrizen

$$A_k := \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix} > 0.$$

Bedingung 2) ist nämlich auch für $(z', 0) = (z_1, \dots, z_k, 0, \dots, 0)$ erfüllt und

$$\overline{z'} A_k (z')^t = \overline{(z', 0)} A (z', 0)^t > 0.$$

Auch diese Untermatrizen haben also nur positive Eigenwerte. Da die Determinante einer hermiteschen Matrix das Produkt ihrer Eigenwerte ist, ist $\det A_k > 0$. Die Bedingung ist also notwendig.

Wir müssen noch zeigen, dass sie auch hinreichend ist. Dazu verwenden wir Induktion nach n. Der Fall $n = 1$ ist klar. Für den Induktionsschritt von $n - 1 \to n$ betrachten wir die Matrix A_{n-1}, welche nach der Induktionsvoraussetzung positiv definit ist.

Nach der Hauptachsentransformation existiert eine unitäre Matrix $S \in \mathrm{U}(n-1)$, so dass $SA_{n-1}\overline{S}^t$ eine Diagonalmatrix mit Einträgen $\lambda_1', \ldots, \lambda_{n-1}' > 0$.

$$\underbrace{\begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix}}_{:=T} \cdot A \cdot \begin{pmatrix} \overline{S}^t & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda_1' & & 0 & b_1 \\ & \ddots & & \vdots \\ 0 & & \lambda_{n-1}' & b_{n-1} \\ \overline{b_1} & \ldots & \overline{b_{n-1}} & c \end{pmatrix} =: A'$$

für gewisse $b_i \in \mathbb{C}, i = 1, 2, \ldots, n-1$, $c \in \mathbb{C}$. Wir betrachten nun:

$$T = \begin{pmatrix} 1 & 0 & \frac{-b_1}{\lambda_1'} \\ & \ddots & \vdots \\ 0 & & 1 \end{pmatrix} \in \mathrm{SL}(n, \mathbb{C}).$$

Es gilt

$$\overline{T}^t A' T = \begin{pmatrix} \lambda_1' & & & 0 \\ & \ddots & & \\ & & \lambda_{n-1}' & \\ 0 & & & c' \end{pmatrix} =: D.$$

Da $0 < \det A = \det A' = \det D$, $\det T = \det \overline{T}^t = 1$ und $\lambda'_1, \ldots, \lambda'_{n-1} > 0$, folgt $c' > 0$.

Für $w = (w_1, \ldots, w_n)^t \neq 0$ gilt daher:

$$\overline{w}^t Dw = \sum_{i=1}^{n-1} w_i^2 \cdot \lambda'_i + w_n^2 c' > 0.$$

Insgesamt ergibt sich demnach:

$$0 < \overline{w}^t Dw = \overline{w}^t \overline{T}^t S' \cdot A \cdot \overline{S'}^t Tw.$$

Nun ist aber $\overline{S'}^t T \in \mathrm{GL}(n, \mathbb{C})$, d.h., $\forall z \neq 0 \; \exists \; w \neq 0$ mit $z = \overline{S'}^t Tw$, also:

$$\overline{z}^t \cdot A \cdot z > 0 \quad \forall z \neq 0 \in \mathbb{C}^n,$$

d.h., A ist positiv definit. $\square$