3.5 Diagonalisierung [Lang & Pucker 3.4.1-2]

Orthogonale Basistransformation (Seite 40):

$$\vec{e}_i' := O\vec{e}_i \quad \Rightarrow \quad \vec{e}_i' \cdot \vec{e}_j' = [e_i']^T[e_j'] = e_i^T O^T O e_j = \delta_{ij},$$

d.h. die neue Basis ist auch orthonormiert, falls $O^T O = \mathbb{1}$ gilt.

Matrixelemente in der neuen Basis:

$$M_{ij}' := \vec{e}_i' \cdot M\vec{e}_j' = [e_i']^T M e_j' = e_i^T O^T M O e_j = (O^T M O)_{ij}.$$

Vektorkomponente in der neuen Basis:

$$v_i' := \vec{e}_i' \cdot \vec{v} = [e_i']^T v = e_i^T O^T v = (O^T v)_i .$$

Kurz zusammengefasst: Spaltenvektoren (v), Zeilenvektoren (v^T) und Matrizen (M) werden als

$$v \to v' = O^T v$$
$$v^T \to v'^T = v^T O$$
$$M \to M' = O^T M O$$

transformiert. „Skalare" Größen wie $v^T v$ oder $v^T M v$ sind invariant.

— o —

Sei jetzt M eine symmetrische $n \times n$-Matrix. Wir nehmen an, dass alle n Eigenwerte ungleich sind oder, falls es Entartung gibt, dass die Eigenvektoren trotzdem als orthogonal (und normiert) gewählt werden können (vgl. Seite 50).

Wir setzen die Eigenvektoren als Spalten in O: $\quad O := (v^{(1)}\, v^{(2)} \dots v^{(n)})$

Es gilt: $\quad MO = (Mv^{(1)}\, Mv^{(2)} \dots Mv^{(n)}) = (\lambda_1 v^{(1)}\, \lambda_2 v^{(2)} \dots \lambda_n v^{(n)})$,

$$O^T M O = \begin{pmatrix} v^{(1)T} \\ v^{(2)T} \\ v^{(n)T} \end{pmatrix} (\lambda_1 v^{(1)}\ \lambda_2 v^{(2)} \dots \lambda_n v^{(n)})$$

$$= \begin{pmatrix} \lambda_1 \overbrace{v^{(1)T} v^{(1)}}^{1} & \lambda_2 \overbrace{v^{(1)T} v^{(2)}}^{0} & \dots & \lambda_n \overbrace{v^{(1)T} v^{(n)}}^{0} \\ \lambda_1 \underbrace{v^{(2)T} v^{(1)}}_{0} & \lambda_2 \underbrace{v^{(2)T} v^{(2)}}_{1} & \ddots & \\ \vdots & & & \\ & & & \lambda_n v^{(n)T} v^{(n)} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_n \end{pmatrix} .$$

D.h., die so gewählte orthogonale Basistransformation diagonalisiert die Matrix M: M' hat nur diagonale Komponente, die genau die Eigenwerte von M (und M') sind.

Eine klassische Anwendung der Diagonalisierung beschäftigt sich mit quadratischen Formen und wird auch Hauptachsentransformation genannt.

Sei $f(x_1,\dots,x_n) := x^T M x$.

> Bemerkung: M kann immer als symmetrisch gewählt werden.
>
> Beweis: $f = \sum_{i,j=1}^{n} x_i M_{ij} x_j = \sum_{i,j=1}^{n} \left(\frac{1}{2} x_i M_{ij} x_j + \frac{1}{2} x_j M_{ji} x_i \right)$
>
> umnenne Indizes!
>
> $$= \sum_{i,j=1}^{n} x_i \frac{M_{ij}+M_{ji}}{2} x_j$$
>
> $$= x^T \left(\frac{M+M^T}{2} \right) x$$
>
> symmetrisch!

Die komplizierte Gleichung $f(x_1,\dots,x_n) = C$ kann in neuen Koordinaten viel einfacher als

$$\underset{\mathbb{1}=OO^T}{x^T}\,\underset{\mathbb{1}=OO^T}{M}\,x = \underbrace{(O^T x)^T}_{x'}\,\underbrace{O^T M O}_{M'}\,\underbrace{(O^T x)}_{x'} = \sum_{i=1}^{n} \lambda_i (x_i')^2 = C$$

ausgedrückt werden.

> „Rang" der quadratischen Form := Zahl von $\lambda_i \neq 0$.
> Falls $\forall \lambda_i > 0$ geht es um die Gleichung eines „Ellipsoids" in Hauptachsenlage.

Beispiel: Skizziere Lösung von $2x^2 + 4xy - y^2 = 6$.

$\Rightarrow$ Schreibe Gl. als $(x\ y)\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 6$.

Eigenwerte und Eigenvektoren von $\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}$: $\lambda_1 = -2,\ v^{(1)} = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -2 \end{pmatrix}$

$\lambda_2 = 3,\ v^{(2)} = \frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$

Die Eigenvektoren sind orthogonal: $[v^{(1)}]^T v^{(2)} = \frac{1}{5}(2-2) = 0$.

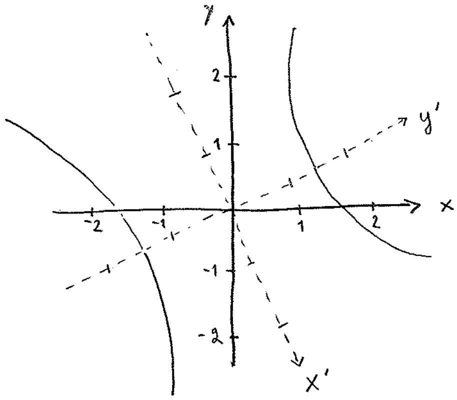
$O = (v^{(1)}\, v^{(2)}) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$; $O^T = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$

$\begin{pmatrix} x' \\ y' \end{pmatrix} = O^T \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} x-2y \\ 2x+y \end{pmatrix}$

$M' = O^T M O = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix}\frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}$.

In neuen Koordinaten gilt also $-2(x')^2 + 3(y')^2 = 6$

$\Leftrightarrow\ y' = \pm\sqrt{2 + \frac{2}{3}(x')^2}$



Verallgemeinerungen (z.B. M nicht symmetrisch)

Eine „Ähnlichkeitstransformation" wird als $M \to M' = P^{-1} M P$ definiert.

Klassische Fragen der linearen Algebra sind, unter welchen Bedingung eine solche P existiert dass M' diagonal wird, und wie diese P konstruiert werden kann?

Falls P existiert, können viele Aussagen gemacht werden:

* die diagonalen Elemente von M' sind genau die Eigenwerte von M.

$$\left[\quad M' e_i' = \lambda_i e_i' \quad \Leftrightarrow \quad P^{-1} M P e_i' = \lambda_i e_i' \quad \Rightarrow \quad M(Pe_i') = \lambda_i (Pe_i') . \ \square \right]$$

(Um genau zu sein sind diese Eigenschaften unabhängig davon ob M' diagonalisiert wird.)

* Spur und Determinante bleiben in der Transformation invariant.

$$\left[\begin{aligned} Sp(P^{-1} M P) &= Sp(P P^{-1} M) = Sp(M) \\ \det(P^{-1} M P) &= \det(P^{-1}) \det(M) \det(P) = \underbrace{\det(P^{-1} P)}_{1} \det(M) \end{aligned} \right]$$

* Summe und „Produkt" von Matrizen transformieren sich wie die Matrizen selbst.

$$\left[\begin{aligned} M' + N' &= P^{-1} M P + P^{-1} N P &= P^{-1} (M+N) P \\ M' N' &= P^{-1} M \underbrace{P P^{-1}}_{\mathbb{1}} N P &= P^{-1} (MN) P \end{aligned} \right]$$

* die ursprüngliche Matrix M kann durch die Spektraldarstellung ausgedrückt werden:

$$\underbrace{\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}}_{M'} = P^{-1} M P \quad \Rightarrow \quad M = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1} .$$

Um P zu konstruieren, könnten wir wieder die Eigenvektoren von M als Spalten setzen (vgl. Seite 51). Die Frage ist dann, ob P^{-1} existiert; dies ist der Fall nur falls es n Eigenvektoren gibt, und diese linear unabhängig sind. Nicht immer der Fall! (vgl. B in Aufgabe 13.1)

Es gibt aber einige einfache hinreichende Bedingungen, z.B. dass alle n Eigenwerte unterschiedlich sind.

[Ohne Beweis; eine bestimmte Ähnlichkeit mit Punkt (ii) auf Seite 50 kann aber bemerkt werden.]

Die in der Physik auftauchenden Matrizen können in der Regel diagonalisiert werden, und zwar mit einer orthogonalen ($P^{-1} = P^T$) bzw. unitären ($P^{-1} = P^+$) Matrix.

[M^{-1} braucht nicht zu existieren; nur P^{-1}!]

— o —

Matrixfunktionen

Wenn wir „Produkte" (d.h. Kompositionen) von Matrizen miteinander linear kombinieren, bekommen wir Matrixfunktionen, z.B. ein Polynom

$$P_m(M) = \sum_{k=0}^{m} a_k M^k \quad , \quad \text{mit } a_k \in \mathbb{R} \text{ und } M^0 := \mathbb{1} .$$

Die höchste Potenz könnte im Prinzip beliebig groß sein; der Satz von Cayley und Hamilton (Seite 49) besagt ja, dass nur Potenzen bis M^{n-1} unabhängig voneinander sein können.

Falls M auch noch diagonalisierbar ist, gilt (vgl. Seite 53)

$$\sum_{k=0}^{m} a_k M^k = P \sum_{k=0}^{m} a_k \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}^k P^{-1} = P \begin{pmatrix} \sum_{k=0}^{m} a_k \lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{m} a_k \lambda_n^k \end{pmatrix} P^{-1} .$$

Falls diese Summen für alle Eigenwerte, also insbesondere für $\max(\{|\lambda_i|\})$, konvergent sind, können wir hier sogar $m \to \infty$ schicken.

Besonders wichtig ist die matrixwertige Exponentialfunktion:

$$\exp(M) := \sum_{k=0}^{\infty} \frac{1}{k!} M^k = \mathbb{1} + M + \frac{1}{2} M^2 + \dots$$

Diese ist immer konvergent, weil $\exp(\lambda)$ auch für alle $\lambda \in \mathbb{R}$ (oder sogar $\lambda \in \mathbb{C}$) konvergent ist.

Einige wichtige Eigenschaften:

* $\det(\exp(M)) = \exp(Sp(M))$

  $\left[\det(\exp(M)) = \det\left(P \begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix} P^{-1} \right) = e^{\lambda_1} e^{\lambda_2} \dots e^{\lambda_n} = e^{\lambda_1 + \lambda_2 + \dots + \lambda_n} = e^{Sp(M)} \right]$

* $\ln(\det(Q)) = Sp(\ln(Q))$

  [$Q := \exp(M)$; definiere matrixwertige ln als Umkehrfunktion von exp, d.h. $M =: \ln(Q)$.]

* $\exp(xM)\exp(xN) = \exp\left\{ x(M+N) + \frac{x^2}{2}[M,N] + \mathcal{O}(x^3) \right\}$,

  wobei $[M,N] := MN - NM$ (vgl. Aufgabe 10.4)

  „Campbell-Baker-Hausdorff-Formel"

  [Ohne Beweis]