Ein bisschen Lineare Algebra und Taylorreihenentwicklung

Alexander Röhnsch

28. Februar 2006

WS 2005/2006
im Seminar
Neuronale Netze und Autonome Roboter
bei Manfred Hild
Lehr- und Forschungsgebiet Künstliche Intelligenz
Institut für Informatik
Mathematisch-Naturwissenschaftliche Fakultät II
Humboldt-Universität zu Berlin

1 Lineare Abbildungen

Ausgehend von Vektorräumen über den reellen und den komplexen Zahlen beschäftigen uns zunächst Abbildungen zwischen Vektorräumen. Dabei soll der Vektor eines Vektorraumes als Vektor eines anderen Vektorraumes ausgedrückt werden, er soll „transformiert" werden.

Formal heißt eine Abbildung $f : V \rightarrow W$ zwischen den Vektorräumen V und W über einem Körper K *Lineare Abbildung*, wenn gilt:

$$\begin{aligned} \forall v_1,\, v_2 \in V &\quad:\quad f(v_1 + v_2) = f(v_1) + f(v_2) \\ \forall v \in V,\, k \in K &\quad:\quad f(k \cdot v) = k \cdot f(v) \end{aligned}$$

Diese so genannten Linearitätsbedingungen lassen erkennen, dass es sich bei einer Linearen Abbildung um einen Homomorphismus von Vektorräumen handelt.

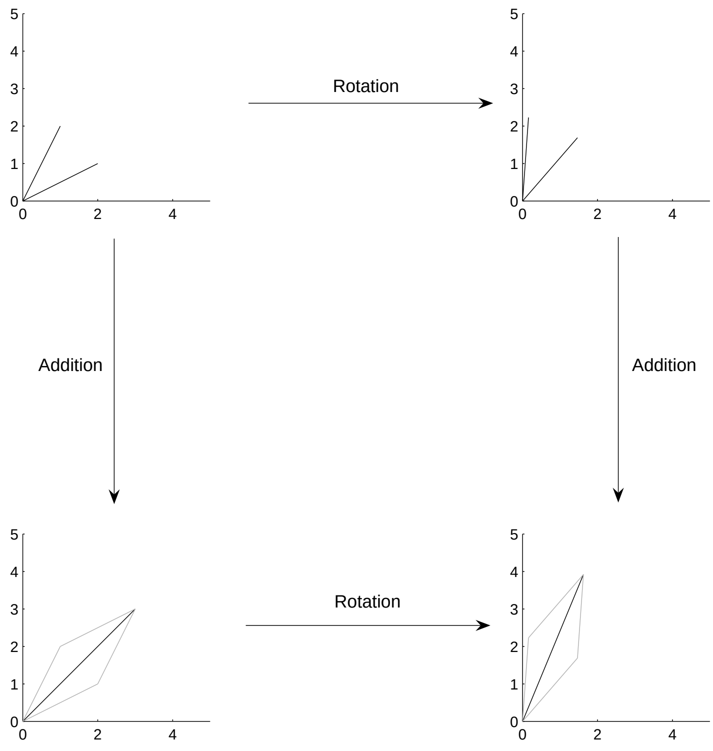


Abbildung 1: Homomorphismus mit Drehen und Addieren

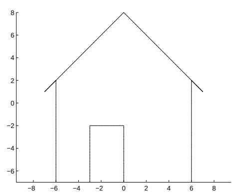
Matrixdarstellung Lineare Abbildungen lassen sich nun durch Matrizen repräsentieren. Jede solche darstellende Matrix beschreibt genau eine Lineare Abbildung bezüglich einer bestimmten Wahl von Basisfamilien für die Vektorräume der Abbildung.

Der Nullvektor wird dabei immer wieder auf den Nullvektor abgebildet.

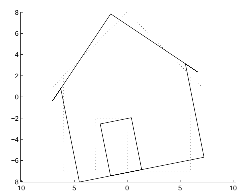
Einzelne Vektoren werden nun vom einen Vektorraum in den anderen transformiert, indem sie mit der Matrix multipliziert werden:

$$M \cdot v = w,\ v \in V,\ w \in W.$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



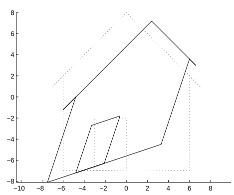
(a) Identische Abbildung



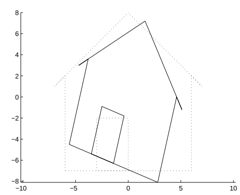
$$\begin{pmatrix} 0{,}98 & -0{,}2 \\ 0{,}2 & 0{,}98 \end{pmatrix}$$

(b) Rotation

$$\begin{pmatrix} 0{,}9 & 0{,}3 \\ 0{,}3 & 0{,}9 \end{pmatrix}$$



(c) Zerrung



$$\begin{pmatrix} 0{,}7 & 0{,}2 \\ -0{,}3 & 0{,}9 \end{pmatrix}$$

(d) Rotation mit Zerrung

Abbildung 2: Beispiele für Lineare Abbildungen auf der Punktmenge (-6;-7), (-6;2), (-7;1), (0;8), (7;1), (6;2), (6;-7), (-3;-7), (-3;-2), (0;-2), (0;-7), (-6;-7).

Ähnliche Matrizen Da die Vektoren von Vektorräumen im Allgemeinen durch verschiedene Basen ausgedrückt werden können, ergeben sich auch für die Formulierung der gleichen Linearen Abbildung mehrere Möglichkeiten, nämlich abhängig von der Wahl der die Vektorräume beschreibenden Basen. Matrizen die auf diese Weise die gleiche Lineare Abbildung beschreiben, heißen *ähnliche Matrizen*.

Zwei Matrizen M, L sind sich genau dann ähnlich, wenn es eine invertierbare Matrix B gibt, so dass gilt:

$$M = B^{-1} \cdot L \cdot B.$$

Da ähnliche Matrizen die gleichen Linearen Abbildungen beschreiben, gleichen sich auch manche ihrer Eigenschaften wie Eigenwerte, Determinante, Rang, charakteristisches Polynom.

Es stellt sich wegen der Vielzahl äquivalenter Matrizen die Frage, ob es in einer Familie ähnlicher Matrizen eine Matrix gibt, die die Lineare Abbildung besonders aussagekräftig beschreibt. Dies ist gerade die Diagonalmatrix, welche etwas später beschrieben wird.

2 Eigenräume

Im Folgenden werden nur quadratische Matrizen betrachtet (und damit Lineare Abbildungen, die von einem Vektorraum in denselben abbilden), da sonst das Eigenwertproblem nicht einfach gelöst werden kann.

Bei der Transformation von Vektoren kann man feststellen, dass im Gegensatz zu den meisten Vektoren manche ihre Richtung beibehalten und nur um einen konstanten Wert gestreckt oder gestaucht und vielleicht gespiegelt werden. Diese Vektoren heißen *Eigenvektoren*, die Werte, um die sie skaliert werden *Eigenwerte*.

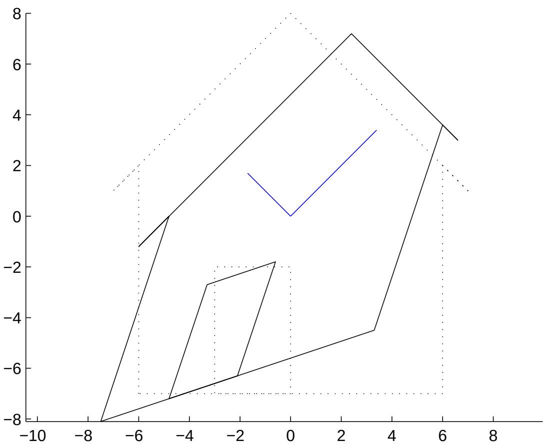


Abbildung 3: Streckung und Stauchung. Die eingezeichneten Eigenvektoren zeigen die Richtung für die Streckung bzw. die Stauchung an. Entlang dieser beiden Vektoren liegen alle Eigenvektoren der Linearen Abbildung.

Eigenvektoren v und Eigenwerte λ einer Matrix M genügen also:

$$M \cdot v = \lambda \cdot v. \tag{1}$$

Die Menge der (Eigen-)Vektoren, die zusammen mit einem bestimmten Eigenwert Gleichung (1) erfüllen, bildet einen Unterraum des Vektorraums und wird *Eigenraum* genannt. Die Summe der Dimensionen aller Eigenräume, bzw. die Anzahl linear unabhängiger Eigenvektoren kann kleiner oder gleich der Dimension der Linearen Abbildung sein.

Berechnung der Eigenwerte Gleichung (1) wird umgestellt zu:

$$(M - \lambda E) \cdot v = 0, \tag{2}$$

wobei E die Einheitsmatrix bezeichnet. Gleichung (2) ist nun genau dann erfüllt, wenn entweder v der Nullvektor ist, oder wenn gilt:

$$det(M - \lambda E) = 0.$$

Dabei heißt $det(M - \lambda E)$ *charakteristisches Polynom* von M. Die Nullstellen $\lambda_1 \ldots \lambda_n$ des charakteristischen Polynoms sind genau die Eigenwerte von M.

Dass es sich dabei auch wirklich um ein Polynom handelt, sieht man gut an den Verfahren zur Berechnung von Determinaten. Durch die Besetzung der Variablen λ auf der Hauptdiagonalen der Matrix $(M - \lambda E)$ stellt sich die Determinante als Summe von λ-Potenzen mit verschiedenen Koeffizienten dar, wobei die höchste auftretende Potenz das Produkt der Hauptdiagonalen ist. Der Grad des Polynoms muss also der Größe n der Matrix entsprechen.

Polynome n-ten Grades haben im Bereich der reellen Zahlen höchstens n Nullstellen, im Bereich der komplexen Zahlen genau n Nullstellen, wobei manche Nullstellen identisch sein können. Bei solchen Stellen gibt die so genannte Vielfachheit an, wieviele Lösungen auf dieselbe Nullstelle fallen. Bei einfachen Nullstellen beträgt ihre Vielfachheit also 1.

Die Vielfachheiten aller unterschiedlicher Nullstellen addieren sich zur Dimension des Polynoms.

Berechnung von Eigenvektoren Über die Gleichung

$$(M - \lambda E) \cdot v = 0$$

lässt sich nun für jeden vorher ermittelten Eigenwert λ eine Menge an Eigenvektoren – ein *Eigenraum* – bestimmen. Die Dimension des Eigenraumes ist dabei mindestens gleich 1, höchstens jedoch gleich der Vielfachheit des entsprechenden Eigenwertes. Aus 1-fachen Eigenwerten ergeben sich also eindimensionale Eigenräume.

Beispielrechnung

Zunächst werden die Eigenwerte der Linearen Abbildung $M = \begin{pmatrix} 0{,}9 & 0{,}3 \\ 0{,}3 & 0{,}9 \end{pmatrix}$ bestimmt:

$$\begin{aligned}
& 0 = \det(M - \lambda E) \\
\Leftrightarrow \quad & 0 = \det \begin{pmatrix} 0{,}9 - \lambda & 0{,}3 \\ 0{,}3 & 0{,}9 - \lambda \end{pmatrix} \\
\Leftrightarrow \quad & 0 = (0{,}9 - \lambda)^2 - 0{,}3^2 \\
\Leftrightarrow \quad & \lambda_1 = 1{,}2;\ \lambda_2 = 0{,}6.
\end{aligned}$$

Als Nächstes folgt die Berechnung der Eigenräume mit Hilfe der Eigenwerte:

$$\begin{aligned}
\lambda_1 \ &: \ 0 = (M - \lambda_1 E) \cdot v = \begin{pmatrix} -0{,}3 & 0{,}3 \\ 0{,}3 & -0{,}3 \end{pmatrix} \cdot v \Rightarrow L_1 = \left\{ v \middle| v = \begin{pmatrix} a \\ a \end{pmatrix},\, a \in \mathbb{R} \right\} \\
\lambda_2 \ &: \ 0 = (M - \lambda_2 E) \cdot v = \begin{pmatrix} 0{,}3 & 0{,}3 \\ 0{,}3 & 0{,}3 \end{pmatrix} \cdot v \Rightarrow L_2 = \left\{ v \middle| v = \begin{pmatrix} a \\ -a \end{pmatrix},\, a \in \mathbb{R} \right\}.
\end{aligned}$$

Aus den erhaltenen Lösungsräumen wählen wir nun je einen Vektor aus, und erhalten so zu den Eigenwerten zwei Eigenvektoren:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix};\ v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Beispielrechnung bei vielfachem Eigenwert

Es werden die Eigenwerte von $M = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$ bestimmt:

$$\begin{aligned} & 0 = \det(M - \lambda E) \\ \Leftrightarrow \quad & 0 = \det \begin{pmatrix} 3-\lambda & 2 & 4 \\ 2 & -\lambda & 2 \\ 4 & 2 & 3-\lambda \end{pmatrix} \\ \Leftrightarrow \quad & 0 = (1+\lambda)^2(8-\lambda) \\ \Leftrightarrow \quad & \lambda_{1/2} = -1;\ \lambda_3 = 8. \end{aligned}$$

Als Nächstes folgt die Berechnung der Eigenräume mit Hilfe der Eigenwerte:

$$\begin{aligned} \lambda_{1/2} : 0 \quad &= \quad (M - \lambda_{1/2}E) \cdot v = \begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix} \cdot v \\ &\Rightarrow \quad L_{1/2} = \left\{ v \middle| v = \begin{pmatrix} a \\ b \\ -a - \frac{1}{2}b \end{pmatrix},\ a,\, b \in \mathbb{R} \right\} \\ \lambda_3 : 0 \quad &= \quad (M - \lambda_3 E) \cdot v = \begin{pmatrix} -5 & 2 & 4 \\ 2 & -8 & 2 \\ 4 & 2 & -5 \end{pmatrix} \cdot v \\ &\Rightarrow \quad L_3 = \left\{ v \middle| v = \begin{pmatrix} 2a \\ a \\ 2a \end{pmatrix},\ a \in \mathbb{R} \right\}. \end{aligned}$$

Aus dem ersten erhaltenen Lösungsraum wählen wir zwei Vektoren aus, aus dem zweiten Lösungsraum einen, und erhalten so zu den Eigenwerten drei Eigenvektoren:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix};\ v_2 = \begin{pmatrix} 0 \\ 1 \\ -\frac{1}{2} \end{pmatrix};\ v_3 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}.$$

3 Diagonalisierung

Zusätzlich zu quadratischen Matrizen setzen wir nun noch voraus, dass die Anzahl linear unabhängiger Eigenvektoren genauso groß sei wie die Dimension der Linearen Abbildung.

Die ermittelten Eigenwerte $\lambda_1 \dots \lambda_n$ einer Matrix M werden nun als Hauptdiagonale einer sonst mit Nullen gefüllten Matrix eingesetzt. Man erhält die *Diagonalmatrix*:

$$D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Die Diagonalmatrix ist der ursprünglichen Matrix ähnlich, denn es lässt sich eine Matrix B finden mit:

$$M = B \cdot D \cdot B^{-1}.$$

Die Spalten dieser Matrix sind n linear unabhängige Eigenvektoren $v_1 \dots v_n$ in der Reihenfolge der zugehörigen Eigenwerte $\lambda_1 \dots \lambda_n$ in der Diagonalmatrix:

$$B = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix} = \begin{pmatrix} v_{11} & \cdots & v_{n1} \\ \vdots & \ddots & \vdots \\ v_{1n} & \cdots & v_{nn} \end{pmatrix}$$

Darüber hinaus ist die Diagonalmatrix ähnlich zu allen Matrizen, die die gleiche Lineare Abbildung beschreiben. Errechnet man aus ihr erneut Eigenvektoren, erhält man eine Orthonormalbasis des Vektorraums.

Die Diagonalmatrix beschreibt somit die Lineare Abbildung besonders einfach und es lässt sich besonders leicht mit ihr rechnen, da nur auf der Hauptdiagonalen Werte ungleich 0 liegen.

4 Jordan-Normalform

Die Diagonalmatrix lässt sich nicht bestimmen, wenn nicht genügend linear unabhängige Eigenvektoren existieren. In solchen Fällen wäre es praktisch, eine ähnlich einfache Matrix zur Verfügung zu haben, die ebenfalls ähnlich zu allen Matrizen ist, die die gleiche Lineare Abbildung beschreiben. Dies ist die Jordan-Matrix, auch Jordan-Normalform genannt.

Zu ihrer Bildung wird die ungenügend große Menge der Eigenvektoren um verallgemeinerte Eigenvektoren, auch *Hauptvektoren* genannt, ergänzt.

$v^{(i)}$ ist Hauptvektor i-ter Stufe zum Eigenwert λ, wenn gilt:

$$(M - \lambda E)^i \cdot v^{(i)} = 0 \wedge (M - \lambda E)^{i-1} \cdot v^{(i)} \neq 0.$$

Die bisherigen Eigenvektoren seien Hauptvektoren erster Stufe. Für vielfache Eigenwerte, für die sich nicht genügend Eigenvektoren ergeben, werden weitere Hauptvektoren aus vorhergehenden mit dem Gleichungssystem

$$(M - \lambda E) \cdot v^{(i)} = v^{(i-1)}$$

schrittweise bei aufsteigendem Index i ermittelt. So werden nacheinander Hauptvektoren ermittelt, bis ausreichend viele Vektoren vorhanden sind.

Die Dimension des aus den gefundenen Eigen- und Hauptvektoren bestehenden *algebraischen Eigenraumes* ist jetzt gleich der Dimension der Linearen Abbildung.

Zur Erzeugung der Matrix B werden die Eigen- und Hauptvektoren verwendet. Die durch:

$$J = B^{-1} \cdot M \cdot B$$

errechnete Matrix J heißt *Jordan-Matrix* und ist wie die Diagonalmatrix ähnlich zu allen Matrizen derselben Linearen Abbildung.

Sie setzt sich aus Blöcken zusammen, wobei die Anzahl der Blöcke der Anzahl linear unabhängiger Eigenvektoren (ohne Hauptvektoren) entspricht. Für eine Matrix mit ϵ linear unabhängigen Eigenvektoren gilt also:

$$J = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_\epsilon \end{pmatrix}.$$

Jeder solcher Jordan-Block J_i ist quadratisch und entspricht einem Eigenwert. Die Vielfachheit eines Eigenwertes bestimmt die Seitenlänge des korrespondierenden Jordan-Blockes. Entlang der Hauptdiagonalen weist er seinen Eigenwert auf, entlang der Diagonalen direkt darüber Einsen, und besteht sonst ausschließlich aus Nullen:

$$J_i = \begin{pmatrix} \lambda_i & 1 & \cdots & 0 \\ 0 & \lambda_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

Ist die Anzahl linear unabhängiger Eigenvektoren der Matrix n so groß wie der Wert ihrer Dimension, so besteht die Jordan-Matrix aus n Blöcken der Größe 1, und ist damit identisch mit der Diagonalmatrix.

Beispielrechnung

Für die Matrix $M = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ errechnen sich die Eigenwerte $\lambda_1 = -2$ mit der Vielfachheit 1 und $\lambda_2 = 1$ mit der Vielfachheit 2. Die Eigenräume dazu sind $\left\{ \begin{pmatrix} 0 \\ a \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\}$ und $\left\{ \begin{pmatrix} a \\ \frac{2}{3} \\ 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\}$. Für einen einfachen Rechenweg wählen wir aus dem ersten den Vektor $v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und aus dem zweiten den Vektor $v_2 = \begin{pmatrix} 1 \\ \frac{2}{3} \\ 0 \end{pmatrix}$ als Eigenvektoren aus.

Da der Eigenwert λ_2 eine Vielfachheit von 2 hat, sich dazu aber nur ein einziger Eigenvektor ermitteln lässt (mehrere wären linear abhängig), benötigen wir dazu noch einen Hauptvektor. Wir lösen für λ_2 und den Eigenvektor v_2 die Gleichung $(M - \lambda_2 E)v' = v_2$ und erhalten mit $(M - \lambda_2 E) = \begin{pmatrix} 0 & 0 & 1 \\ 2 & -3 & -1 \\ 0 & 0 & 0 \end{pmatrix}$ durch Einsetzen den Lösungsraum $\left\{ \begin{pmatrix} \frac{5}{6} + \frac{3}{2}a \\ a \\ 1 \end{pmatrix} \middle| a \in \mathbb{R} \right\}$. Daraus wählen wir wiederum den Vektor $v_3 = \begin{pmatrix} \frac{7}{3} \\ 1 \\ 1 \end{pmatrix}$ als Hauptvektor aus.

Die ermittelten Eigen- und Hauptvektoren ergeben spaltenweise die Matrix $B = \begin{pmatrix} 0 & 1 & \frac{7}{3} \\ 1 & \frac{2}{3} & 1 \\ 0 & 0 & 1 \end{pmatrix}$. Mit $J = B^{-1} \cdot M \cdot B$ ergibt sich somit die Jordanmatrix $J = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, die aus den zwei Blöcken $J_1 = (-2)$ und $J_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ besteht, jeweils den Eigenwerten $\lambda_1 = -2$ und $\lambda_2 = 1$ entsprechend.

5 Taylorreihen

Benötigt man die Näherung einer beliebigen Funktion in der Umgebung einer bestimmten Stelle, so kann man sich der *Taylorentwicklung* bedienen. Als Näherungsfunktion wird dabei ein Polynom benutzt, das so genannte *Taylorpolynom*.

Die Funktionswerte einer Funktion f sind äquivalent berechenbar durch:

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \ldots \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, \end{aligned}$$

wobei x_0 eine beliebige aber bestimmte Stelle bezeichnet, die *Entwicklungsstelle*.

Um ein Taylorpolynom für eine Näherung n-ten Grades zu erhalten, werden nur die ersten n Summanden der Formel genommen:

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Als Voraussetzung muss die zu nähernde Funktion an der Entwicklungsstelle x_0 n-mal stetig differenzierbar sein.

Zwar steigt mit höherem Näherungsgrad die Näherungsgüte in der Umgebung der Entwicklungsstelle, oftmals werden aber nur ganz wenige der ersten Entwicklungsglieder benötigt, häufig nur die ersten zwei oder drei, womit sich lineare bzw. quadratische Polynome ergeben. Der Berechnungsaufwand kann so stark reduziert werden.

Spezielle Taylorpolynome Bei den Taylorreihen einiger Funktionen treten durch die Eigenschaften der Funktionen an gewissen Entwicklungsstellen weitere Vereinfachungen auf. So berechnen sich die Taylorpolynome für die Exponentialfunktion und die Sinusfunktion an der Stelle $x_0 = 0$ wie folgt:

$$\begin{aligned} f(x) = e^x \quad &\Rightarrow \quad T_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \\ f(x) = \sin x \quad &\Rightarrow \quad T_n(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!}. \end{aligned}$$

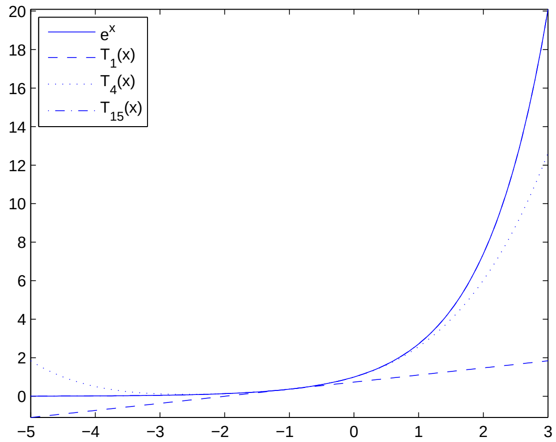


Abbildung 4: Annäherung der Exponentialfunktion $f(x) = e^x$ durch ein lineares Taylorpolynom, ein Taylorpolynom vierten und eines fünfzehnten Grades. Das Polynom fünfzehnten Grades nähert in dieser Umgebung bereits so genau, dass sein Graph nicht mehr von der Ursprungsfunktion unterschieden werden kann.

6 Literatur

BENKER, HANS: *Mathematik mit MATLAB* (Berlin [u.a.]: Springer, 2000)

MATHEWS, JOHN H.: *Numerical methods using Matlab* (Upper saddle River, NJ: Pearson Prentice Hall, 2004)

PAREIGIS, BODO: *Lineare Algebra für Informatiker* (Berlin [u.a.]: Springer, 2000)

STRANG, GILBERT: *Lineare Algebra* (Berlin [u.a.]: Springer, 2003)

JOOS, G. UND KALUZA, TH.: *Höhere Mathematik für den Praktiker* (Leipzig: Johann Ambrosius Barth, 1952)

BRONSTEIN, I.N. UND SEMENDJAJEW, K.A.: *Taschenbuch der Mathematik* (Leipzig: Teubner Verlagsgesellschaft, 1960)

University Level Mathematics (University of Surrey, England) (30. November 2005): http://www.maths.surrey.ac.uk/interactivemaths/emmaspages/