

Methoden für Interpolation und Approximation von Funktionen auf \mathbb{R}

Diplomarbeit zur Erlangung des akademischen Grades eines
Magister der naturwissenschaftlichen Fakultät der Leopold-
Franzens-Universität Innsbruck

vorgelegt von Robert Morandell

betreut von em. Univ. Prof. Dr. Gilbert Helmberg

Institut für Mathematik der Universität Innsbruck

Juni 2000

Danken möchte ich allen, die zum Gelingen dieser Arbeit beigetragen haben. Allen voran meinen Eltern für die Unterstützung und die Geduld und meinem Betreuer Dr. Gilbert Helmberg für die Themenstellung und Beratung. Weiters bin ich noch Gregor Thalhammer und Frau Bombassaro Dank schuldig, die mich in \LaTeX bzw. Matlab eingeführt haben; schließlich Irene Stelzel für die vielen Kleinigkeiten, die sich so leicht anhäufen.

Inhaltsverzeichnis

Einleitung	7
1 Polynominterpolation	9
1.1 Lagrange-Interpolation	9
1.2 Newton-Interpolation	11
1.3 Beispiele und Bilder	12
1.4 Ausschwingen	13
1.5 Rationale Interpolation	17
1.6 Beispiel und Bild	19
2 Tschebyscheff-Interpolation	21
2.1 Tschebyscheff-Polynome	22
2.2 Interpolationsverfahren	23
2.3 Beispiele und Bilder	26
3 Spline-Interpolation	29
3.1 Lineare Splines	29
3.1.1 Spezielle Eigenschaften	29
3.1.2 Beispiele und Bilder	30
3.2 Kubische Splines	33
3.2.1 Verschiedene Splinearten	37
3.2.2 Minimaleigenschaft der kubischen Splineinterpolierenden	45
3.2.3 Beispiele und Bilder	47

4	Bézier-Kurven	51
4.1	Bernstein-Polynome	51
4.2	Bézier-Polynome	54
4.3	Bézier-Splines	60
4.3.1	Arten von Bézier-Splines	63
4.3.2	Beispiel und Bild	64
5	Methode der kleinsten Quadrate	67
5.1	Beschreibung der Methode	68
5.2	Berechnung in Matrizenschreibweise	69
5.3	Gewichtete Methode der kleinsten Quadrate	71
5.4	Interpolative Approximation mittels der Methode der kleinsten Quadrate	72
5.5	Allgemeinere Ansatzfunktionen	73
5.6	Beispiele und Bilder	75
6	Fourier-Analyse	77
6.1	Ergebnisse aus der Fourier-Analysis	77
6.2	Diskrete Fourier-Analyse	78
6.2.1	Interpolation und Approximation durch Fourier-Polynome	81
6.2.2	Beispiele und Bilder	85
7	Approximation mittels Taylorformel	90
7.1	Beschreibung der Methode	90
7.2	Beispiele und Bilder	91
8	Zusammenfassung	97
	Literaturverzeichnis	99
	Lebenslauf	100

Abbildungsverzeichnis

1.1	Relativ gutes Ergebnis eines Interpolationspolynoms.	12
1.2	Oszillieren eines Polynoms 4. Grades.	13
1.3	Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 2. Grades.	14
1.4	Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 4. Grades.	15
1.5	Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 8. Grades; die Funktion beginnt an den Enden bereits „unruhig“ zu werden.	15
1.6	Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 14. Grades; Oszillation ist bereits sehr stark ausgeprägt.	16
1.7	Die Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 9. Grades ist „schief.“	16
1.8	Zusammen mit der Polstelle bei $x = \pi$ erinnert diese Kurve stark an eine rationale Funktion.	19
1.9	Approximation einer Funktion anhand der Gestalt durch ein entsprechendes rationales Polynom.	20
2.1	$\frac{1}{1+x^2}$ approximiert durch ein Tschebyscheff-Polynom aus 15 Stützstellen. .	26
2.2	$\frac{1}{1+x^2}$ approximiert durch ein Tschebyscheff-Polynom aus 16 Stützstellen. .	27
3.1	$\frac{1}{1+x^2}$ linear approximiert durch 4 Punkte.	31
3.2	$\frac{1}{1+x^2}$ linear approximiert durch das Maximum und wenige symmetrische Punkte.	31
3.3	$\frac{1}{1+x^2}$ linear approximiert durch 27 Punkte.	32
3.4	Lineare Spline-Interpolation für $\frac{1}{2}\log(x-1)$	33
3.5	Noch nicht überzeugende Spline-Interpolation durch 5 Stützstellen für die Funktion $\frac{1}{1+x^2}$	47
3.6	Auch für größere Stützstellenanzahl ist die Spline-Interpolation geeignet. .	48

3.7	Periodische Spline-Interpolation von $\sin(2 \cos(x))$ mit gutem Ergebnis. . .	49
4.1	Die Bernstein-Polynome $B_i^3(\lambda)$ für $\lambda \in [0, 1]$, $i = 0, \dots, 3$	52
4.2	Kubisches Bézier-Polynom mit den Bézier-Punkten b_i und der zugehörigen linearen Hülle.	55
4.3	Kubische Bézier-Polynome mit unterschiedlichem b_1	57
4.4	Bézier-Polynomapproximation für den Viertelkreis.	60
4.5	Geometrische Deutung der Anschlussbedingungen.	62
4.6	Ein 2 mal stetig differenzierbarer Bézier-Spline.	64
5.1	Interpolation der beiden Nullstellen und Approximation der restlichen Stützstellen mit Hilfe der Methode der kleinsten Quadrate.	72
5.2	Approximation von $\frac{1}{1+x^2}$ bei geringen Messfehlern.	75
5.3	Approximation von $\frac{1}{1+x^2}$ bei größeren Messfehlern.	76
6.1	$\sin(2 \cos(x))$ approximiert durch ein Fourier-Polynom 4. Grades.	85
6.2	$\frac{1}{1+x^2}$ approximiert durch ein Fourier-Polynom 4. Grades.	86
6.3	$\frac{1}{1+x^2}$ approximiert durch ein Fourier-Polynom 8. Grades.	87
6.4	$ x $ approximiert durch ein Fourier-Polynom 4. Grades.	88
6.5	$ x $ approximiert durch ein Fourier-Polynom 8. Grades.	89
7.1	Approximation von $f(x) = x^2 \cdot x $ mittels eines Taylorpolynoms höchstmöglichen Grades.	92
7.2	Der Bereich der „guten“ Approximation vergrößert sich für Taylorpolynome mit dem Erhöhen des Grades nur langsam; zusätzlich wächst der Fehler ausserhalb dieses Bereichs.	93
7.3	Taylorpolynome hohen wie niedrigen Grades schwingen nicht im Bereich um den Entwicklungspunkt und bieten auch schon ab einem niedrigen Grad eine passable Approximation in einem gewissen Bereich.	94
7.4	Taylorapproximation einer periodischen Funktion im Symmetriepunkt. .	95
7.5	Taylorapproximation einer periodischen Funktion an einer beliebigen, nicht speziellen Stelle, z.B.: $m = \left(\frac{\pi}{3}, f\left(\frac{\pi}{3}\right)\right)$	96
7.6	Taylorapproximation von $\sin(x)$	96

Einleitung

Die Approximation ist ein wichtiger Teilbereich der numerischen Mathematik, da in der Praxis vielfach nicht exakt, sondern nur annäherungsweise gerechnet werden kann. Wegen des großen Umfangs ist es wichtig, sich auf einen Teilbereich zu beschränken. Selbst „Interpolation und Approximation von Funktionen auf \mathbb{R} “ ist noch zu weitläufig und daher möchte ich mich auf die Aufgabenstellung konzentrieren, eine durch eine Tabelle von Werten gegebene, eindimensionale Funktion anzunähern. „Annähern“ ist dabei das richtige Wort, um die Allgemeinheit der Aufgabenstellung geeignet zu beschreiben. Spezielle Anwendungsgebiete wie Integration oder Ableitung von in tabellarischer Form gegebenen Funktionen werden nicht behandelt, da dies den Rahmen dieser Arbeit sprengen würde. Für diese Problemstellungen muss auf weitere Literatur verwiesen werden. Was die Quellen zu dieser Arbeit betrifft, so werden zu einem großen Teil die Ergebnisse aus *Numerische Mathematik* von Hans Rudolf Schwarz [9] und *Numerische Mathematik für Ingenieure und Naturwissenschaftler* von Friedrich Weller [12] verwendet. Die restlichen Beiträge sind im Literaturverzeichnis zu finden.

Interpolieren heisst das Erstellen einer Funktion, deren Graph durch gegebene Punkte (x_i, y_i) laufen soll; eine Approximationsfunktion andererseits ist eine Funktion, die nicht genau durch die Punkte (x_i, y_i) geht, sondern den offensichtlich auftretenden Fehler nach gewissen Bedingungen minimiert.

Der erste Teil der Arbeit beschäftigt sich mit der Interpolation. Das sind einmal die Methoden von Newton und Lagrange, die beide die Koeffizienten der durch $n + 1$ Punkte gegebenen Polynomfunktion n -ten Grades eindeutig bestimmen. Eine weitere Möglichkeit bieten Tschebyscheff-Polynome, die durch Vorgeben der Stützstellen den Fehler minimal zu halten versuchen. Von einem völlig anderen Standpunkt geht die Spline-Methode aus, die für jedes Teilintervall $[x_i, x_{i+1}]$ ein eigenes Polynom niedrigen Grades berechnet, wobei an den Randpunkten gewisse Bedingungen erfüllt werden müssen. Als letzte Variante werden noch Bézier-Polynome beschrieben, die es auch in einer Bézier-Splinevariante gibt. Bézier-Polynome bieten den Vorteil, dass sie geometrisch anschaulich zu deutende Punkte enthalten, die auch nachträglich noch einfach und anschaulich verändert werden können.

Die diskrete Fourier-Analyse stellt einen Sonderfall dar, da ein und dieselbe Methode sowohl eine Interpolation als auch eine Approximation nach der ebenfalls beschriebenen Methode der kleinsten Quadrate ermöglicht. Die Methode der kleinsten Quadrate ist es

auch, die den zweiten Teil über die Approximation beginnt. Schließlich wird die Taylor-entwicklung als Approximationsmethode beschrieben.

Die verschiedenen Methoden werden aber nicht nur beschrieben, sondern in einem zweiten Schritt anhand der Theorie und vor allem anhand von anschaulichen Beispielen verglichen, um die Vor- und Nachteile der einzelnen Methoden zu ergründen.

Prinzipiell gibt es zwei mögliche Wege zu Daten zu kommen. Einmal durch Messen von Werten, für die eine Funktion zur Bestimmung der Punkte zwischen den Messwerten gefunden werden soll, andererseits besteht auch die Möglichkeit der rein mathematischen Nutzung der Interpolation bzw. Approximation, mit dem Ziel für schwierig und aufwendig zu berechnende Funktionen eine einfachere, möglichst gute Alternative zu bieten. Für beide Anwendungsbereich gilt, dass je mehr über die Funktion bekannt ist, desto besser. Sind aus der Theorie Extrema, Wendepunkte, Nullstellen, Pole und Ähnliches bekannt, so sind nach Möglichkeit diese Punkte zu tabellieren. Auch wenn gewissen Methoden wie Tschebyscheff und Fourier es nicht gestatten, die Stützpunkte frei zu wählen, so erleichtert oder ermöglicht dieses Wissen doch die Analyse der berechneten Funktion und die Entscheidung, ob man mit dem Ergebnis zufrieden ist oder nicht. Letztendlich ist die Entscheidung einerseits von der eventuell bestehenden Theorie und zweitens, und das ist meist das wichtigste, vom „Auge“ abhängig. Ist man nicht zufrieden, so stehen zwei Möglichkeiten zur Wahl: Entweder versucht man durch mehr Information wie zum Beispiel durch mehr Messwerte oder durch Verändern von Parametern das Ergebnis noch zu verbessern, oder man benützt eine andere Methode und hofft auf ein besseres bzw. den Erwartungen eher entsprechendes Ergebnis.

An dieser Stelle möchte ich noch auf folgende Konvention in den Abbildungen hinweisen. Mehrmals wird die interpolierte beziehungsweise die approximierte Funktion der exakten direkt in einem Bild gegenübergestellt. In der ganzen Arbeit ist dabei die Ausgangsfunktion strichliert gezeichnet, die angenäherte prinzipiell durchgehend. Nur wenn mehr als 2 approximierte Funktionen in einer Grafik dargestellt sind, kommt es vor, dass Approximationen nicht durchgehend gezeichnet sind.

1 Polynominterpolation

Ausgegangen werde von folgender Situation: Eine Funktion f sei in tabellarischer Form gegeben durch $n+1$ Paare von Werten (x_i, y_i) . Diese Funktion f soll nun durch eine Polynomfunktion P angenähert werden. Der Grad des Interpolationspolynoms beträgt für $n+1$ Stützpunkte gleich n .

Im Folgenden werden die Verfahren von Lagrange und Newton vorgestellt. Obwohl beide Methoden von verschiedenen Ansätzen ausgehen, ist das Ergebnis trotzdem das gleiche. Das ist auch nicht weiter verwunderlich, da es zu $n+1$ Punkten nur ein Polynom von maximal n -tem Grad gibt, das die Interpolationsbedingungen erfüllt. Die Methode nach Newton hat den Vorteil, dass noch beliebig viele Punkte hinzugenommen werden können, ohne dass alle bereits berechneten Koeffizienten wieder berechnet werden müssen, da nur noch der Koeffizient des höchsten Grades bestimmt werden muss.

Interpolationspolynome haben allerdings den Nachteil, dass sich durch Erhöhen der Stützstellenanzahl die Approximation im Allgemeinen nicht unbedingt verbessert, da Polynomfunktionen höheren Grades meist „unruhig“ werden und oszillieren (vgl. Abschnitt 1.4).

Des Weiteren wird auch die Interpolation mit rationalen Funktionen untersucht, die besonders für Funktionen mit Polstellen gut geeignet sind. Auch diese Methode ist aber nicht unproblematisch, da es zum einen für rationale Polynome keinen eindeutigen sondern mehrere mögliche Ansätze gibt und da zum anderen nicht für jedes Ansatzpolynom eine Lösung des Interpolationsproblems existieren muss.

1.1 Lagrange-Interpolation

Wir folgen [9] und definieren das Polynom $P_n(x)$ durch

$$P_n(x) := \sum_{i=0}^n f(x_i) L_i(x) \tag{1.1}$$

und

$$\begin{aligned} L_i(x) &= \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)} \\ &= \prod_{\substack{k=0 \\ k \neq i}}^n \left(\frac{x-x_k}{x_i-x_k} \right). \end{aligned} \quad (1.2)$$

Aus den Definitionen für P_n und L_i ist ersichtlich, dass der Grad des Polynoms P_n gleich n ist. Weiters gilt

$$L_i(x_j) = \delta_{ij} = \begin{cases} 1, & \text{für } i = j \\ 0, & \text{für } i \neq j \end{cases}, \quad (1.3)$$

da aus (1.2) und $i = j$ folgt, dass

$$L_i(x_i) = \prod_{\substack{k=0 \\ k \neq i}}^n \left(\frac{x_i-x_k}{x_i-x_k} \right) = 1.$$

Andererseits gilt unter der Bedingung $i \neq j$, dass

$$L_i(x_j) = \left[\prod_{\substack{k=0 \\ k \neq i \\ i \neq j}}^n \left(\frac{x_j-x_k}{x_i-x_k} \right) \right] \cdot \underbrace{\left(\frac{x_j-x_j}{x_i-x_j} \right)}_{=0} = 0.$$

Als Nächstes muss noch gezeigt werden, dass das Polynom P_n die Punkte (x_i, y_i) , $i = 0, \dots, n$ interpoliert. Dies ist aber leicht einsichtig, da aus (1.3) folgt, dass

$$\sum_{i=0}^n L_i(x_j) = 1$$

und somit gilt für alle Interpolationspunkte x_j ($j = 0, \dots, n$)

$$\begin{aligned} P_n(x_j) &= f(x_0) \underbrace{L_0(x_j)}_{=0} + \dots + f(x_{j-1}) \underbrace{L_{j-1}(x_j)}_{=0} + \\ &+ f(x_j) \underbrace{L_j(x_j)}_{=1} + f(x_{j+1}) \underbrace{L_{j+1}(x_j)}_{=0} + \dots + f(x_n) \cdot \underbrace{L_n(x_j)}_{=0} = y_j. \end{aligned}$$

1.2 Newton-Interpolation

Die Darstellung des Interpolationspolynoms nach Newton geht unter der Voraussetzung von $n+1$ paarweise verschiedenen Stützpunkten von folgendem Ansatz mit den unbestimmten Koeffizienten c_0, c_1, \dots, c_n aus:

$$P_n(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \dots + c_n(x - x_0)(x - x_1)(x - x_2)(x - x_3) \dots (x - x_{n-1}) \quad (1.4)$$

Die einzelnen Koeffizienten können nun folgendermaßen sukzessive berechnet werden. Sei $n = 0$, dann folgt aus dem obigen Ansatz, dass

$$P_0(x_0) = y_0 = c_0. \quad (1.5)$$

Wenn man nun noch einen weiteren Punkt hinzunimmt ($n = 1$), so soll nach (1.4) gelten:

$$P_1(x) = y_1 = c_0 + c_1(x - x_0). \quad (1.6)$$

Das Polynom P_1 interpoliert (x_0, y_0) , da für beliebige c_1 gilt, dass

$$P_1(x_0) = c_0 + c_1(x_0 - x_0) = c_0 = y_0. \quad (1.7)$$

Den Koeffizienten c_1 gilt es nun so zu bestimmen, dass die Gleichung $P_1(x_1) = y_1$ erfüllt ist:

$$P_1(x_1) = y_1 = c_0 + c_1(x_1 - x_0), \quad c_1 = \frac{y_1 - c_0}{x_1 - x_0}. \quad (1.8)$$

Analog verfährt man mit den nächsten Punkten, die hinzugenommen werden. Hat man auf diese Weise das Interpolationspolynom für n Punkte berechnet und nimmt nun noch einen weiteren Punkt hinzu, so geschieht dies nach demselben Schema. Die Idee ist, dass zum bereits berechneten Polynom ein Polynom addiert wird, dessen Funktionswert an allen vorhergehenden Stützstellen x_i gleich 0 ist, damit die Interpolationseigenschaft für diese Punkte erhalten bleibt. Der neue Koeffizient c_n wird so gewählt, dass auch der nächste Punkt (x_{n+1}, y_{n+1}) mit dem Polynom P_n inzidiert.

$$P_i(x) = c_0 + c_1(x - x_0) + c_2(x - x_0)(x - x_1) + \dots + c_n \underbrace{(x - x_0)(x - x_1)(x - x_2)(x - x_3) \dots (x - x_{n-2})(x - x_{n-1})}_{= 0 \quad \text{für } x=x_k \quad k=0, \dots, n-1} \quad (1.9)$$

Durch Umformen von

$$P_n(x) = P_{n-1}(x) + c_n \prod_{k=0}^{n-1} (x - x_k) = y_n$$

zu

$$c_n = \frac{y_n - P_{n-1}(x_n)}{\prod_{k=0}^{n-1} (x_n - x_k)}$$

folgt die sukzessive Koeffizientenberechnung:

$$\begin{aligned} c_0 &= y_0 \\ c_n &= \frac{y_n - P_{n-1}(x_n)}{\prod_{k=0}^{n-1} (x_n - x_k)} \quad (n \geq 1). \end{aligned} \quad (1.10)$$

1.3 Beispiele und Bilder

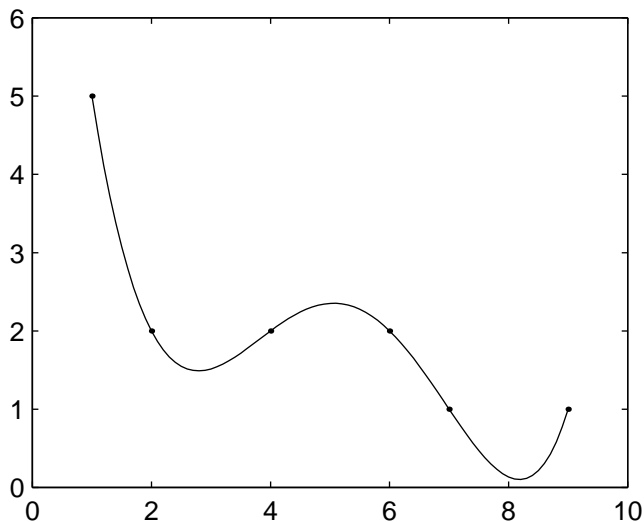


Abbildung 1.1: Relativ gutes Ergebnis eines Interpolationspolynoms.

Beispiel 1. Interpolationspolynome können gute Ergebnisse liefern wie in Abbildung 1.1, wo die 6 Punkte $(1, 5)$, $(2, 2)$, $(4, 2)$, $(6, 2)$, $(7, 1)$ und $(9, 1)$ durch ein Polynom 5. Grades interpoliert werden, nämlich durch das Polynom $P_5(x) = \frac{1}{1260}x^5 + \frac{2}{315}x^4 - \frac{67}{180}x^3 + \frac{1027}{315}x^2 - \frac{1081}{105}x + \frac{62}{5}$, dessen Koeffizienten auf 4 Stellen nach dem Komma gerundet sind. Allerdings fällt auf, dass die Lösung am rechten Rand eine unerwartet große Auslenkung nach unten hat.

Beispiel 2. Dieses Schwingen oder Oszillieren am Rande der Funktion, das ansatzweise in Abbildung 1.1 zu sehen ist, ist besonders typisch für Interpolationspolynome höheren

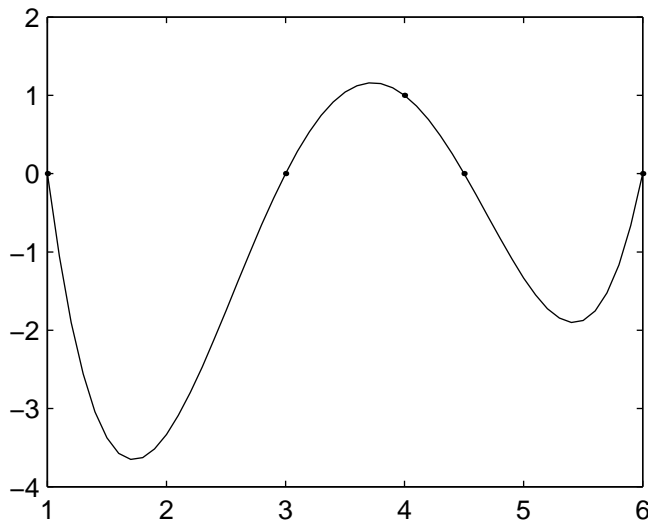


Abbildung 1.2: Oszillieren eines Polynoms 4. Grades.

Grades. Es kann aber auch bei Polynomen niederen Grades vorkommen [8]. Abbildung 1.2 zeigt zum Beispiel ein Interpolationspolynom 4. Grades durch die Punkte $(1, 0)$, $(3, 0)$, $(4, 1)$, $(4.5, 0)$ und $(6, 0)$. Das zugehörige Interpolationspolynom lautet $P_4(x) = \frac{1}{3}x^4 - \frac{29}{6}x^3 + 24x^2 - 46.5x + 27$. Auffällig ist, dass die Punkte mit einer Ausnahme auf der y-Achse liegen und dass das Interpolationpolynom deutlich negative Werte annimmt. Ob überhaupt negative y-Werte zulässig sind, ist von der Aufgabenstellung abhängig. Gegebenenfalls muss die Situation durch mehr Information geklärt werden.

1.4 Ausschwingen

Was das Hinzufügen von Stützpunkten für diese Interpolationsmethode bedeutet, zeigen die nächsten Beispiele und Abbildungen.

Beispiel 3. Die Funktion $f(x) = \frac{1}{1+x^2}$ soll durch 3, 5, 9 und 15 Stützpunkte interpoliert werden. Auffällig ist, wie sehr die Polynomfunktion zu oszillieren beginnt, je mehr Stützpunkte verwendet werden. Die erste Annäherung durch die 3 Stützpunkte (vgl. Abb. 1.3) $(-5, \frac{1}{26})$, $(0, 1)$, und $(5, \frac{1}{26})$ ergibt das Polynom $P_2(x) = -0.0385x^2 + 1$. Dieses kann nur sehr grob die Tendenz der eigentlichen Funktion widerspiegeln. So gibt es zum Beispiel noch keine Wendepunkte.

Nach Hinzunahme von 2 weiteren Punkten $(-3, \frac{1}{10})$ und $(3, \frac{1}{10})$ enthält das Interpolationspolynom $P_4(x) = 0.0038x^4 - 0.1346x^2 + 1$ in Abbildung 1.4 nun Wendepunkte, dafür muss man aber auch 2 Minima in Kauf nehmen, die die Funktion f nicht besitzt. Daher haben auch die Steigungen in den Umgebungen der Ränder jeweils die falschen Vorzeichen.

1 Polynominterpolation

Wir nehmen nun 4 weitere Punkte hinzu, nämlich $(-4, \frac{1}{17})$, $(-2, \frac{1}{5})$, $(2, \frac{1}{5})$ und $(4, \frac{1}{17})$, und sehen (Abb. 1.5), dass im Intervall $[-2, 2]$ die Approximation recht gut ist; außerhalb dieses Intervalls beginnt die Funktion aber, je weiter sie sich den Rändern nähert, mehr und mehr zu oszillieren. Das zugehörige Interpolationspolynom lautet $P_8(x) = 0.00004x^8 - 0.00249x^6 + 0.04633x^4 - 0.34842x^2 + 1$.

Dieser Effekt ist im Polynom 14. Grades in Abb. 1.6 dominierend und beeinträchtigt sogar die bereits gute Interpolation von P_8 im Intervall $[-2, 2]$. Es wurden die Stützstellen $(-4.5, \frac{1}{21.25})$, $(-1, \frac{1}{2})$, $(-0.5, \frac{1}{1.25})$, $(0.5, \frac{1}{1.25})$, $(1, \frac{1}{2})$ und $(4.5, \frac{1}{21.25})$ hinzugefügt, um

$$P_{14} = -0.00000096x^{14} + 0.00007x^{12} - 0.00190x^{10} + 0.026512x^8 - 0.18236x^6 + 0.59559x^4 - 0.93791x^2 + 1 \quad (1.11)$$

zu erhalten.

Die Funktion f ist symmetrisch zur y-Achse. Ist dies bekannt oder vermutbar, so muss möglichst ein Polynom geraden Grades verwendet werden, sonst ist das Ergebnis nicht sehr ansprechend, wie Abbildung 1.7 zeigt. Als Interpolationspunkte wurden $(-5, \frac{1}{26})$, $(-4, \frac{1}{17})$, $(-3, \frac{1}{10})$, $(-2, \frac{1}{5})$, $(0, 1)$, $(1, \frac{1}{2})$, $(2, \frac{1}{5})$, $(3, \frac{1}{10})$, $(4, \frac{1}{17})$ und $(5, \frac{1}{26})$ verwendet. Daraus erhält man das Polynom

$$P_9(x) = -0.00002x^9 + 0.00004x^8 + 0.00122x^7 - 0.00249x^6 - 0.02192x^5 + 0.04633x^4 + 0.15104x^3 - 0.34842x^2 - 0.32579x + 1. \quad (1.12)$$

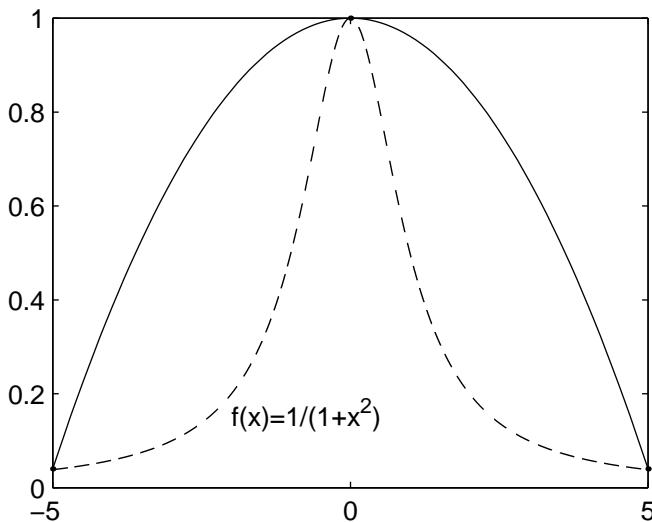


Abbildung 1.3: Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 2. Grades.

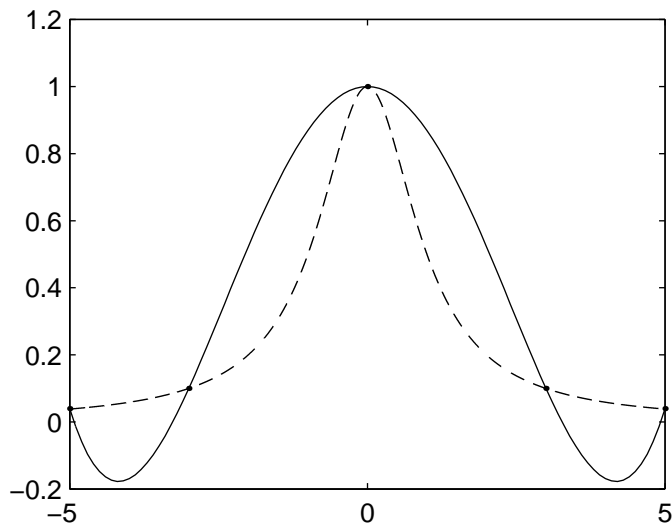


Abbildung 1.4: Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 4. Grades.

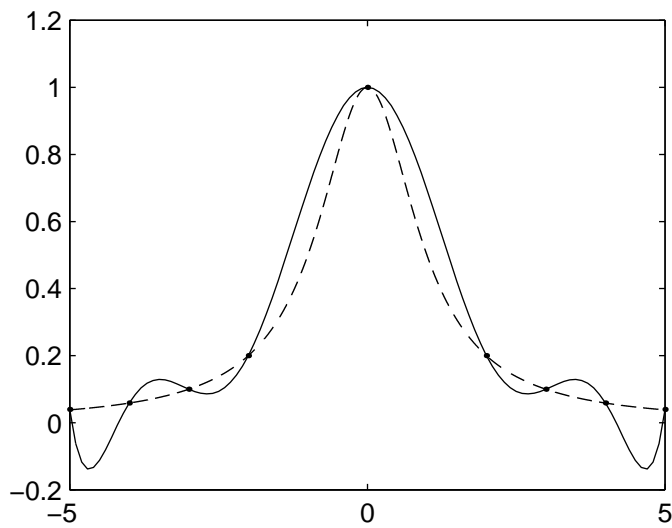


Abbildung 1.5: Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 8. Grades; die Funktion beginnt an den Enden bereits „unruhig“ zu werden.

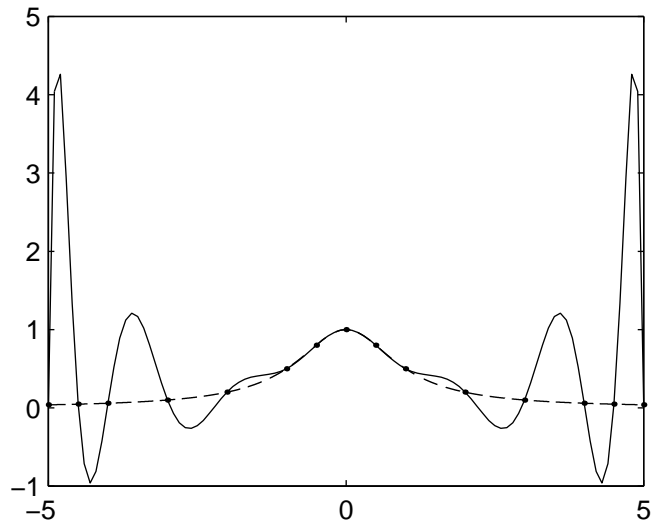


Abbildung 1.6: Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 14. Grades; Oszillation ist bereits sehr stark ausgeprägt.

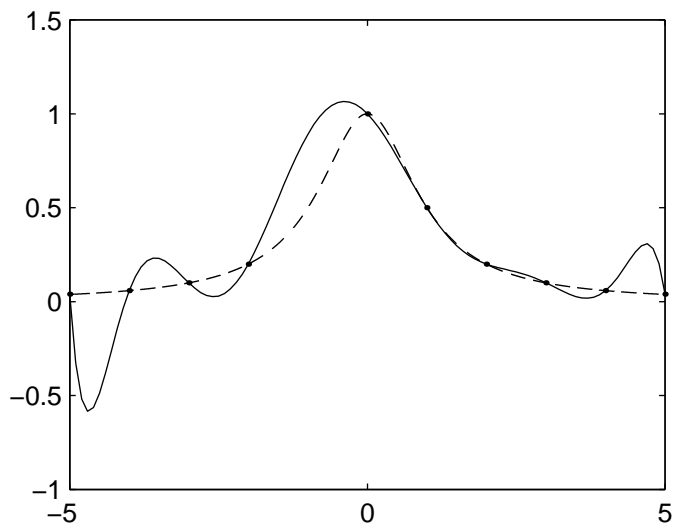


Abbildung 1.7: Die Approximation von $f(x) = \frac{1}{1+x^2}$ durch ein Polynom 9. Grades ist „schief.“

1.5 Rationale Interpolation

Wenn man anhand der zu interpolierenden Punkte vermuten kann, dass die gesuchte Funktion Polstellen oder Asymptoten für $x \rightarrow -\infty$ oder $x \rightarrow \infty$ besitzt, so kann dem Rechnung getragen werden, indem man mit einer rationalen Funktion interpoliert. Wir folgen dafür der Argumentation von [9].

Gegeben seien $n+1$ paarweise verschiedene Punkte (x_i, y_i) , so sucht man nun eine rationale Funktion $R(x)$ von der Form

$$R(x) = \frac{p_0 + p_1x + \cdots + p_\alpha x^\alpha}{q_0 + q_1x + \cdots + q_\beta x^\beta} = \frac{P_\alpha}{Q_\beta}, \quad (1.13)$$

wobei P_α und Q_β Polynome vom Grad α bzw. β sind. Aus den gegebenen Punkten lassen sich die Bedingungen $R(x_i) = \frac{P_\alpha(x_i)}{Q_\beta(x_i)} = y_i$, $i = 0, \dots, n$ für die Interpolation ableiten.

Diesen $n+1$ Gleichungen stehen $\alpha + \beta + 2$ Koeffizienten $p_0, \dots, p_\alpha, q_0, \dots, q_\beta$ gegenüber. Da Zähler und Nenner aber mit einer beliebigen Zahl ungleich Null multipliziert werden können, ist es zulässig einen Koeffizienten zu normieren und gleich Eins zu setzen. Ist die Bedingung $\alpha + \beta = n$ erfüllt, so muß ein Gleichungssystem mit $n+1$ Unbekannten und $n+1$ Bedingungen gelöst werden.

Es zeigt sich aber, dass - im Gegensatz zur Polynominterpolation - nicht immer eine rationale Funktion $R(x)$ existiert, die Lösung des Interpolationsproblems ist, auch wenn für das Gleichungssystem immer eine nicht-triviale Lösung existiert. Aus (1.13) folgt

$$\begin{aligned} P_\alpha(x_i) - y_i Q_\beta(x_i) &= 0 \\ &= p_0 + p_1 x_i + \cdots + p_\alpha x_i^\alpha - y_i (q_0 + q_1 x_i + \cdots + q_\beta x_i^\beta) = 0 \\ &\quad i = 0, \dots, n. \end{aligned} \quad (1.14)$$

Es ist nämlich möglich, dass eine Stützstelle x_k eine Nullstelle des Nenner-Polynoms Q_β ist. Aus (1.14) folgt in diesem Fall, dass auch $P_\alpha(x_k) = 0$ gelten muss. Somit kann durch $(x - x_k)$ gekürzt werden und die Interpolationsbedingung $R(x_k) = y_k$ wird möglicherweise nicht erfüllt. So ein Punkt wird *unerreichbarer Punkt* genannt.

Beispiel 4. Gegeben seien die Punkte $(-1, 2)$, $(1, 3)$ und $(2, 2)$, die durch eine Funktion mit dem Ansatz

$$R(x) = \frac{P_1(x)}{Q_1(x)} = \frac{p_0 + p_1 x}{q_0 + q_1 x}$$

interpoliert werden sollen. Aus (1.14) und der Angabe ergibt sich das Gleichungssystem

$$\begin{aligned} p_0 - p_1 - 2q_0 + 2q_1 &= 0 \\ p_0 + p_1 - 3q_0 - 3q_1 &= 0 \\ p_0 + 2p_1 - 2q_0 - 4q_1 &= 0. \end{aligned}$$

Da das Gleichungssystem bis auf einen Faktor bestimmt ist, setzen wir $q_0 = 1$ und erhalten dann die Lösungen $p_0 = 2$, $p_1 = -2$ und $q_1 = -1$. Es folgt somit

$$R(x) = \frac{2(1-x)}{(1-x)} = 2.$$

Diese Funktion erfüllt aber nicht die Interpolationsbedingung für den Punkt $(1, 3)$.

Auswege:

- Weglassen des „kritischen Punktes“:

Die berechnete Funktion $R(x)$ erfüllt diese Bedingung bereits. Sie ist aber meist nicht sehr „zufriedenstellend“.

- Neuerliche Berechnung mit einer anderen Ansatzfunktion:

Wählt man im obigen Beispiel den Ansatz

$$\tilde{R}(x) = \frac{\tilde{P}_0(x)}{\tilde{Q}_3(x)} = \frac{\tilde{p}_0}{\tilde{q}_0 + \tilde{q}_1 x + \tilde{q}_2 x^2}$$

und setzt $\tilde{q}_2 = 1$, so erhält man nach Lösen des Gleichungssystems die Werte $\tilde{p}_0 = 12$, $\tilde{q}_0 = 4$ und $\tilde{q}_1 = -1$. Die Funktion

$$\tilde{R}(x) = \frac{12}{4 - x + x^2}$$

erfüllt nun wirklich alle Interpolationsbedingungen.

Bemerkung 1. Es ist möglich, dass der Eins gesetzte Koeffizient gleich Null ist. Dann ist das Gleichungssystem nach dem angegebenen Verfahren natürlich nicht lösbar. Ist das Gleichungssystem nicht lösbar, so bedeutet dies, dass der Koeffizient gleich Null ist. Das daraus resultierende Gleichungssystem liefert dann die gewünschten Koeffizienten nach dem beschriebenen Verfahren.

Beispiel 5. Gegeben seien die Punkte $(-1, \frac{1}{2})$, $(0, 1)$ und $(1, \frac{1}{2})$, die durch eine rationale Funktion $R(x) = \frac{p_0}{q_0 + q_1 x + q_2 x^2}$ interpoliert werden. Setzt man die gegebenen Punkte ein, so erhält man das Gleichungssystem

$$\begin{aligned} 2p_0 - q_0 + q_1 - q_2 &= 0 \\ 2p_0 - q_0 - q_1 - q_2 &= 0 \\ p_0 - q_0 &= 0. \end{aligned}$$

Setzt man $q_1 = 1$, so sieht man sofort, dass Gleichung 1 und Gleichung 2 keine gemeinsamen reellen Schnittpunkte besitzen; daraus folgt, dass das Gleichungssystem nicht mehr zu lösen ist. Daher setzen wir $q_1 = 0$ und erhalten als neues Gleichungssystem

$$2p_0 - q_0 - q_2 = 0$$

$$p_0 - q_0 = 0.$$

Dieses Gleichungssystem ist wieder bis auf einen Faktor eindeutig bestimmt, so dass es zulässig ist, einen Koeffizienten gleich Eins zu setzen. (Gelänge man so wieder auf ein nichtlösbares System, würde das wieder bedeuten, dass auch dieser Koeffizient gleich Null zu setzen ist.) Wir setzen nun $p_0 = 1$ und erhalten somit $q_0 = 1$ und $q_2 = 1$. Die rationale Funktion $R(x) = \frac{1}{1+x^2}$ erfüllt alle Interpolationsbedingungen.

1.6 Beispiel und Bild

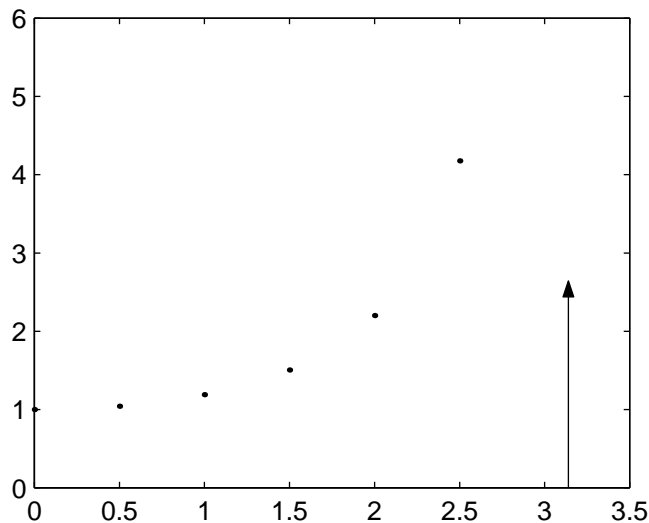


Abbildung 1.8: Zusammen mit der Polstelle bei $x = \pi$ erinnert diese Kurve stark an eine rationale Funktion.

Beispiel 6. Gegeben sei die Funktion f durch die (gerundeten) Punkte $(0, 1)$, $(0.5, 1.04)$, $(1, 1.19)$, $(1.5, 1.50)$, $(2, 2.20)$, $(2.5, 4.18)$. Darüberhinaus vermutet man einen Pol an der Stelle π . Da uns die Kurve (vgl. Abb. 1.8) sehr an ein rationales Polynom erinnert und rationale Polynome Polstellen erzeugen können, versuchen wir daher diese Funktion mittels einer mit dem Ansatz $R(x) = \frac{1}{ax+b}$ zu approximieren. Da $\lim_{x \rightarrow \pi} R(x) = \infty$ setzen wir den Nenner gleich Null. Als Bedingung erhalten wir daraus $b = -a\pi$. Unser neuer

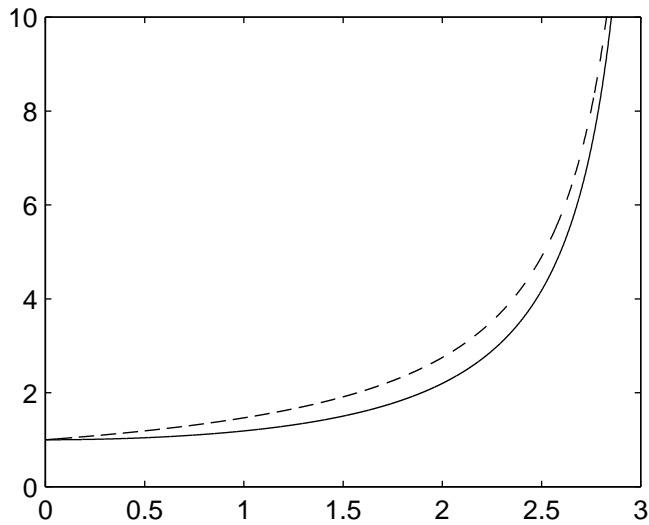


Abbildung 1.9: Approximation einer Funktion anhand der Gestalt durch ein entsprechendes rationales Polynom.

Ansatz lautet also $R(x) = \frac{1}{ax - a\pi}$. Da uns nur noch eine Variable zu bestimmen bleibt, können wir auch nur mehr für einen Punkt die Interpolation fordern. Wir entscheiden uns noch für den Punkt $(0,1)$ und erhalten das rationale Polynom $R(x) = \frac{1}{-\frac{x}{\pi} + 1}$ als Approximation für die $f(x) = \frac{x}{\sin(x)}$. Wie der direkte Vergleich in Abbildung 1.9 erkennen lässt, ist die Approximation für so ein „einfaches“ Polynom gar nicht schlecht. Zusammen mit der „Methode der kleinsten Quadrate“ (vgl. Kapitel 5), ließe sich der Fehler gleichmäßig über alle Stützstellen verteilen.

2 Tschebyscheff-Interpolation

Wir wollen nun eine mindestens $n + 1$ mal stetig differenzierbare Funktion $f(x)$ im Intervall $[a, b]$ untersuchen. Approximiert man die Funktion mit einem Interpolationspolynom, so lässt sich der Fehler wie folgt berechnen [12]:

Satz 1.

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{i=0}^n (x - x_i). \quad (2.1)$$

ξ bezeichne dabei eine von x abhängige Stelle mit $\xi \in [a, b]$.

Beweis. Wegen der Interpolationsbedingungen hat $f(x) - P_n(x)$ mindestens $n + 1$ Nullstellen x_0, x_1, \dots, x_n , so dass gilt

$$f(x) - P_n(x) = g(x)(x - x_0)(x - x_1) \dots (x - x_n). \quad (2.2)$$

Sei $\bar{x} \neq x_i$, $\bar{x} \in [a, b]$ ein beliebiger, aber fester Wert, dann ist

$$F(x) := f(x) - P_n(x) - g(\bar{x})(x - x_0)(x - x_1) \dots (x - x_n) \quad (2.3)$$

mindestens $n + 1$ mal stetig differenzierbar. Offensichtlich gilt $F(\bar{x}) = 0$. Nach dem Satz von Rolle gilt, dass es zwischen 2 Nullstellen einer Funktion mindestens eine Nullstelle der Ableitung dieser Funktion gibt. F' hat somit über dem Intervall $[a, b]$ mindestens $n + 1$ Nullstellen. Durch neuerliches Anwenden des Satzes von Rolle ist einsichtig, dass F'' mindestens n Nullstellen und F''' mindestens $n - 1$ Nullstellen besitzt. Daraus folgt, dass $F^{(n+1)}$ mindestens eine Nullstelle hat und somit gibt es ein ξ für das $F^{(n+1)}(\xi) = 0$. Für $F^{(n+1)}$ gilt

$$F^{(n+1)}(x) = f^{(n+1)}(x) - g(\bar{x})(n+1)!, \quad (2.4)$$

da P_n nach der $n + 1$ -ten Ableitung verschwindet und von $\prod_{i=0}^n (x - x_i)$ nur mehr $(n + 1)!$ übrig bleibt. Setzt man in diese Gleichung für x den Wert $\xi(x)$ ein, so erhält man

$$g(\bar{x}) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}. \quad (2.5)$$

Für jedes $\bar{x} \in [a, b]$ gilt somit

$$f(\bar{x}) - P_n(\bar{x}) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{i=0}^n (\bar{x} - x_i). \quad (2.6)$$

Da diese Gleichung trivialerweise auch für $\bar{x} = x_i$, $i = 0, \dots, n$ erfüllt ist, gilt die Gleichung für alle $x \in [a, b]$, sodass sich der Fehler durch

$$f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{i=0}^n (x - x_i) \quad (2.7)$$

abschätzen lässt.

□

Nun stellt sich die Frage, wie man den Fehler von vornherein möglichst gering halten kann. Nachdem die zu interpolierende Funktion nicht beeinflussbar ist, bietet sich der Weg über die Stützstellen an, der uns zu den Tschebyscheff-Polynomen führen wird.

2.1 Tschebyscheff-Polynome

Definition 1. Das Polynom $T_n(x)$, $x \in [-1, 1]$

$$T_n(x) = T_n(\cos(\varphi)) := \cos(n\varphi), \quad x = \cos\varphi, \quad \varphi \in [0, \pi] \quad (2.8)$$

heisst *Tschebyscheff-Polynom* oder *T-Polynom* vom Grad n .

Rekursionsformel: Aus dem Additionstheorem $\cos((n+1)\varphi) + \cos((n-1)\varphi) = 2 \cos(\varphi) \cos(n\varphi)$ folgt für Tschebyscheff-Polynome für $n > 1$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad T_0(x) = 1, \quad T_1(x) = x. \quad (2.9)$$

Eigenschaften der Tschebyscheff-Polynome

1. $|T_n(x)| \leq 1$ für $x \in [-1, 1]$.
 2. $T_n(x)$ besitzt $n+1$ Extrema an den Stellen $x_k^{(e)} = \cos\left(\frac{k\pi}{n}\right)$, $k = 0, 1, \dots, n$.
 3. $T_n(x)$ besitzt die n Nullstellen $x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right)$, $k = 1, \dots, n$. Diese Werte werden auch *Tschebyscheff-Abszissen* genannt.
- (2.10)

Tschebyscheff-Polynome haben eine weitere und bemerkenswerte Eigenschaft, die als *Minimax*-Eigenschaft bekannt ist. Als Vorbereitung sei hier die Definition von *normierten Polynomen* angegeben.

Definition 2. Ein Polynom heißt *normiert*, wenn der höchste Koeffizient des Polynoms gleich Eins ist, falls also für ein Polynom P

$$P(x) = \underbrace{a_n}_{=1} x^n + a_{n-1} x^{n-1} + \dots + a_1 x^1 + a_0 \quad (2.11)$$

gilt.

Satz 2 (Minimax-Eigenschaft). *Unter allen normierten Polynomen n -ten Grades ist das Tschebyscheff-Polynom $T_n(x)/2^{n-1}$ minimal bezüglich der Maximumnorm $\|P_n\|_{\max} := \max_{x \in [-1,1]} |P_n(x)|$ im Intervall $[-1, 1]$.*

Beweis. Indirekt:

Angenommen, es gebe ein normiertes Polynom n -ten Grades P_n für das $|P_n(x)| < \frac{1}{2^{n+1}}$ für alle $x \in [-1, 1]$ gilt. Dann gilt aber für das Polynom $Q(x) = P_n(x) - T_n(x)/2^{n-1}$ an den $n+1$ Extremstellen $x_k^{(e)}$ von T_n , dass es abwechselnd positiv und negativ wäre und daher mindestens n Nullstellen in $[-1, 1]$ hätte. Da aber P_n und T_n normierte Polynome sind, ist das Polynom Q höchstens vom Grad $n-1$. Ist $Q \neq 0$, dann kann Q keine n Nullstellen besitzen. Für $Q = 0$ sind T_n und P_n identisch. Damit ist die Annahme falsch und der Satz bewiesen.

□

Auf diesen Satz baut die Tschebyscheff-Interpolation auf.

2.2 Interpolationsverfahren

Voraussetzung des Verfahrens ist eine mindestens $n+1$ mal differenzierbare Funktion f definiert auf dem Intervall $[-1, 1]$. Die Idee der Tschebyscheff-Interpolation basiert darauf, den Fehler $f(x) - P_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \prod_{i=0}^n (x - x_i)$ möglichst klein zu halten. Dies ist durch geeignete Wahl der x_i möglich. Sei $\prod_{i=0}^n (x - x_i)$ ein normiertes Polynom $(n+1)$ -ten Grades. Nach dem Minimax-Satz ist das Betragsmaximum dieses Polynoms für alle $x \in [-1, 1]$ genau dann minimal, wenn die $n+1$ Stützstellen gleich den $n+1$ Nullstellen von T_{n+1} sind. Dann gilt nämlich $\max_{x \in [-1,1]} \left| \prod_{i=0}^n (x - x_i) \right| = 2^{-n}$. Für das Interpolationspolynom

2 Tschebyscheff-Interpolation

$P_n^*(x)$, dessen Stützstellen gleich den Tschebyscheff-Abszissen von T_{n+1} sind, erhält man somit den kleinstmöglichen Interpolationsfehler.

Als Darstellung von $P_n^*(x)$ bietet sich eine Linearkombination von Tschebyscheff-Polynomen an:

$$P_n^*(x) = \frac{1}{2}\gamma_0 T_0(x) + \sum_{j=1}^n \gamma_j T_j(x). \quad (2.12)$$

Für die Bestimmung der Koeffizienten $\gamma_0, \gamma_1, \dots, \gamma_n$ aus den Interpolationsbedingungen

$$\frac{1}{2}\gamma_0 T_0(x_l) + \sum_{j=1}^n \gamma_j T_j(x_l) = f(x_l), \quad l = 1, 2, \dots, n+1 \quad (2.13)$$

benötigen wir den folgenden Satz.

Satz 3. Seien x_l die $n+1$ Nullstellen von $T_{n+1}(x)$, dann gilt für $0 \leq k, j \leq n$

$$\sum_{i=1}^{n+1} T_k(x_i) T_j(x_i) = \begin{cases} 0, & \text{falls } k \neq j \\ \frac{1}{2}(n+1), & \text{falls } k = j > 0 \\ n+1, & \text{falls } k = j = 0. \end{cases} \quad (2.14)$$

Beweis. Setzen wir $h := \frac{\pi}{n+1}$, so erhalten wir

$$T_k(x_l) = \cos\left(k \frac{2l-1}{n+1} \frac{\pi}{2}\right) = \cos\left(kh\left(l - \frac{1}{2}\right)\right). \quad (2.15)$$

Mit Hilfe des Additionstheorems $\cos \alpha \cos \beta = \frac{1}{2}(\cos(\alpha - \beta) + \cos(\alpha + \beta))$ folgt

$$\begin{aligned} \sum_{l=1}^{n+1} T_k(x_l) T_j(x_l) &= \sum_{l=1}^{n+1} \cos\left(kh\left(l - \frac{1}{2}\right)\right) \cos\left(jh\left(l - \frac{1}{2}\right)\right) = \\ &= \frac{1}{2} \sum_{l=1}^{n+1} \left[\cos\left((k-j)h\left(l - \frac{1}{2}\right)\right) + \cos\left((k+j)h\left(l - \frac{1}{2}\right)\right) \right] = \\ &= \frac{1}{2} \operatorname{Re} \left[\underbrace{\sum_{l=1}^{n+1} e^{i(k-j)h(l-\frac{1}{2})}}_{=:q_1} + \underbrace{\sum_{l=1}^{n+1} e^{i(k+j)h(l-\frac{1}{2})}}_{=:q_2} \right]. \end{aligned} \quad (2.16)$$

Für $j \neq k$ gelten wegen $0 \leq k, j \leq n$ die Ungleichungen $1 \leq |k-j| \leq n$ und $1 \leq k+j < 2n$. Daher gelten wegen $\frac{\pi}{n+1} = h \leq |(k-j)h| \leq nh = \frac{\pi n}{n+1}$ und $\frac{\pi}{n+1} = h \leq (k+j)h < 2nh = \frac{2\pi n}{n+1}$ die Ungleichungen

$$|(k-j)h| \neq 2\pi l, \quad (k+j)h \neq 2\pi l \quad \text{für } l \in \mathbb{N}, \quad (2.17)$$

2 Tschebyscheff-Interpolation

sodass $q_1 \neq 1$ und $q_2 \neq 1$ folgt. Für die erste Summe erhalten wir mit $h(n+1) = \pi$ und wegen der Formel für die geometrische Reihe

$$\begin{aligned} \sum_{l=1}^{n+1} e^{i(k-j)h(l-\frac{1}{2})} &= e^{\frac{1}{2}i(k-j)h} \sum_{l=1}^{n+1} e^{i(k-j)h(l-1)} = e^{\frac{1}{2}i(k-j)h} \sum_{l=0}^n e^{i(k-j)hl} = \\ &= e^{\frac{1}{2}i(k-j)h} \cdot \frac{e^{i(k-j)h(n+1)} - 1}{e^{i(k-j)h} - 1} = \frac{(-1)^{k-j} - 1}{2i \sin(\frac{1}{2}(k-j)\frac{\pi}{n+1})}. \end{aligned} \quad (2.18)$$

Für die zweite Summe erhalten wir durch analoge Rechnung fast dasselbe Ergebnis. Im obigen Ergebnis ist nur $(j-k)$ durch $(j+k)$ zu ersetzen. Da in beiden Ausdrücken der Realteil Null ist, ist der Satz für $j \neq k$ bewiesen.

Für $j = k > 0$ liefert die erste Summe in (2.16) wegen $q_1 = 1$ den Wert $(n+1)$, während der zweite wegen $\frac{\pi}{n+1} < (k+j)h \leq \frac{2\pi n}{n+1}$ als geometrische Reihe gemäß des obigen Resultats mit $k = j$ den Wert Null hat. Daraus erhält man wegen (2.16) das Ergebnis $\frac{1}{2}(n+1)$.

Der Fall $j = k = 0$ folgt wegen $T_0 = 1$ direkt aus dem Satz.

□

Explizite Darstellung von γ_k

Multipliziert man die l -te Interpolationsbedingung (2.13) mit $T_k(x_l)$, wobei k ein fester Index mit $0 \leq k \leq n$ ist und addiert dann alle $n+1$ Gleichungen, so erhält man für $k = 0, 1, \dots, n$

$$\begin{aligned} \gamma_k &= \frac{2}{n+1} \sum_{l=1}^{n+1} f(x_l) T_k(x_l) = \\ &= \frac{2}{n+1} \sum_{l=1}^{n+1} f\left(\cos\left(\frac{(2l-1)\pi}{(n+1)2}\right)\right) \cos\left(k \frac{(2l-1)\pi}{(n+1)2}\right). \end{aligned} \quad (2.19)$$

Bemerkung 2. Da Tschebyscheff-Polynome nur über dem Intervall $[-1, 1]$ die Minimax-Eigenschaft besitzen, muss die Funktion $f(z)$, die über dem Intervall $[a, b]$ definiert ist, durch die lineare Transformation

$$x = \frac{2z - b - a}{b - a}, \quad z = \frac{x(b - a) + b + a}{2} \quad (2.20)$$

in eine Funktion auf das gewünschte Intervall $[-1, 1]$ übergeführt werden.

Bemerkung 3. Für die Tschebyscheff-Interpolation sind also die Stützstellen vorgegeben. Das bedeutet, dass dieses Verfahren nur für den Fall geeignet ist, dass man sich

die Stützstellen aussuchen kann. Nachträgliches Hinzunehmen von Messwerten ist nicht so einfach möglich, da die Tschebyscheff-Abszissen wieder neu berechnet und an diesen Stellen dann gemessen bzw. ausgewertet werden müssen. Dasselbe gilt, falls ein Polynom anderen Grades gewünscht wird. Da dies wohl meist zu umständlich ist, ist die Tschebyscheff-Interpolation eher im Bereich der Funktionsapproximation von Bedeutung, da hier die Funktion f , die approximiert werden soll, bekannt ist, wenn sie auch meist umständlich zu berechnen ist, und daher an beliebigen Stellen ausgewertet werden kann. Der Aufwand dabei ist „nur“ ein rechnerischer.

Bemerkung 4. Tschebyscheff-Polynome sind aber auch dem Phänomen der Oszillation bei hohem Grad unterworfen, da sie eben auch Polynome sind. Der Unterschied liegt nur darin, dass durch die Stützstellenwahl von vornherein versucht wird, den möglichen maximalen Fehler zu minimieren.

2.3 Beispiele und Bilder

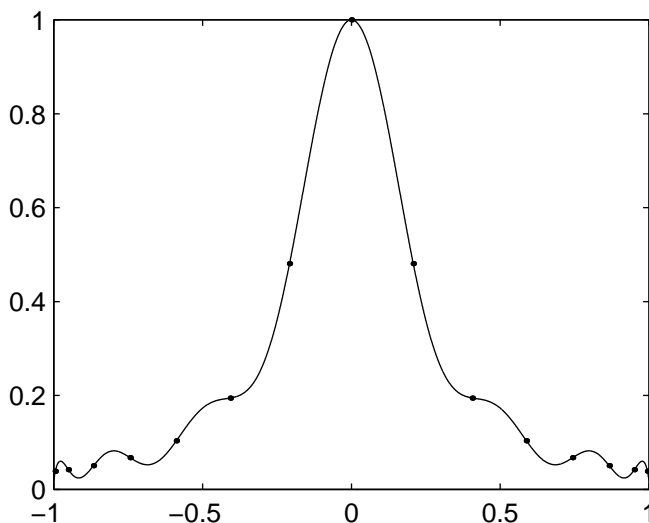


Abbildung 2.1: $\frac{1}{1+x^2}$ approximiert durch ein Tschebyscheff-Polynom aus 15 Stützstellen.

Beispiel 7. Betrachten wir wie in Kapitel 1 wieder die Aufgabe, die Funktion $f(z) = \frac{1}{1+z^2}$, $z \in [-5, 5]$ zu approximieren; diesmal mit einem Tschebyscheff-Polynom. Für ein Interpolationspolynom 14. Grades bestimmen wir die Nullstellen von T_{15} nach (2.10), Eigenschaft 3 und erhalten die 15 Interpolationsstellen $x_k = \cos\left(\frac{(2k-1)\pi}{30}\right)$ für $k = 1, \dots, 15$, an denen die transformierte Funktion, nämlich $\frac{1}{1+(5x_k)^2}$, $x \in [-2, 1]$, aus-

2 Tschebyscheff-Interpolation

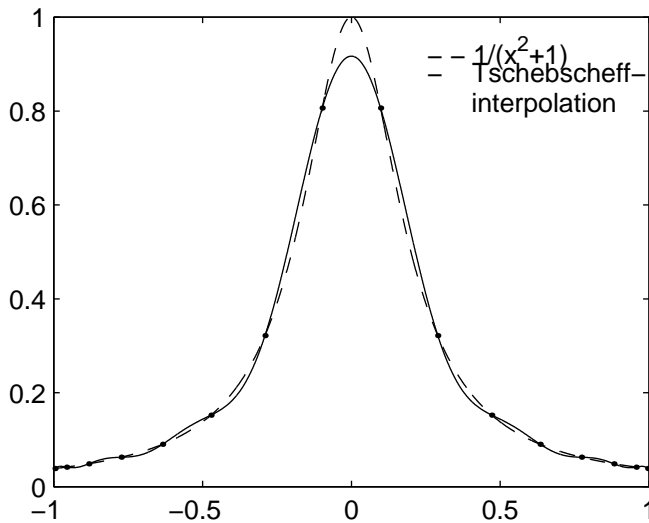


Abbildung 2.2: $\frac{1}{1+x^2}$ approximiert durch ein Tschebyscheff-Polynom aus 16 Stützstellen.

gewertet werden soll. Die zugehörige Tabelle lautet

x_k	$f(x_k)$	x_k	$f(x_k)$
0	1	± 0.743144	0.0675375
± 0.207911	0.480613	± 0.866025	0.0506329
± 0.406736	0.194709	± 0.951056	0.0423501
± 0.587784	0.103763	± 0.994521	0.0388699

Daraus berechnen wir γ_k nach (2.19). Eingesetzt in den Ansatz nach (2.13) ergibt sich das Polynom P_{14} , das in der üblichen Schreibweise

$$P_{14}(x) = -333.6537x^{14} + 1264.5349x^{12} - 1927.3623x^{10} + 1510.7608x^8 - 646.9421x^6 + 149.0472x^4 - 17.3659x^2 + 1$$

lautet. Abbildung 2.1 zeigt das Tschebyscheff-Polynom 14. Grades. Wie man eindeutig sehen kann, ist die Funktion verglichen mit dem Interpolationpolynom (Abb. 1.6) viel „ruhiger“. Das Oszillieren beschränkt sich auch an den Rändern auf ein erträglicheres Maß verglichen mit dem zitierten Interpolationspolynom.

Beispiel 8. Sehen wir uns nun ein Tschebyscheff-Polynom ungeraden Grades an. Die Stützstellen sind mit $x_k = \left(\frac{(2k-1)\pi}{32}\right)$ für $k = 1, \dots, 16$, die Funktionswerte mit $f(x_k) = \frac{1}{1+(5x)^2}$ vorgegeben. Das Polynom lautet

$$P(x) = -108.7969x^{14} + 239.5323x^{12} - 724.7714x^{10} + 627.4367x^8 - 305.6649906x^6 + 83.6621x^4 - 12.2794x^2 + 0.9168.$$

2 Tschebyscheff-Interpolation

Verglichen mit dem ungeraden Interpolationspolynom von Kapitel 1 ist das Tschebyscheff-Polynom durch 16. Stützstellen nicht so schief und nicht so unruhig. Das ist auch nicht verwunderlich, denn dieses Polynom ist wegen der symmetrischen liegenden Stützstellen auch selbst symmetrisch. Andererseits wird die Funktion f im Bereich des Punktes $(0, 1)$ nicht so gut angenähert. Das hat natürlich damit zu tun, dass das Interpolationspolynom in Abbildung 1.7 den Punkt $(0, 1)$ interpoliert. Während das Tschebyscheff-Polynom den y -Wert 1 nicht erreicht, schießt das Interpolationspolynom links des Punktes $(0, 1)$ zu weit hinaus. Zum Vergleich ist hier auch noch einmal die Funktion $f = \frac{1}{1+x^2}$ in strichlierter Form eingezeichnet.

3 Spline-Interpolation

Approximiert man eine tabellarisch gegebene Funktion mit einem Interpolationspolynom, so muss man selbst für Tschebyscheff-Polynome mit „minimaler“ Oszillation rechnen. Dieses Oszillieren tritt aber erst bei Polynomen höheren Grades auf. Die Spline-Methode unterbindet genau dies, indem zwischen den einzelnen Stützstellen jeweils ein eigenes Polynom niedrigen Grades berechnet wird. Die nun gesuchte Funktion soll aus mehreren aneinandergereihten Polynomen bestehen, die in den Stützpunkten bestimmte Eigenschaften zu erfüllen haben, damit Stetigkeit und Differenzierbarkeit auch in diesen Punkten gewährleistet sind. Im Prinzip kann der Grad der Polynome beliebig gewählt werden. Hier werden aber nur lineare und kubische Splines behandelt.

Lineare Splines haben den Vorteil, dass sie sehr einfach gebaut sind, man mit ihnen leicht rechnen kann und absolut keine unerwünschten Nebeneffekte, wie zum Beispiel das Oszillieren auftreten können, wodurch sie auch als Approximation für Funktionen, über die man so gut wie nichts weiß, geeignet sind.

Kubische Splines hingegen sind nicht nur stetig, sondern auch in den Stützstellen 2 mal stetig differenzierbar. Das bedeutet, dass sie auch „für das Auge“ eine gute Approximation bieten, aber Dank des niedrigen Grades doch noch nicht oszillieren.

3.1 Lineare Splines

Nach [7] seien $n + 1$ Punkte (x_i, y_i) , $i = 0, \dots, n$ gegeben, wobei $x_0 < x_1 < \dots < x_{n-1} < x_n$ gilt, die durch einen linearen Spline $s(x)$ zu interpolieren sind. Auf den Teilintervallen $[x_i, x_{i+1}]$ lässt sich die Funktion $s(x) \mid_{[x_i, x_{i+1}]} = s_i(x)$ wie folgt darstellen:

$$s_i(x) = \frac{(x_{i+1} - x)y_i + (x - x_i)y_{i+1}}{x_{i+1} - x_i}. \quad (3.1)$$

Offensichtlich ist $s_i(x)$ eine lineare Funktion, die wegen $s_i(x_i) = y_i$ und $s_i(x_{i+1}) = y_{i+1}$ stetig ist und die Interpolationsbedingungen erfüllt.

3.1.1 Spezielle Eigenschaften

Lineare Splines bestechen gerade durch ihre „Einfachheit.“ Für eine „gute“ Approximation müssen daher viele Stützpunkte gewählt werden. Dafür gibt es aber absolut kein

Oszillieren. Ganz im Gegenteil - Dank der Linearität sind Bereiche in der Umgebung von Asymptoten oder Polen kein Problem, da die zweite Ableitung an diesen Stellen gegen Null geht und die Krümmung von linearen Splines gleich Null ist. Das bedeutet andererseits aber auch, dass in Umgebungen starker Krümmung oder zumindest vermuteter starker Krümmung mehr Stützpunkte gewählt werden müssen, damit eine sinnvolle Interpolation möglich ist. Ebenso ist es natürlich ratsam, in Umgebungen steiler Anstiege einer Kurve die Abstände $h_i = x_{i+1} - x_i$ gering zu halten.

Zusätzlich haben lineare Splineinterpolierende Eigenschaften, die Splines höherer Ordnung nicht mehr bieten können [11]:

- **Positivität bleibt erhalten:**
Gilt $\forall i \ y_i > 0$, so ist auch der lineare Spline $s(x) > 0 \ \forall i$.
- **Monotonie bleibt erhalten:**
Sei $y_j < y_{j+1} < \dots < y_{k-1} < y_k$, $1 \leq j < k \leq 0, \dots, n$ für $j < k$, so gilt im gesamten Intervall $[x_j, x_k]$, dass $s'(x) > 0 \ \forall x \in [x_j, x_k]$.
- **Konvexität bleibt erhalten:**
 $s'_i := s'(x)|_{(x_i, x_{i+1})}$ sei die Ableitung des Splines auf dem Teilintervall (x_i, x_{i+1}) . Sind die Daten konvex im Sinne von $s'_j < s'_{j+1} < \dots < s'_{k-1} < s'_k$, wobei für j und k die Ungleichungen $0 \leq j < k < n$ gelten, so ist die Splinefunktion s eingeschränkt auf $[x_j, x_k + 1]$ ebenfalls konvex.

3.1.2 Beispiele und Bilder

Die Anzahl der Stützpunkte für einen linearen Spline ist von großer Bedeutung. Im Gegensatz zu Interpolationspolynomen gilt, je mehr Punkte desto besser. Eine geringe Stützstellenzahl kann, besonders wenn nicht wichtige Punkte wie Extrema oder Wendepunkte, bzw. Punkte in deren Nähe angegeben werden, nur eine sehr ungenaue und ungefähre Vorstellung des Kurvenverlaufs vermitteln.

Beispiel 9. Die Funktion $f(x) = \frac{1}{1+x^2}$ soll durch die 4 Punkte $(-5, \frac{1}{26})$, $(-3, \frac{1}{10})$, $(\frac{1}{2}, \frac{1}{1.25})$ und $(1, \frac{1}{2})$ approximiert werden (Abb. 3.1). Selbst wenn durch das Maximum und wenige weitere Punkte interpoliert wird, ist das Ergebnis auch noch nicht viel besser. In Abbildung 3.2 sind die Punkte $(-5, \frac{1}{26})$, $(-3, \frac{1}{10})$, $(0, 1)$, $(3, \frac{1}{10})$ und $(5, \frac{1}{26})$ und der zugehörige lineare Spline gezeichnet. Allerdings sind die Symmetrie und die allgemeine Tendenz schon einigermaßen erkenn- oder vermutbar.

Beispiel 10. Approximationen durch lineare Splines können sich - wie die nächsten Beispiele zeigen - durchaus sehen lassen. Zum Vergleich zeigt Abbildung 3.3 die Funktion

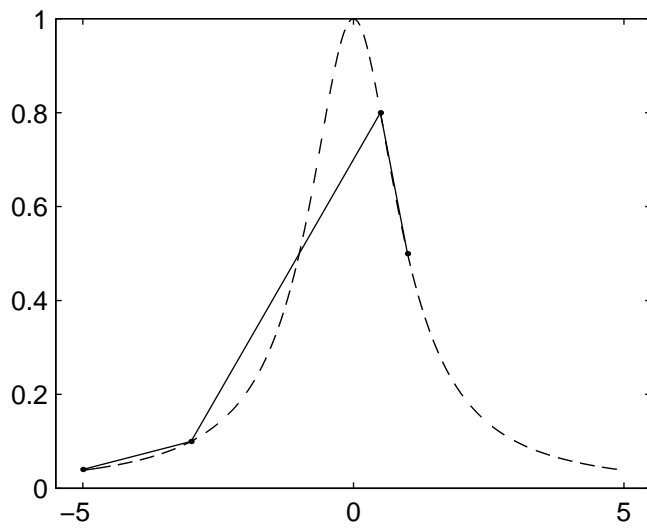


Abbildung 3.1: $\frac{1}{1+x^2}$ linear approximiert durch 4 Punkte.

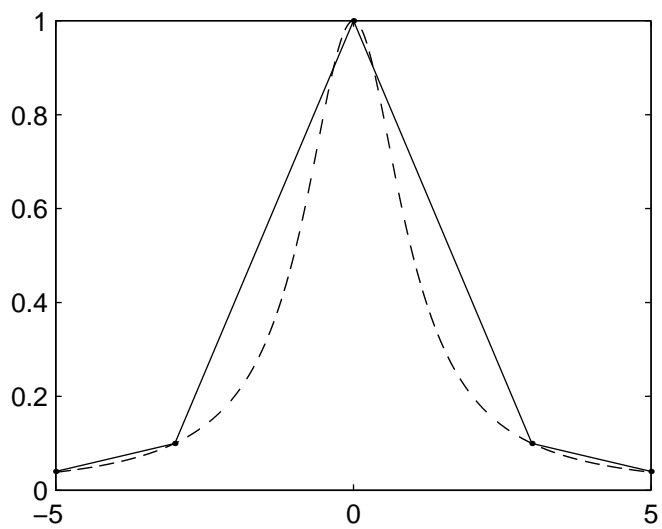


Abbildung 3.2: $\frac{1}{1+x^2}$ linear approximiert durch das Maximum und wenige symmetrische Punkte.

3 Spline-Interpolation

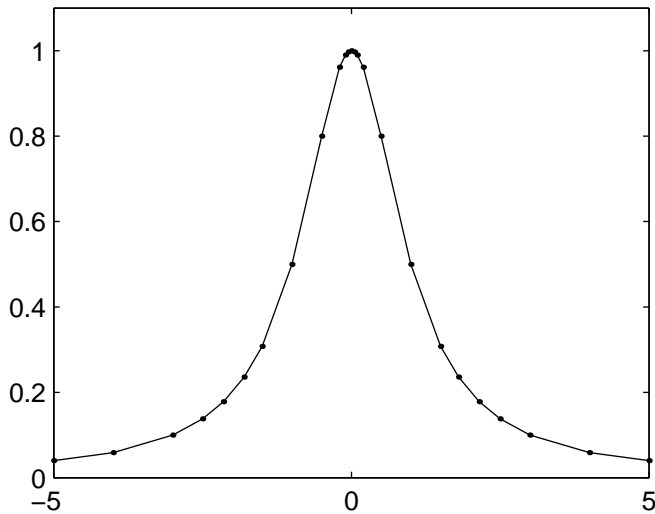


Abbildung 3.3: $\frac{1}{1+x^2}$ linear approximiert durch 27 Punkte.

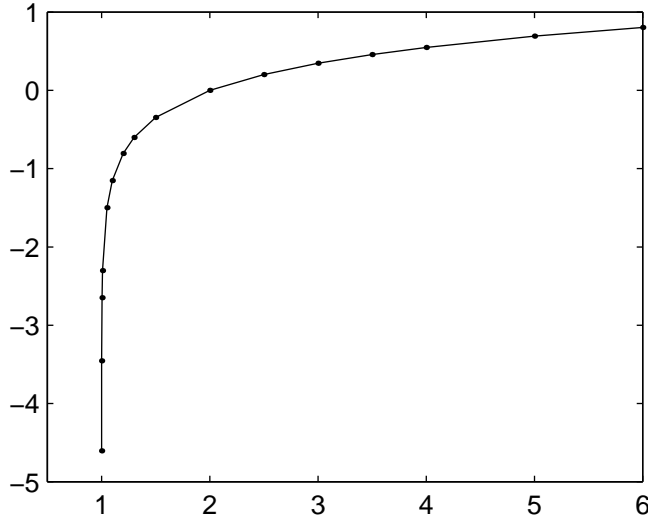
$f(x) = \frac{1}{1+x^2}$ nun durch die folgenden 27 symmetrisch liegenden Punkte approximiert.

x	$f(x)$	x	$f(x)$	x	$f(x)$	x	$f(x)$
0	1	± 0.5	0.8	± 1.8	0.2358	± 4	0.0625
± 0.05	0.9975	± 1	0.5	± 2.15	0.1779	± 5	0.04
± 0.1	0.9901	± 1.3	0.3717	± 2.5	0.1379		
± 0.2	0.9615	± 1.5	0.3077	± 3	0.1		

Beispiel 11. Auch die Funktion $\frac{1}{2}\log(x-1)$ ist, durch einen linearen Spline aus 16 Punkten interpoliert, als logarithmische Funktion wiederzuerkennen (vgl. 3.4). Die zugehörige Wertetabelle lautet:

x	$f(x)$	x	$f(x)$	x	$f(x)$	x	$f(x)$
1.0001	-4.6052	1.05	-1.4979	1.5	-0.3466	3.5	0.4581
1.001	-3.4539	1.1	-1.1513	2	0	4	0.5493
1.005	-2.6492	1.2	-0.8047	2.5	0.2027	5	0.6931
1.01	-2.3026	1.3	-0.6020	3	0.3466	6	0.8047

Bemerkung 5. Lineare Splines sind auch unter dem Begriff „lineare Interpolation“ bekannt.

Abbildung 3.4: Lineare Spline-Interpolation für $\frac{1}{2}\log(x-1)$

3.2 Kubische Splines

Für kubische Splines folgen wir [12]. Gegeben seien wieder $n+1$ Punkte (x_i, y_i) , $i = 0, \dots, n$, wobei $x_0 < x_1 < \dots < x_{n-1} < x_n$ gelten muss. Das Intervall $[x_0, x_n]$ wird nun in die Teilintervalle $[x_i, x_{i+1}]$ mit der Länge

$$h_i = x_{i+1} - x_i \quad (3.2)$$

zerlegt. Für jedes dieser Intervalle wird ein Polynom 3. Grades berechnet, für das folgender Ansatz gewählt wird:

$$s_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (3.3)$$

Gesucht sind also die Koeffizienten a_i, b_i, c_i und d_i , die sich aus folgenden Überlegungen ableiten lassen: Der Funktionswert an der Stellen x_i und x_{i+1} muss gleich y_i beziehungsweise y_{i+1} sein:

$$s_i(x_i) = 0 + 0 + 0 + d_i = y_i \quad (3.4)$$

$$s_i(x_{i+1}) = a_i h_i^3 + b_i h_i^2 + c_i h_i + d_i = y_{i+1} \quad (3.5)$$

Weiters sollen aber nicht nur der Funktionswert im Berührungspunkt der Polynome s_i und s_{i+1} , sondern auch die erste und zweite Ableitung der sich berührenden Polynome s_i und s_{i+1} übereinstimmen. Wir verwenden im Weiteren folgende, in diesem Zusammenhang übliche Schreibweise:

Schreibweise 1.

$$y_i'' := s_i''(x_i), \quad i = 0, 1, \dots, n-1, \quad (3.6)$$

$$y_n'' := s_{n-1}(x_n). \quad (3.7)$$

y_i'' , $i = 0, 1, \dots, n$ ist also die zweite Ableitung im Punkt x_i .

Wird s_i differenziert und an den Stellen x_i und x_{i+1} ausgewertet, so erhält man:

$$s_i'(x_i) = 0 + 0 + c_i \quad (3.8)$$

$$s_i'(x_{i+1}) = 3a_i h_i^2 + 2b_i h_i + c_i = s_{i+1}'(x_{i+1}) \quad (3.9)$$

$$s_i''(x_i) = 0 + 2b_i \quad (3.10)$$

$$s_i''(x_{i+1}) = 6a_i h_i + 2b_i = s_{i+1}''(x_{i+1}). \quad (3.11)$$

Aus diesen Gleichungen lassen sich durch Einsetzen die Koeffizienten berechnen. Trivialerweise ist wegen (3.4)

$$d_i = y_i. \quad (3.12)$$

b_i erhält man aus (3.10).

$$b_i = y_i''/2. \quad (3.13)$$

Setzt man (3.13) in (3.11) ein, so erhält man

$$a_i = \frac{1}{6h_i}(y_{i+1}'' - y_i''). \quad (3.14)$$

Schließlich werden noch alle so berechneten Koeffizienten in (3.5) eingesetzt:

$$c_i = \frac{1}{h_i}(y_{i+1} - y_i) - \frac{h_i}{6}(y_{i+1}'' + 2y_i''). \quad (3.15)$$

Nachdem die Funktionswerte y_i , $i = 0, \dots, n$ gegeben sind, fehlen nur mehr die Werte der zweiten Ableitungen y_i'' , $i = 1, \dots, n-1$. Hier verwenden wir die letzte der Interpolationsbedingungen, nämlich die Stetigkeit der ersten Ableitung. Laut Voraussetzung ist $s_i'(x_{i+1}) = y_{i+1}'$ und $s_{i+1}'(x_{i+1}) = y_{i+1}'$. Die in (3.13) bis (3.15) berechneten Koeffizienten werden in (3.9) eingesetzt, sodass man

$$s_i'(x_{i+1}) = \frac{1}{h_i}(y_{i+1} - y_i) + \frac{h_i}{6}(2y_{i+1}'' + y_i'') \quad (3.16)$$

3 Spline-Interpolation

erhält. Analog muss natürlich auch für die Polynomfunktion s_{i-1} gelten:

$$s'_{i-1}(x_i) = \frac{1}{h_{i-1}}(y_i - y_{i-1}) + \frac{h_{i-1}}{6}(2y''_i + y''_{i-1}). \quad (3.17)$$

Aus der Bedingung $s'_{i-1}(x_i) = y'_i = s'_i(x_i)$ folgt wegen (3.17), (3.8) und (3.15)

$$\frac{1}{h_{i-1}}(y_i - y_{i-1}) + \frac{h_{i-1}}{6}(2y''_i + y''_{i-1}) = \frac{1}{h_i}(y_{i+1} - y_i) - \frac{h_i}{6}(y''_{i+1} + 2y''_i). \quad (3.18)$$

Bringt man alles auf eine Seite, multipliziert mit 6 und ordnet nach y''_{i-1} , y''_i und y''_{i+1} , so erhält man

$$h_{i-1}y''_{i-1} + 2(h_{i-1} + h_i)y''_i + h_iy''_{i+1} - \frac{6}{h_i}(y_{i+1} - y_i) + \frac{6}{h_{i-1}}(y_i - y_{i-1}) = 0 \quad (3.19)$$

als Bedingung für die inneren Stützstellen x_i , $i = 1, \dots, n-1$. Abgesehen von den beiden Unbekannten y''_0 und y''_n muss für die Bestimmung der Unbekannten y''_i nun nur mehr ein Gleichungssystem gelöst werden.

Definition 3. Um das zu lösende Gleichungssystem in Matrizenform schreiben zu können, definieren wir die „Splinematrix“ S und die Vektoren y'' und h :

$$S := \begin{pmatrix} 2(h_0 + h_1) & h_1 & & & & \\ h_1 & 2(h_1 + h_2) & h_2 & & & \\ & h_2 & 2(h_2 + h_3) & h_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & h_{n-3} & 2(h_{n-3} + h_{n-2}) & h_{n-2} \\ & & & & h_{n-2} & 2(h_{n-2} + h_{n-1}) \end{pmatrix} \quad (3.20)$$

$$y'' := \begin{pmatrix} y''_1 \\ y''_2 \\ y''_3 \\ \vdots \\ y''_{n-2} \\ y''_{n-1} \end{pmatrix}, \quad h := \begin{pmatrix} \frac{6}{h_1}(y_2 - y_1) - \frac{6}{h_0}(y_1 - y_0) - h_0y''_0 \\ \frac{6}{h_2}(y_3 - y_2) - \frac{6}{h_1}(y_2 - y_1) \\ \frac{6}{h_3}(y_4 - y_3) - \frac{6}{h_2}(y_3 - y_2) \\ \vdots \\ \frac{6}{h_{n-2}}(y_{n-1} - y_{n-2}) - \frac{6}{h_{n-3}}(y_{n-2} - y_{n-1}) \\ \frac{6}{h_{n-1}}(y_n - y_{n-1}) - \frac{6}{h_{n-2}}(y_{n-1} - y_{n-2}) - h_{n-1}y''_n \end{pmatrix}. \quad (3.21)$$

3 Spline-Interpolation

Damit lautet das Gleichungssystem:

$$S \cdot y'' = h. \quad (3.22)$$

Schreibweise 2. Da die Matrix S und der Vektor h in dieser Gestalt unhandlich beziehungsweise unübersichtlich sind, möchte ich folgende 2 Schreibweisen einführen:

$$\begin{aligned} 1.) \quad h_{i-1}^i &:= h_i + h_{i-1} \\ 2.) \quad \Delta_i &:= y_i - y_{i-1} \end{aligned} \quad (3.23)$$

Somit sehen die Matrix S und der Vektor h folgendermaßen aus:

$$S = \begin{pmatrix} 2h_0^1 & h_1 & & & \\ h_1 & 2h_1^2 & h_2 & & \\ & h_2 & 2h_2^3 & h_3 & \\ & & \ddots & \ddots & \ddots \\ & & & h_{n-3} & 2h_{n-3}^{n-2} & h_{n-2} \\ & & & & h_{n-2} & 2h_{n-2}^{n-1} \end{pmatrix}, \quad h = \begin{pmatrix} \frac{6}{h_1}\Delta_2 - \frac{6}{h_0}\Delta_1 - h_0 y_0'' \\ \frac{6}{h_2}\Delta_3 - \frac{6}{h_1}\Delta_2 \\ \frac{6}{h_3}\Delta_4 - \frac{6}{h_2}\Delta_3 \\ \vdots \\ \frac{6}{h_{n-2}}\Delta_{n-1} - \frac{6}{h_{n-3}}\Delta_{n-2} \\ \frac{6}{h_{n-1}}\Delta_n - \frac{6}{h_{n-2}}\Delta_{n-1} - h_{n-1} y_n'' \end{pmatrix}. \quad (3.24)$$

Die Matrix S hat wegen ihrer besonderen Gestalt maximalen Rang. Um dies zu zeigen, bedarf es der folgenden Definition sowie des nachfolgenden Satzes.

Definition 4. Eine Matrix A heißt *diagonal dominant*, falls der Betrag jedes Diagonalelementes größer ist, als die Summe der restlichen Elemente in derselben Zeile, also falls

$$|a_{ii}| > \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|, \quad i = 1, 2, \dots, n. \quad (3.25)$$

Satz 4. Jede *diagonal dominante Matrix* hat maximalen Rang.

Beweis. Gezeigt wird, dass nach einem Gauss-Eliminationsschritt die Matrix immer noch diagonal dominant ist. \bar{a}_{ik} bezeichne das neue Element an der Stelle von a_{ik} nach dem ersten Schritt. Wir beginnen mit Element a_{11} und eliminieren die restlichen Elemente in der ersten Spalte. Wegen

$$\begin{pmatrix} a_{11} & a_{1k} \\ a_{i1} & a_{ik} \end{pmatrix} \rightarrow \begin{pmatrix} a_{11} & a_{1k} \\ 0 & a_{ik} - \frac{a_{i1}a_{1k}}{a_{11}} \end{pmatrix} \quad (3.26)$$

gilt allgemein für die reduzierten Elemente

$$\bar{a}_{ik} = a_{ik} - \frac{a_{i1}a_{1k}}{a_{11}}. \quad (3.27)$$

Für die Diagonalelemente folgt daraus die Abschätzung

$$|\bar{a}_{ii}| = \left| a_{ii} - \frac{a_{i1}a_{1i}}{a_{11}} \right| \geq |a_{ii}| - \left| \frac{a_{i1}a_{1i}}{a_{11}} \right|, \quad i, k = 2, 3, \dots, n. \quad (3.28)$$

Für die Summe der Beträge der Nichtdiagonalelemente gilt

$$\begin{aligned} \sum_{\substack{k=2 \\ k \neq i}}^n |\bar{a}_{ik}| &= \sum_{\substack{k=2 \\ k \neq i}}^n \left| a_{ik} - \frac{a_{i1}a_{1k}}{a_{11}} \right| \leq \sum_{\substack{k=2 \\ k \neq i}}^n |a_{ik}| + \left| \frac{a_{i1}}{a_{11}} \right| \sum_{\substack{k=2 \\ k \neq i}}^n |a_{1k}| = \\ &= \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| - |a_{i1}| + \left| \frac{a_{i1}}{a_{11}} \right| \left[\sum_{k=2}^n |a_{1k}| - |a_{1i}| \right] < \\ &< |a_{ii}| - |a_{i1}| + \left| \frac{a_{i1}}{a_{11}} \right| [|a_{11}| - |a_{1i}|] = |a_{ii}| - \left| \frac{a_{i1}a_{1i}}{a_{11}} \right| \leq |\bar{a}_{ii}|. \end{aligned} \quad (3.29)$$

Damit ist gezeigt, dass die neue Matrix \bar{A} wieder diagonal dominant ist und somit ist der Satz bewiesen.

□

Aus diesem Satz folgt, dass das Gleichungssystem $S \cdot y'' = h$ eindeutig lösbar ist, wenn y_0'' und y_n'' gegeben sind. Dieser Fall ist in der Praxis eher selten. Die verschiedenen Möglichkeiten, mit diesem Problem umzugehen, werden im nächsten Abschnitt behandelt. Dann wird es sich als nützlich erweisen, nicht nur die spezielle „tridiagonale“ Form der Splinematrix sondern die allgemeine, diagonal dominante Form betrachtet zu haben. Es wird sich nämlich herausstellen, dass auch die im Folgenden auftretenden Matrizen diagonal dominant sind.

3.2.1 Verschiedene Splinearten

Je nachdem wie man mit den noch unbekannten Werten y_0'' und y_n'' umgeht, lassen sich verschiedene Arten von Splines ableiten:

- natürliche Splines
- allgemeine Splines
- periodische Splines

- Splines mit vorgegebener erster Ableitung
- Splines mit vorgegebener dritter Ableitung
- Splines mit not-a-knot Randbedingung
- Parabolische Splines
- Parametrische kubische Splines

Bemerkung 6. Im Folgenden wird mehrmals davon geredet, dass gewisse Werte „bekannt“ sind. In der Praxis heißt das natürlich, dass nicht nur bekannte, sondern auch „gut“ abschätzbare Werte verwendet werden können. Letztlich muss man immer noch überprüfen, ob das Ergebnis „gefällt“ oder nicht. Ist man nicht zufrieden, so ist eine andere Interpolationsmethode zu wählen, beziehungsweise, falls möglich, sind noch Werte für die nicht gefälligen Bereiche zu erheben.

Natürliche Splines

Diese Möglichkeit ist die einfachste, da y_0'' und y_n'' gleich Null gesetzt werden und nur mehr das Gleichungssystem mit den $n - 1$ Variablen y_i'' gelöst werden muss. Ausgegangen wird von der Annahme, dass die Krümmung im Anfangs- und Endpunkt gleich Null ist. Man stelle sich vor, eine biegsame Latte wird gezwungen, gewisse Punkte (die Stützstellen) zu durchlaufen, wobei die Latte am Anfangs- und Endpunkt frei ist und sie die innere Spannung minimiert, indem sie die geringste Krümmung „wählt.“ (Siehe Abschnitt 3.2.2).

Allgemeine Splines

Sind die Werte y_0'' und y_n'' bekannt beziehungsweise gut abschätzbar, so ist es natürlich möglich, diese Werte einzusetzen und das Gleichungssystem $A \cdot y = h$ lösen.

Periodische Splines

Diese Möglichkeit ist für den Fall zu wählen, dass die zu interpolierende Funktion periodisch mit der Periode τ ist. Also gilt für diese Funktion $f(x) = f(x + \tau)$. Natürlich gelten nicht die gleichen Randbedingungen wie bei der allgemeinen Spline-Interpolation. Einmal müssen Anfangs- und Endpunkt übereinstimmen:

$$f(x_0) = f(x_0 + \tau) = f(x_n) \quad (3.30)$$

Anfangs- und Endpunkt der Periode müssen also gegeben sein. Aus $y_0 = y_n$ folgt natürlich $\Delta_n = y_n - y_{n-1} = y_0 - y_{n-1} = \Delta_0$, was in den folgenden Gleichungen zu beachten ist.

3 Spline-Interpolation

Zweitens sollen natürlich wieder die ersten und zweiten Ableitungen der Polynome s_0 und s_{n-1} übereinstimmen. Aus (3.19) folgt wegen der Periodizität

$$h_0 y_n'' + 2(h_0 + h_1) y_1'' + h_1 y_2'' = \frac{6}{h_1} \Delta_2 - \frac{6}{h_0} \Delta_1, \quad (3.31)$$

$$h_{n-1} y_{n-1}'' + 2(h_{n-1} + h_n) y_0'' + h_n y_1'' = \frac{6}{h_0} \Delta_1 - \frac{6}{h_{n-1}} \Delta_0. \quad (3.32)$$

Daraus ergibt sich für das Gleichungssystem $P \cdot y = h_p$

$$P := \begin{pmatrix} 2h_1^0 & h_1 & & & & & & h_0 \\ h_1 & 2h_2^1 & h_2 & & & & & \\ & h_2 & 2h_3^2 & h_3 & & & & \\ & & & & \ddots & \ddots & \ddots & \\ & & & & & & & h_{n-2} & 2h_{n-1}^{n-2} & h_{n-1} \\ h_0 & & & & & & & h_{n-1} & 2h_n^{n-1} \end{pmatrix} \quad (3.33)$$

$$y := \begin{pmatrix} y_1'' \\ \vdots \\ y_i'' \\ \vdots \\ y_n'' \end{pmatrix}, \quad h_p := \begin{pmatrix} \frac{6\Delta_2}{h_1} - \frac{6\Delta_1}{h_0} \\ \vdots \\ \frac{6\Delta_{i+1}}{h_i} - \frac{6\Delta_i}{h_{i-1}} \\ \vdots \\ \frac{6\Delta_1}{h_0} - \frac{6\Delta_0}{h_{n-1}} \end{pmatrix}. \quad (3.34)$$

Da P diagonal dominant ist, kann das Gleichungssystem eindeutig gelöst werden.

Splines mit vorgegebener erster Ableitung

Sind in den Randpunkten die ersten Ableitungen y_0' und y_n' bekannt, so lässt sich das Gleichungssystem wieder eindeutig lösen. Der anschauliche Vergleich dieser Methode mit der Latte ist auch hier möglich. Man muss sich die Enden wieder in den Randpunkten fixiert vorstellen und zwar so, dass eben genau die Steigungen y_0' und y_n' entstehen.

Differenziert man den Ansatz (3.3), so erhält man

$$s_0'(x_0) = c_0 = y_0'.$$

3 Spline-Interpolation

Wegen (3.15) und (3.2.1) gilt $c_0 = \frac{1}{h_0}(y_1 - y_0) - \frac{h_0}{6}(y_1'' + 2y_0'') = y_0'$ (3.15). Daraus folgt

$$h_0 y_0'' = \frac{3(y_1 - y_0)}{h_0} - 3y_0' - \frac{1}{2}h_0 y_1''. \quad (3.35)$$

Analog gilt

$$\begin{aligned} y_n' &= s_{n-1}'(x_n) = 3a_{n-1}(h_{n-1})^2 + 2b_{n-1}h_{n-1} + c_{n-1} = \\ &= \frac{1}{6}y_{n-1}''h_{n-1} + \frac{1}{3}h_{n-1}y_n'' + \frac{y_n - y_{n-1}}{h_{n-1}}. \end{aligned}$$

Daraus folgt

$$h_{n-1}y_n'' = 3(y_n' - \frac{1}{6}y_{n-1}''h_{n-1} - \frac{y_n - y_{n-1}}{h_{n-1}}). \quad (3.36)$$

Wird (3.35) in Gleichung (3.19) für $i = 1$ und analog (3.36) ebenfalls in Gleichung (3.19) aber diesmal für $i = n - 1$ eingesetzt, so erhält man

$$\begin{aligned} (\frac{3}{2}h_0 + 2h_1)y_1'' + h_1y_2'' &= 6\frac{y_2 - y_1}{h_1} - 9\frac{y_1 - y_0}{h_0} + 3y_0', \\ h_{n-2}y_{n-2}'' + (2h_{n-2} + \frac{3}{2}h_{n-1})y_{n-1}'' &= 9\frac{y_n - y_{n-1}}{h_{n-1}} - 6\frac{y_{n-1} - y_{n-2}}{h_{n-2}} - 3y_n'. \end{aligned} \quad (3.37)$$

Zu lösen ist also das Gleichungssystem

$$\begin{pmatrix} \frac{3}{2}h_0 + 2h_1 & h_1 & & & & \\ h_1 & 2h_2^1 & h_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & h_{i-1} & 2h_{i-1}^i & h_i & \\ & & & \ddots & \ddots & \ddots \\ & & & & h_{n-3} & 2h_{n-2}^{n-2} & h_{n-2} \\ & & & & h_{n-2} & 2h_{n-2} + \frac{3}{2}h_{n-1} & \end{pmatrix} \begin{pmatrix} y_1'' \\ y_2'' \\ \vdots \\ y_i'' \\ \vdots \\ y_{n-2}'' \\ y_{n-1}'' \end{pmatrix} = \begin{pmatrix} 6\frac{\Delta_2}{h_1} - 9\frac{\Delta_1}{h_0} + 3y_0' \\ \frac{6\Delta_3}{h_2} - \frac{6\Delta_2}{h_1} \\ \vdots \\ \frac{6\Delta_{i+1}}{h_i} - \frac{6\Delta_i}{h_{i-1}} \\ \vdots \\ \frac{6\Delta_{n-1}}{h_{n-2}} - \frac{6\Delta_{n-2}}{h_{n-3}} \\ 9\frac{\Delta_n}{h_{n-1}} - 6\frac{\Delta_{n-1}}{h_{n-2}} - 3y_n' \end{pmatrix}, \quad (3.38)$$

das wegen seiner diagonal dominanten Matrix eindeutig lösbar ist.

Splines mit vorgegebener dritter Ableitung

Für den Fall, dass die dritten Ableitungen an den Rändern gegeben sind, ist der Spline auch eindeutig lösbar. Sind die Randbedingungen $s'''(x_0) = y_0'''$ und $s'''(x_n) = y_n'''$ gegeben, so folgen aus (3.14), eingesetzt in die dritte Ableitung des Ansatzes (3.3), die Ausdrücke

3 Spline-Interpolation

$$\begin{aligned} s'''(x_0) &= s_0'''(x_0) = y_0''' = 6a_0 = (y_1'' - y_0'') \frac{1}{h_0}, \\ s'''(x_n) &= s_{n-1}'''(x_n) = y_n''' = 6a_{n-1} = (y_n'' - y_{n-1}'') \frac{1}{h_{n-1}}. \end{aligned}$$

Daraus ergeben sich die Gleichungen

$$y_0'' = y_1'' - h_0 y_0''', \quad y_n'' = y_{n-1}'' + h_{n-1} y_n''', \quad (3.39)$$

die eingesetzt in (3.19) für $i = 1$ bzw. $i = n - 1$

$$\begin{aligned} (3h_0 + 2h_1)y_1'' + h_1 y_2'' &= \frac{6(y_2 - y_1)}{h_1} + \frac{6(y_1 - y_0)}{h_0} + (h_0)^2 y_0''', \\ h_{n-2} y_{n-2}'' + (2h_{n-2} + 3h_{n-1})y_{n-1}'' &= \frac{6(y_n - y_{n-1})}{h_{n-1}} - \frac{6(y_{n-1} - y_{n-2})}{h_{n-2}} - (h_{n-1})^2 y_n''' \end{aligned} \quad (3.40)$$

als erste bzw. letzte Gleichung ergeben. Zusammen mit den inneren Gleichungen für Splines (3.19)

$$h_{i-1} y_{i-1}'' + 2(h_{i-1} + h_i) y_i'' + h_i y_{i+1}'' - \frac{6}{h_i} (y_{i+1} - y_i) + \frac{6}{h_{i-1}} (y_i - y_{i-1}) = 0 \quad (3.41)$$

für $i = 2, \dots, n - 2$.

ist diese Splineart nun eindeutig gegeben.

Splines mit not-a-knot Randbedingung

Diese Variante ist dann anzuwenden, wenn man ausser den Stützstellen keine weitere Information besitzt. y_0'' und y_n'' sind also unbekannt und es gibt keinen Grund sie gleich Null zu setzen. Weiters ist die Funktion nicht periodisch. In diesem Fall fehlen 2 Bedingungen, um das Gleichungssystem eindeutig zu lösen.

Man kann sich aber damit behelfen, dass man nur *eine* Polynomfunktion über die Intervalle $[x_0, x_2]$ bzw. $[x_{n-2}, x_n]$ aufstellt. Es gelte also

$$s_0(x) = s_1(x) \quad \text{und} \quad s_{n-2}(x) = s_{n-1}(x). \quad (3.42)$$

Als *Knoten* oder englisch *knots* bezeichnet man jene Stützstellen x_i , an denen zwei Polynome zusammenstoßen. Der Name 'not-a-knot' bezieht sich darauf, dass bei diesen Splines nicht alle Stützstellen auch Knoten sind. Die Stützpunkte x_1 und x_{n-1} sind nämlich keine Knoten, da diese Stellen eben nicht Randpunkte von zwei zusammenstoßenden Polynomen sind, sondern im Inneren eines Polynoms liegen. Daher folgt aus

3 Spline-Interpolation

dem gewählten Ansatz (3.3) $a_0 = a_1$ und $a_{n-2} = a_{n-1}$. Aus den Bedingungen für die Koeffizienten (vgl. (3.14)) folgt

$$\frac{1}{6h_0}(y_1'' - y_0'') = \frac{1}{6h_1}(y_2'' - y_1''),$$

$$y_0'' = y_1'' + \frac{h_0}{h_1}(y_1'' - y_2''). \quad (3.43)$$

Analog gilt auch

$$y_n'' = y_{n-1}'' + \frac{h_{n-1}}{h_{n-2}}(y_{n-1}'' - y_{n-2}''). \quad (3.44)$$

Nach (3.19) gilt

$$h_0 y_0'' + 2(h_0 + h_1)y_1'' + h_1 y_2'' = \frac{6}{h_1}(y_2 - y_1) - \frac{6}{h_0}(y_1 - y_0).$$

Substituiert man (3.43), so erhält man folgende linke Seite:

$$\begin{aligned} & h_0(y_1'' + \frac{h_0}{h_1}(y_1'' - y_2'')) + 2(h_0 + h_1)y_1'' + h_1 y_2'' = \\ & = h_0 y_1'' + \frac{(h_0)^2}{h_1}(y_1'' - y_2'') + 2(h_0 + h_1)y_1'' + h_1 y_2'' = \\ & = y_1''(h_0 + \frac{(h_0)^2}{h_1} + 2(h_0 + h_1)) + y_2''(h_1 - \frac{(h_0)^2}{h_1}) = \\ & = \frac{h_1 h_0 + (h_0)^2 + 2h_0 h_1 + 2(h_1)^2}{h_1} y_1'' + \frac{(h_1)^2 - (h_0)^2}{h_1} y_2'' = \\ & = (h_0 + 2h_1) \frac{h_1 + h_0}{h_1} y_1'' + (h_1 - h_0) \frac{h_1 + h_0}{h_1} y_2''. \end{aligned}$$

Multipliziert man jetzt noch mit $\frac{h_1}{h_1 + h_0}$, so heißt die umgeformte und vereinfachte Gleichung

$$(h_0 + 2h_1)y_1'' + (h_1 - h_0)y_2'' = \frac{6(y_2 - y_1)}{h_1 + h_0} - \frac{6h_1(y_1 - y_0)}{(h_1 + h_0)h_0}. \quad (3.45)$$

Durch eine analoge Umformung erhält man

$$(h_{n-2} - h_{n-1})y_{n-2}'' + (2h_{n-2} + h_{n-1})y_{n-1}'' = \frac{6h_{n-2}(y_n - y_{n-1})}{(h_{n-1} + h_{n-2})h_{n-1}} - \frac{6(y_{n-1} - y_{n-2})}{(h_{n-1} + h_{n-2})}. \quad (3.46)$$

Die inneren Gleichungen (vgl. 3.19) können wieder übernommen werden, sodass dieses Gleichungssystem eindeutig gegeben und wegen der diagonalen Dominanz eindeutig lösbar ist.

$$h_{i-1}y_{i-1}'' + 2(h_{i-1} + h_i)y_i'' + h_i y_{i+1}'' - \frac{6}{h_i}(y_{i+1} - y_i) + \frac{6}{h_{i-1}}(y_i - y_{i-1}) = 0 \quad (3.47)$$

für $i = 2, \dots, n-2$.

Parabolisch-kubische Splines

Ein andere Möglichkeit [6] für den Fall, dass keine weiteren Daten gegeben sind, stellen parabolisch-kubische Splines dar. Man macht aus der Not eine Tugend und verwendet an den Randstellen nicht kubische sondern parabolische Polynome. Dadurch verringert sich die Anzahl der Unbekannten um 2. Das heißt, dass die Splines eindeutig gegeben sind. Der Ansatz für die Randpolynome lautet also

$$\begin{aligned} s_0(x) &= b_0(x - x_0)^2 + c_0(x - x_0) + d_0, \\ s_{n-1}(x) &= b_{n-1}(x - x_{n-1})^2 + c_{n-1}(x - x_{n-1}) + d_{n-1}. \end{aligned} \quad (3.48)$$

Differenziert man die Ansatzpolynome der Ränder zwei mal, so erhält man

$$\begin{aligned} s_0''(x) &= 2b_0 = y_0'', \\ s_{n-1}''(x) &= 2b_{n-1} = y_{n-1}''. \end{aligned} \quad (3.49)$$

Offensichtlich sind die zweiten Ableitungen konstant und daher gilt:

$$y_0'' = y_1'' \quad \text{und} \quad y_{n-1}'' = y_n''. \quad (3.50)$$

Setzt man dies in (3.19) ein, erhält man die beiden Gleichungen

$$\begin{aligned} (3h_0 + 2h_1)y_1'' + h_1y_2'' &= \frac{6(y_1 - y_0)}{h_0} - \frac{6(y_2 - y_1)}{h_1}, \\ h_{n-2}y_{n-2}'' + (2h_{n-2} + 3h_{n-1})y_{n-1}'' &= \frac{6(y_{n-1} - y_{n-2})}{h_{n-2}} - \frac{6(y_n - y_{n-1})}{h_{n-1}}. \end{aligned} \quad (3.51)$$

Zusammen mit den inneren Gleichungen

$$\begin{aligned} h_{i-1}y_{i-1}'' + 2(h_{i-1} + h_i)y_i'' + h_iy_{i+1}'' - \frac{6}{h_i}(y_{i+1} - y_i) + \frac{6}{h_{i-1}}(y_i - y_{i-1}) &= 0 \\ \text{für } i &= 2, \dots, n-2 \end{aligned} \quad (3.52)$$

sind parabolisch-kubische Splines damit eindeutig gegeben. Auch diese Matrix ist diagonal dominant; das Gleichungssystem ist also eindeutig lösbar.

Parametrische kubische Splines

Parametrische kubische Splines [4] stellen einen Sonderfall dar, weil sie sich von anderen Splines nicht dadurch unterscheiden, wie die Unbekannten y_0'' und y_n'' behandelt werden. Vielmehr setzen sie sich aus 2 gleichen Splinearten zusammen. Diese Möglichkeit ist anzuwenden, wenn die gegebenen Punkte (x_i, y_i) nicht streng monoton geordnet sind und eine explizite Darstellung der Form $y = f(x)$ nicht möglich ist. Das ist zum Beispiel bei geschlossenen Kurven oder Kurven mit Doppelpunkten der Fall.

3 Spline-Interpolation

Die Koordinaten von x und y werden als Funktion eines Parameters $t \in [a, b]$ aufgefasst, wobei das Parameterintervall $[a, b]$ sowie der Parameterwert im Prinzip beliebig gewählt werden können:

$$x = x(t), \quad y = y(t). \quad (3.53)$$

Eine mögliche Wahl des Parameters ist zum Beispiel die Länge der Sehne von einem Stützpunkt zum nächsten:

$$t_0 = 0, \quad t_{i+1} = t_i + \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}. \quad (3.54)$$

Zu beachten ist, dass die Reihenfolge der Punkte beibehalten wird, da sich sonst der Kurvenverlauf verändert. Somit ist $x_i = x(t_i)$ bzw. $y_i = y(t_i)$. Die Funktionen $x(t)$ und $y(t)$ können nun in Form von Wertetabellen

$$\begin{array}{c|cccc} t & t_0 & t_1 & \cdots & t_n \\ \hline x & x_0 & x_1 & \cdots & x_n \end{array} \quad \text{und} \quad \begin{array}{c|cccc} t & t_0 & t_1 & \cdots & t_n \\ \hline y & y_0 & y_1 & \cdots & y_n \end{array}$$

dargestellt werden. Beide Funktionen $x(t)$ und $y(t)$ sind durch Splines zu interpolieren, sodass das Ergebnis als

$$S(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} s_x(t) \\ s_y(t) \end{pmatrix} \quad (3.55)$$

geschrieben werden kann.

Wahl der Splines

Bei gewünschten geschlossenen und glatten Kurven sind periodische kubische Splines für s_x und s_y zu wählen. Soll bei einer geschlossenen Kurve im Anfangs- bzw. Endpunkt ein nicht differenzierbarer Punkt auftreten, so können auch andere Splinearten gewählt werden. Für nicht geschlossene Kurven sind natürlich auch alle Splinearten möglich. Bei der Wahl der einzelnen Arten sind die gleichen Überlegungen anzustellen wie bei nicht parametrischen Splines.

Bemerkung 7. Theoretisch wäre es natürlich auch möglich die Funktionen $x(t)$ und $y(t)$ mittels Polynomfunktionen statt mit Splines zu interpolieren. Wegen der meist unerwünschten „Oszillation“ wird dies aber in der Praxis nicht gemacht.

3.2.2 Minimaleigenschaft der kubischen Splineinterpolierenden

Wir folgen dabei generell der Argumentation von [12]. Wie bereits erwähnt, lässt sich ein kubischer Spline als biegsame Latte deuten, leider aber nur mit gewissen Abstrichen, was wir jetzt genauer betrachten wollen. $l(x)$ beschreibe die Biegelinie der Latte, die an den Stützstellen x_i gelenkig gelagert ist und auf die dort keine äußeren Kräfte einwirken. Die Biegeenergie der Latte ist durch

$$E = c \int_a^b \left(\frac{l''(x)}{(1+l'(x)^2)^{\frac{3}{2}}} \right)^2 dx \quad (3.56)$$

gegeben, wobei c eine von geometrischen und physikalischen Größen abhängige Konstante ist. Die Latte verformt sich so, dass die Biegeenergie E minimal ist. Für kleine Auslenkungen $l'(x)$, für die $l'^2 \ll 1$ gilt, ist der Wert von E durch

$$E \approx c \int_a^b (l'')^2 dx \quad (3.57)$$

annäherbar. Der nächste Satz zeigt, dass in diesem Fall die Energie E minimal wird, wenn die Latte die Form $l(x) = s(x)$, also die Form bestimmter kubischer Splines, hat.

Die im folgenden Satz erwähnte Bedingung (3.58) ist, wie in der nachfolgenden Bemerkung 8 ausgeführt wird, für eine Reihe von Splines natürlicherweise erfüllt.

Satz 5. *Sei $s(x)$ eine kubische Splinefunktion, die eine durch die Stützpunkte (x_i, y_i) gegebene Funktion $f(x)$ an den Stützstellen $a = x_0 < x_1 < \dots, x_n = b$ interpoliert, und $l(x)$ eine beliebige in $[a, b]$ zweimal stetig differenzierbare Funktion, die denselben Interpolationsbedingungen wie $s(x)$ genüge, so dass*

$$[s''(x)(l'(x) - s'(x))]_{x=a}^b = 0. \quad (3.58)$$

Dann gilt

$$\int_a^b (l''(x))^2 dx \geq \int_a^b (s''(x))^2 dx, \quad (3.59)$$

d.h. für die kubische Splinefunktion $s(x)$ ist das Integral

$$\int_a^b (l''(x))^2 dx, \quad (3.60)$$

das bis auf eine Konstante dem angenäherten Wert E entspricht, minimal, falls die Bedingung (3.58) erfüllt ist.

3 Spline-Interpolation

Beweis. Da alle Funktionen in diesem Beweis Funktionen von x sind, wollen wir wegen der besseren Übersichtlichkeit statt $s(x)$, $l(x)$, $s'(x)$, $l'(x)$, $s''(x)$, $l''(x)$ und $s'''(x)$ nur s , l , s' , l' , s'' , l'' und s''' schreiben.

Aus der trivialen Aussage $l = s + (l - s)$ folgt $l'' = s'' + (l'' - s'')$ und somit

$$\int_a^b (l'')^2 dx = \int_a^b (s'')^2 dx + 2 \int_a^b s''(l'' - s'') dx + \int_a^b (l'' - s'')^2 dx.$$

Wegen $\int_a^b (l'' - s'')^2 dx \geq 0$ ist die Behauptung (3.59) bewiesen, falls das mittlere Integral verschwindet. Dies wird nun gezeigt. Mittels partieller Integration erhält man

$$\int_a^b s''(l'' - s'') dx = \underbrace{[s''(l' - s')]_a^b}_{=0} - \int_a^b s'''(l' - s') dx. \quad (3.61)$$

Dabei folgt $[s''(l' - s')]_a^b = 0$ wegen der Voraussetzung (3.58) des Satzes. Da $s(x)$ eine kubische Splinefunktion ist, ist $s''' := s_i$ eingeschränkt auf das Intervall $[x_i, x_{i+1})$ jeweils konstant. Somit gilt

$$\int_a^b s'''(l' - s') dx = \sum_{i=0}^{n-1} s_i \int_{x_i}^{x_{i+1}} (l' - s') dx = \sum_{i=0}^{n-1} s_i [l - s]_{x_i}^{x_{i+1}} = 0, \quad (3.62)$$

da für $l(x_i) = s(x_i)$. Das Gleichheitszeichen in (3.59) gilt nur für $l'' = s''$.

□

Bemerkung 8. Die Bedingung (3.58) des obigen Satzes ist insbesondere erfüllt für

- i) $s''(a) = s''(b) = 0$ (natürliche kubische Splines)
- ii) $s'(a) = l'(a)$ und $s'(b) = l'(b)$ (kubische Splines mit gegebener erster Ableitung)
- iii) $s(a) = s(b)$, $s'(a) = s'(b)$ und $s''(a) = s''(b)$ (periodische Splines)

Für die anderen Splinearten ist diese Bedingung nicht erfüllt. Der Unterschied zwischen den einzelnen Splinearten ist bei einer dem Problem entsprechend angepassten Wahl der noch zu bestimmenden Variablen eher gering. Während an den Rändern speziell durch Vorgabe bestimmter, stark differierender Werte die Unterschiede noch größer sein können, so wird der Unterschied im Inneren des Splines mehr und mehr vernachlässigbar gering. Hinzu kommt, dass speziell für eine größere Anzahl an Stützstellen unterschiedliche Randbedingungen auf die inneren Polynome immer weniger Auswirkung haben.

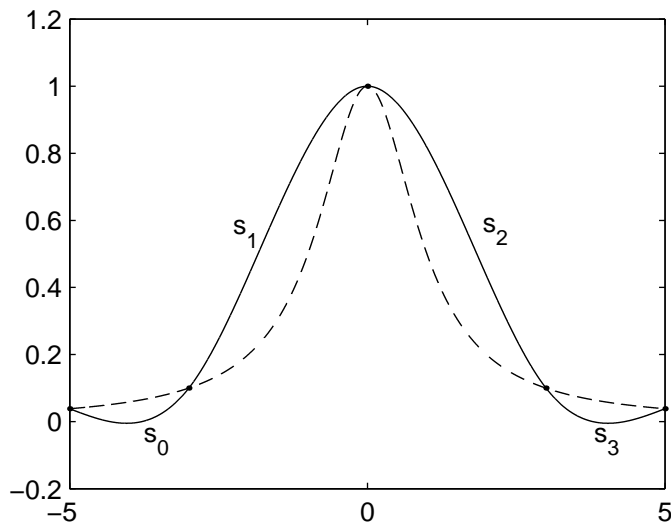


Abbildung 3.5: Noch nicht überzeugende Spline-Interpolation durch 5 Stützstellen für die Funktion $\frac{1}{1+x^2}$.

Bemerkung 9. Wichtig ist noch zu erwähnen, dass im Gegensatz zu einem Interpolationspolynom die Splinefunktion für wachsende Stützstellenanzahl $n + 1$ gegen die zu interpolierende Funktion konvergiert. Ebenso konvergieren die ersten und zweiten Ableitungen. Mit wachsender Knotenanzahl verringert sich also der Fehler [4]. Dies kann im Allgemeinen nicht für Interpolationspolynome behauptet werden.

3.2.3 Beispiele und Bilder

Wie bei der Polynominterpolation untersuchen wir auch diesmal die Spline-Methode anhand der Funktion $f(x) = \frac{1}{1+x^2}$. Die Ausgangswerte entsprechen denen aus Kapitel 3, Beispiel 3. Für 3 Punkte ist noch kein Unterschied zu merken; bei 5 Interpolationspunkten (Abb. 3.5) kann man schon eindeutig eine Verbesserung erkennen; man ist aber noch nicht zufrieden. Bereits bei 9 Punkten (ohne Abb.) kann man bereits mit der Annäherung zufrieden sein. Der Spline passt sich ohne unerwünschtes Schwingen an die Funktion f an. Original und Interpolationsfunktion sind in diesem Maßstab nur mehr schwer zu trennen. Dies ist auch in Abbildung 3.6 der Fall, in der ein Spline durch die folgenden 11 Punkte dargestellt ist.

x	$f(x)$	x	$f(x)$
0	1	± 3	$\frac{1}{10}$
± 1	$\frac{1}{2}$	± 4	$\frac{1}{17}$
± 2	$\frac{1}{5}$	± 5	$\frac{1}{26}$

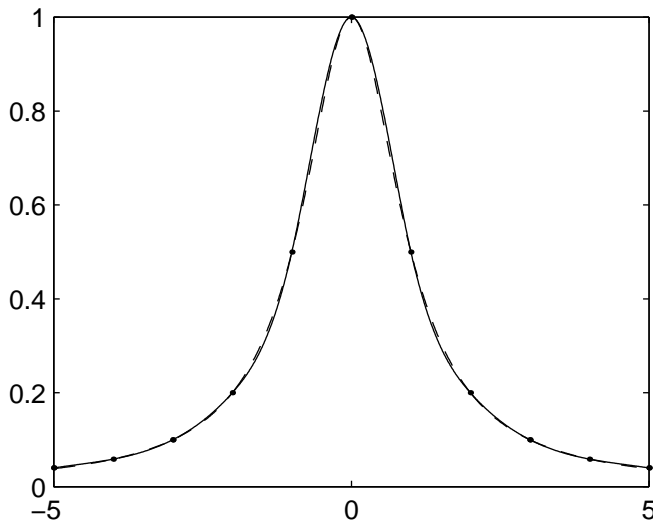


Abbildung 3.6: Auch für größere Stützstellenanzahl ist die Spline-Interpolation geeignet.

An diesem Beispiel erkennt man, dass der Spline auch für mehr Interpolationspunkte nicht zu oszillieren beginnt.

Beispiel 12. Exemplarisch soll an dieser Stelle die Splineberechnung für 5 Stützpunkte gezeigt werden: Ausgegangen wird von den Punkten $(x_0, y_0) = (-5, \frac{1}{26})$, $(x_1, y_1) = (-3, \frac{1}{10})$, $(x_2, y_2) = (0, 1)$, $(x_3, y_3) = (3, \frac{1}{26})$ und $(x_4, y_4) = (5, \frac{1}{26})$, die durch einen natürlichen Spline interpoliert werden sollen. Die zweiten Ableitungen in den Randpunkten sollen also gleich Null sein: $y_0'' = 0$ und $y_4'' = 0$. Aus der Angabe lassen sich die Abstände zwischen den Punkten berechnen:

$$h_0 = x_1 - x_0 = 2, \quad h_1 = x_2 - x_1 = 3, \quad h_2 = x_3 - x_2 = 3, \quad h_3 = x_4 - x_3 = 3.$$

Gemäß (3.20) ergibt sich daher das Gleichungssystem

$$\begin{pmatrix} 10 & 3 & 0 \\ 3 & 12 & 3 \\ 0 & 3 & 10 \end{pmatrix} \cdot \begin{pmatrix} y_1'' \\ y_2'' \\ y_3'' \end{pmatrix} = \begin{pmatrix} \frac{21}{13} \\ -\frac{18}{5} \\ \frac{21}{13} \end{pmatrix},$$

dessen Lösungsvektor

$$y'' = \begin{pmatrix} \frac{327}{1105} \\ -\frac{99}{221} \\ \frac{327}{1105} \end{pmatrix}$$

3 Spline-Interpolation

lautet. Aus den Gleichungen (3.12) bis (3.15),

$$\begin{aligned} a_i &= \frac{1}{6h_i}(y''_{i+1} - y''_i), \\ b_i &= y''_i/2, \\ c_i &= \frac{1}{h_i}(y_{i+1} - y_i) - \frac{h_i}{6}(y''_{i+1} + 2y''_i), \\ d_i &= y_i, \end{aligned}$$

lassen sich nun die Lösungen berechnen:

$$\begin{aligned} s_0(x) &= \frac{327}{1105}(x+5)^3 - \frac{15}{221}(x+5) + \frac{1}{26}, \\ s_1(x) &= -\frac{137}{3315}(x+3)^3 + \frac{327}{1105}(x+3)^2 + \frac{252}{1105}(x+3) + \frac{1}{10}, \\ s_2(x) &= \frac{137}{3315}x^3 - \frac{99}{442}x^2 + 1, \\ s_3(x) &= -\frac{109}{4420}(x-3)^3 + \frac{327}{2210}(x-3)^2 - \frac{252}{1105}(x-3) + \frac{1}{10}. \end{aligned}$$

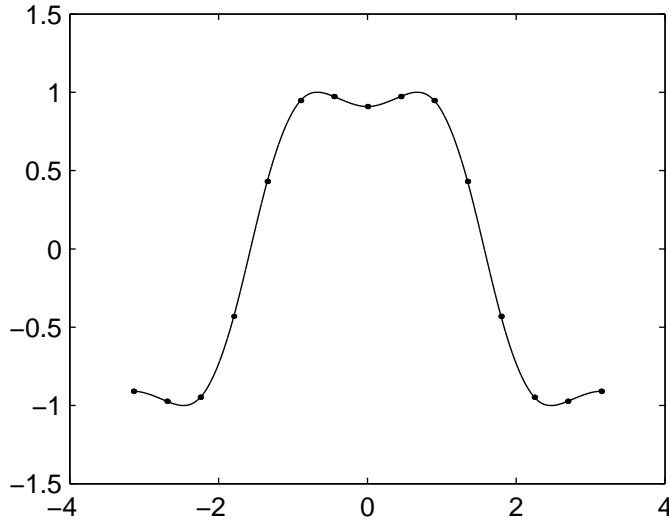


Abbildung 3.7: Periodische Spline-Interpolation von $\sin(2 \cos(x))$ mit gutem Ergebnis.

Beispiel 13. Abschließend möchte ich noch ein Beispiel für eine periodische Spline-Interpolation zeigen. Die Funktion $f(x) = \sin(2 \cos(x))$ wird durch 15 äquidistante Stützstellen interpoliert. Der Grund für äquidistante Stützstellen liegt einzig und allein darin, dass in diesem Fall die Splinelösung besser mit der Lösung der Fourier-Methode (vgl.

3 Spline-Interpolation

Kapitel 6) vergleichbar ist, da diese nur für äquidistante Stützstellen anwendbar ist.

x	$f(x)$	x	$f(x)$
0	0.9093	± 1.9635	-0.6928
± 0.3927	0.9619	± 2.3562	-0.9878
± 0.7854	0.9878	± 2.7489	-0.9619
± 1.1781	0.6928	± 3.1416	-0.9093
± 1.5708	0		

Auch in diesem Fall ist die Approximation so hervorragend, dass der Spline von der eigentlichen Funktion in diesem Maßstab nicht zu unterscheiden ist. Die Funktion ist zum Vergleich in Kapitel 7, Abbildung 7.4 gezeichnet.

4 Bézier-Kurven

Bézier-Polynome sind Linearkombinationen von Bernstein-Polynomen. Diese spezielle Darstellungsweise hat einige Vorteile. Die Kurve ist durch Punkte b_i charakterisiert, die sich anschaulich geometrisch deuten lassen und deren Veränderung sich vor allem in der Umgebung dieser Punkte auswirkt. Mit speziellen Programmen lassen sich somit Kurven konstruieren, die durch Manipulation der Punkte b_i verändert werden können. Im Gegensatz zu den Interpolationspolynomen und Splines sind sie also nachträglich noch veränderbar, ohne dass die Stützstellen verändert werden müssen.

Als Möglichkeiten der Interpolation bzw. Approximation werden Bézier-Polynome und kubische Bézier-Splines, die wie Spline-Kurven segmentweise definiert sind, vorgestellt [12].

4.1 Bernstein-Polynome

Definition 5. Das i -te *Bernstein-Polynom* $B_i^n(\lambda)$ vom Grad n bezüglich des Intervalls $[0, 1]$ ist definiert als

$$B_i^n(\lambda) := \begin{cases} \binom{n}{i} \lambda^i (1 - \lambda)^{n-i}, & n \in \mathbb{N} \text{ und } i = 0, 1, \dots, n. \\ 0, & \text{sonst.} \end{cases} \quad (4.1)$$

Abbildung 4.1 zeigt das Bild von B_i^3 für $i = 0, \dots, 3$.

Bemerkung 10. Es ist üblich, Bernstein-Polynome über dem Einheitsintervall $[0, 1]$ zu verwenden. Da sich jedes Intervall $[a, b]$ durch die Transformation $\lambda = \frac{t-a}{b-a}$ auf das Einheitsintervall abbilden lässt, ist dies keine Einschränkung. Die Definition für ein Bernstein-Polynom über dem Intervall $[a, b]$ wäre

$$B_i^n(t; a, b) := B_i^n\left(\frac{t-a}{b-a}\right) = \frac{1}{(b-a)^n} \binom{n}{i} (t-a)^i (b-t)^{n-i}. \quad (4.2)$$

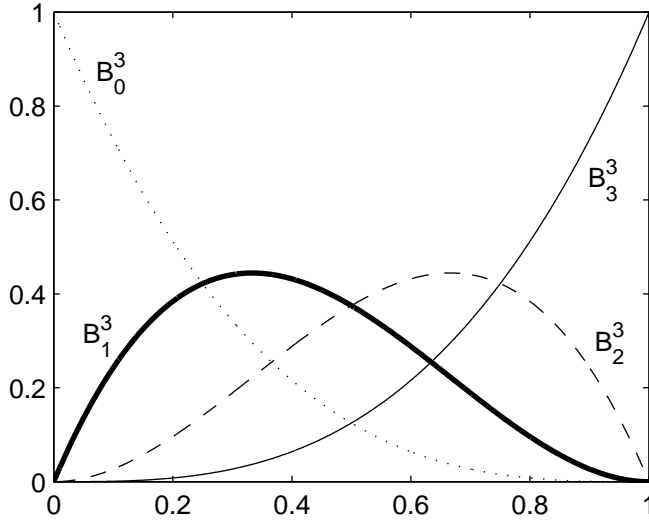


Abbildung 4.1: Die Bernstein-Polynome $B_i^3(\lambda)$ für $\lambda \in [0, 1]$, $i = 0, \dots, 3$.

Bemerkung 11. In der Folge wird der Parameter λ immer für das Intervall $[0, 1]$ und t für ein beliebiges Intervall $[a, b]$ verwendet.

Der nächste Satz gibt eine Übersicht über wichtige Eigenschaften von Bernstein-Polynomen.

Satz 6. *Eigenschaften der Bernstein-Polynome:*

1. $\lambda = 0$ ist i -fache Nullstelle von B_i^n .
2. $\lambda = 1$ ist $(n - i)$ -fache Nullstelle von B_i^n .
3. $B_n^n(\lambda) = \lambda^n$, $B_0^n(\lambda) = (1 - \lambda)^n$, $B_0^0(\lambda) = 1$
4. $(1 - \lambda)B_0^n = B_0^{n+1}$, $\lambda B_n^n = B_{n+1}^{n+1}$.
5. $B_i^n(\lambda) \geq 0$ für $\lambda \in [0, 1]$.
6. $\sum_{i=0}^n B_i^n(\lambda) = 1$ für $\lambda \in \mathbb{R}$.
7. $B_i^n(\lambda) = B_{n-i}^n(1 - \lambda)$, $i = 0, 1, \dots, n$ (Symmetrie der Bernstein-Polynome).
8. $B_i^n(\lambda)$ hat genau ein Maximum in $(0, 1)$ und zwar bei $\lambda = \frac{i}{n}$, $i = 1, \dots, n - 1$.
9. $\frac{\partial}{\partial \lambda} B_i^n(\lambda) = n[B_{i-1}^{n-1}(\lambda) - B_i^{n-1}(\lambda)]$, $i = 0, \dots, n$.

10. Für Bernstein-Polynome gilt die Rekursionsformel

$$B_i^n(\lambda) = \lambda B_{i-1}^{n-1} + (1 - \lambda) B_i^{n-1}(\lambda),$$

$$i = 1, \dots, n.$$

11. Die $n + 1$ Bernstein-Polynome B_i^n für $i = 0, 1, \dots, n$ sind linear unabhängig.

Beweis. Die Aussagen 1 bis 5 ergeben sich direkt aus der Definition (4.1).

ad 6. Wegen des Binomischen Lehrsatzes gilt

$$1 = \left((1 - \lambda) + \lambda \right)^n = \sum_{i=0}^n \binom{n}{i} (1 - \lambda)^{n-i} \lambda^i = \sum_{i=0}^n B_i^n(\lambda).$$

ad 7.

$$B_{n-i}^n(1 - \lambda) = \underbrace{\binom{n}{n-i}}_{=\binom{n}{i}} \underbrace{(1 - (1 - \lambda))^{n-(n-i)}}_{=\lambda^i} (1 - \lambda)^{n-i} = B_i^n(\lambda).$$

ad 8. Die erste Ableitung von $B_i^n(\lambda)$ lautet

$$\begin{aligned} \frac{\partial}{\partial y} B_i^n(\lambda) &= \binom{n}{i} \left[-(n-i)(1-\lambda)^{n-i-1} \lambda^i + i(1-\lambda)^{n-i} \lambda^{i-1} \right] = \\ &= \binom{n}{i} (1-\lambda)^{n-i-1} \lambda^{i-1} \left[-(n-i)\lambda + i(1-\lambda) \right] = \\ &= \binom{n}{i} (1-\lambda)^{n-i-1} \lambda^{i-1} (i - n\lambda). \end{aligned}$$

Gesucht ist die Menge aller $\lambda \in (0, 1)$, für die gilt $\frac{\partial}{\partial y} B_i^n(\lambda) = 0$. Offensichtlich erfüllt nur $\lambda = \frac{i}{n}$ diese Bedingung.

ad 9. Wir leiten das Bernstein-Polynom $B_i^n(\lambda)$ ab und erhalten

$$\begin{aligned} \frac{\partial}{\partial y} B_i^n(\lambda) &= -(n-i) \binom{n}{i} (1-\lambda)^{n-i-1} \lambda^i + i \binom{n}{i} \lambda^{i-1} (1-\lambda)^{n-i} = \\ &= -\frac{(n-i)n!}{i!(n-i)!} \lambda^i (1-\lambda)^{n-i-1} + \frac{i \cdot n!}{i!(n-i)!} \lambda^{i-1} (1-\lambda)^{n-i} = \\ &= -n \frac{(n-1)!}{i!(n-1-i)!} \lambda^i (1-\lambda)^{n-i-1} + \frac{n \cdot (n-1)!}{(i-1)!(n-i)!} \lambda^{i-1} (1-\lambda)^{n-i} = \\ &= -n \binom{n-1}{i} \lambda^i (1-\lambda)^{n-i-1} + n \binom{n-1}{i-1} \lambda^{i-1} (1-\lambda)^{n-i} = \\ &= n [B_{i-1}^{n-1}(\lambda) - B_i^{n-1}(\lambda)]. \end{aligned}$$

ad 10. Es gilt

$$\binom{n-1}{i-1} + \binom{n-1}{i} = \frac{(n-1)!}{(i-1)!(n-i)!} + \frac{(n-1)!}{i!(n-i-1)!} = \binom{n}{i},$$

und somit folgt die Rekursionsformel

$$\begin{aligned} B_i^n(\lambda) &= \binom{n}{i} (1-\lambda)^{n-i} \lambda^i = \binom{n-1}{i-1} (1-\lambda)^{n-i} \lambda^i + \binom{n-1}{i} (1-\lambda)^{n-i} \lambda^i = \\ &= \lambda \binom{n-1}{i-1} (1-\lambda)^{(n-1)-(i-1)} \lambda^{i-1} + (1-\lambda) \binom{n-1}{i} (1-\lambda)^{n-1-i} \lambda^i = \\ &= \lambda B_{i-1}^{n-1}(\lambda) + (1-\lambda) B_i^{n-1}(\lambda). \end{aligned}$$

ad 11. Um die lineare Unabhängigkeit der $n+1$ Bernstein- Polynome zu zeigen, ist nachzuweisen, dass aus

$$\sum_{i=0}^n c_i B_i^n(\lambda) = 0, \quad c_i \in \mathbb{R}, \lambda \in \mathbb{R},$$

$c_i = 0$ für alle i folgt. Setzen wir $\lambda = 1$, so folgt aus der Definition, dass $B_i^n(1) = 0$ für $i = 0, \dots, n-1$ und $B_n^n(1) = 1$. Somit gilt $c_n = 0$. Nun muss aber auch für die Ableitung an der Stelle 1 die Gleichung $\left. \frac{\partial}{\partial \lambda} \sum_{i=0}^n c_i B_i^n(\lambda) \right|_{\lambda=1} = 0$ gelten.

$$\begin{aligned} 0 &= \left. \frac{\partial}{\partial \lambda} \sum_{i=0}^n c_i B_i^n(\lambda) \right|_{\lambda=1} = \sum_{i=0}^n c_i \left. \frac{\partial}{\partial \lambda} B_i^n(\lambda) \right|_{\lambda=1} = \\ &= 0 + \dots + c_{n-1} \underbrace{\binom{n}{n-1} \lambda^n (1-\lambda)^0 (n-1-\lambda n)}_{\neq 0 \text{ für } \lambda=1}. \end{aligned}$$

Daher folgt $c_{n-1} = 0$. Für die höheren Ableitungen argumentieren wir analog bis wir $c_i = 0$ für alle $i = 0, \dots, n$ erhalten.

□

4.2 Bézier-Polynome

Aus der linearen Unabhängigkeit der Bernstein-Polynome B_i^n folgt, dass sich Polynome P vom Grad n eindeutig als

$$P(\lambda) = \sum_{i=0}^n b_i B_i^n(\lambda) \tag{4.3}$$

darstellen lassen. Ebene Kurven können somit durch die sogenannte *Bézier-Darstellung*

$$x(\lambda) = \begin{pmatrix} x_1(\lambda) \\ x_2(\lambda) \end{pmatrix} := \sum_{i=0}^n \begin{pmatrix} b_{i1} \\ b_{i2} \end{pmatrix} B_i^n(\lambda) = \sum_{i=0}^n b_i B_i^n(\lambda) \text{ für } \lambda \in [0, 1] \quad (4.4)$$

beziehungsweise

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} := \sum_{i=0}^n \begin{pmatrix} b_{i1} \\ b_{i2} \end{pmatrix} B_i^n(t; a, b) = \sum_{i=0}^n b_i B_i^n(t; a, b) \text{ für } t \in [a, b] \quad (4.5)$$

approximiert werden. Die Punkte b_i nennt man *Bézier-Punkte*.

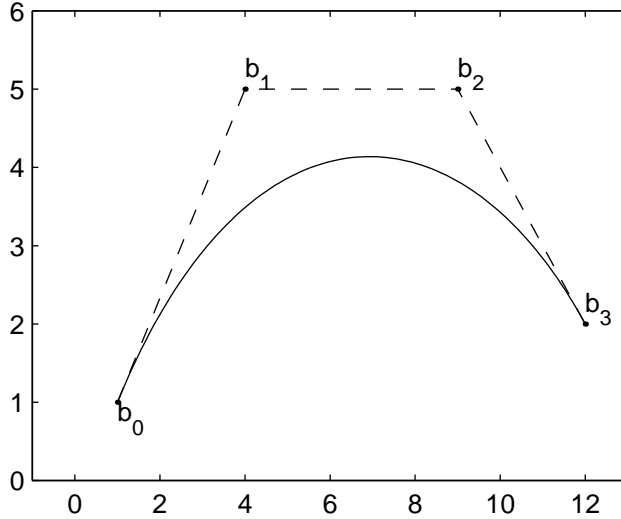


Abbildung 4.2: Kubisches Bézier-Polynom mit den Bézier-Punkten b_i und der zugehörigen linearen Hülle.

Eigenschaften der Bézier-Punkte

Bézier-Punkte lassen sich geometrisch einfach deuten. Verbindet man aufeinanderfolgende Punkte b_i und b_{i+1} , so erhält man das *Bézier-Polygon*. Dies ist in Abbildung 4.2 anhand eines kubischen Bézier-Polynoms verdeutlicht.

Satz 7. *Der Graph des Bézier-Polynoms verläuft innerhalb der konvexen Hülle der Bézier-Punkte bzw. des Bézier-Polygons.*

Beweis. Eine Menge M heißt konvex, wenn für alle $u, v \in M$ gilt, dass auch alle Punkte ihrer Verbindungsstrecke Teil der Menge sind, also wenn

$$\{(1 - \mu)u + \mu v \mid \mu \in [0, 1]\} \subset M \quad (4.6)$$

gilt.

4 Bézier-Kurven

Eine konvexe Hülle ist gegeben durch

$$\left\{ u = \sum_{i=0}^m \mu_i u_i \mid m \in \mathbb{N}, u_i \in M, \mu_i \geq 0, \sum_{i=0}^m \mu_i = 1 \right\}. \quad (4.7)$$

Ersetzt man nun μ_i durch $B_i^n(\lambda)$, das die Eigenschaften $B_i^n(\lambda) \geq 0$ und $\sum_{i=0}^n B_i^n(\mu) = 1$ hat, so ist der Satz bewiesen. □

Satz 8. Für die Randpunkte $\lambda = 0$ und $\lambda = 1$ einer Bézier-Kurve $x(\lambda) = \sum_{i=0}^n b_i B_i^n(\lambda)$ gilt

$$\begin{aligned} x(0) &= b_0, & x(1) &= b_n \\ \dot{x}(0) &= n(b_1 - b_0), & \dot{x}(1) &= n(b_n - b_{n-1}) \\ \ddot{x}(0) &= n(n-1)(b_2 - 2b_1 + b_0), & \ddot{x}(1) &= n(n-1)(b_n - 2b_{n-1} + b_{n-2}). \end{aligned} \quad (4.8)$$

Beweis. Die ersten beiden Aussagen folgen durch Einsetzen in die Definition. Die erste Ableitung lautet nach Satz 5

$$\begin{aligned} \dot{x}(\lambda) &= n \sum_{i=0}^n b_i (B_{i-1}^{n-1}(\lambda) - B_i^{n-1}(\lambda)) = \\ &= -nb_0 B_0^{n-1}(\lambda) + n \sum_{i=1}^{n-1} [b_i (B_{i-1}^{n-1}(\lambda) - B_i^{n-1}(\lambda))] + nb_n B_{n-1}^{n-1}(\lambda). \end{aligned}$$

Durch Auflösen des Summenzeichens und Umgruppieren erhält man

$$\begin{aligned} \dot{x}(\lambda) &= \dots = n \left(\underbrace{-b_0 B_0^{n-1}(\lambda) + b_1 B_0^{n-1}(\lambda)}_{(b_1 - b_0) B_0^{n-1}(\lambda)} \underbrace{-b_1 B_1^{n-1}(\lambda) + b_2 B_1^{n-1}(\lambda)}_{(b_2 - b_1) B_1^{n-1}(\lambda)} - b_2 B_1^{n-1}(\lambda) + \dots \right. \\ &\quad \left. \dots \underbrace{-b_{n-1} B_{n-1}^{n-1}(\lambda) + b_n B_{n-1}^{n-1}(\lambda)}_{(b_n - b_{n-1}) B_{n-1}^{n-1}(\lambda)} \right) = \\ &= n \sum_{i=0}^{n-1} (b_{i+1} - b_i) B_i^{n-1}(\lambda). \end{aligned}$$

Differenziert man $\dot{x}(\lambda)$ ein zweites Mal und fasst analog alle $B_i^{n-2}(\lambda)$ zusammen, so erhält man

$$\ddot{x}(\lambda) = n(n-1) \sum_{i=0}^{n-2} (b_{i+2} - 2b_{i+1} + b_i) B_i^{n-2}(\lambda).$$

Der Satz ist durch Einsetzen von $\lambda = 0$ bzw. $\lambda = 1$ gezeigt. □

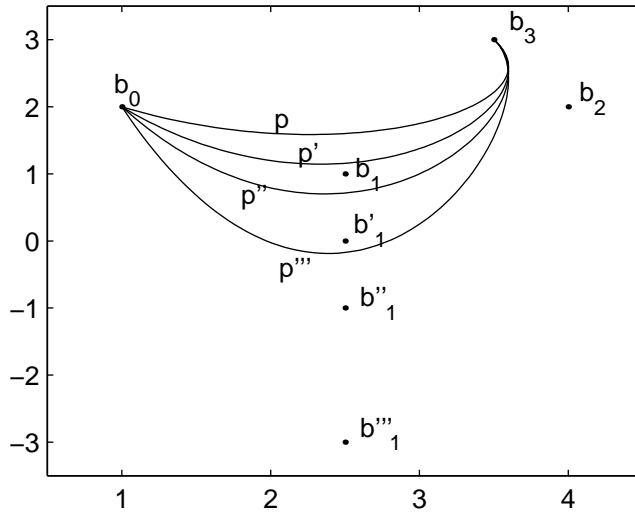


Abbildung 4.3: Kubische Bézier-Polynome mit unterschiedlichem b_1 .

Bemerkung 12. Im Folgenden schreiben wir \dot{x} für die Ableitung der Kurve x , weil wir später noch einen oberen Index einführen und x'^j weniger übersichtlich ist als \dot{x}^j .

Bézier-Polynome sind durch ihre Bézier-Punkte eindeutig gegeben. Im Gegensatz zu Koeffizienten von Interpolationspolynomen haben sie aber eine geometrische Bedeutung. Deshalb kann eine Bézier-Kurve durch Veränderung eines Bézier-Punktes in einer gewünschten Art verändert werden. Möchte man ein Interpolationspolynom durch Veränderung der Koeffizienten dazu bewegen, eine bestimmte Form anzunehmen, so ließe sich dies nur durch die Methode „Versuch und Irrtum“ bewerkstelligen, was für höhergradige Polynome absolut aussichtslos ist.

Beispiel 14. Gehen wir von den Bézier-Punkten $b_0 = (1, 2)$, $b_1 = (2.5, 1)$, $b_2 = (4, 2)$ und $b_3 = (3.5, 3.5)$ aus. Wie Abbildung 4.3 zeigt, bewirkt eine Veränderung eines Bézier-Punktes, eine - mit etwas Erfahrung - vorhersehbare Veränderung der Kurve. In unserem Beispiel wird b_1 „nach unten gezogen“. $b'_1 = (2.5, 0)$, $b''_1 = (2.5, -1)$ und $b'''_1 = (2.5, -3)$ stellen verschiedene Werte dar, die durch die Kurven p' bis p''' realisiert sind. Die immer kleiner werdenden Ordinaten-Werte des Bézier-Punktes b_1 schlagen sich anschaulich in der Kurve nieder. Diese wird ebenso „nach unten gezogen“ - allerdings wie ein Gummiband - denn je mehr wir uns den anderen Bézier-Punkten nähern, desto weniger ausgeprägt ist dieser Effekt.

Anfangs- und Endpunkt eines Bézier-Polynoms sind eindeutig bestimmt. Die ersten Ableitung in den Randpunkten ist wegen Satz 8 auch sofort ersichtlich. Nach Satz 8 ist die Richtung der erste Ableitung in den Randpunkten gleich der Verbindung der Punkte b_0 und b_1 bzw. b_n und b_{n-1} , was sich optisch leicht umsetzen lässt. Das bedeutet, dass sich

mit Hilfe eines Computerprogrammes die Möglichkeit eröffnet, zuerst die Bézier-Punkte per Hand zu wählen und anschließend die Punkte solange zu verändern, bis die Kurve in der gewünschten Form erscheint. Für eine schnelle Auswertung des Bézier-Polynoms folgt der nächste Algorithmus.

Verfahren von de Casteljau

Dieses Verfahren dient der Berechnung von $x(\lambda)$ für $\lambda \in [0, 1]$ aus der Bézier-Darstellung $x(\lambda) = \sum_{i=0}^n b_i B_i^n(\lambda)$ bei gegebenen Bézier-Punkten b_i .

Satz 9. Definieren wir das Bézier-Polynom $b_{r,\dots,s}(\lambda)$ wie folgt

$$b_{r,\dots,s}(\lambda) := \sum_{i=r}^s b_i B_{i-r}^{s-r}(\lambda), \quad (4.9)$$

dann gilt an der Stelle λ für das Bézier-Polynom $b_{r,\dots,s}(\lambda)$ die Rekursionsformel

$$b_{r,\dots,s}(\lambda) = (1 - \lambda)b_{r,\dots,s-1}(\lambda) + \lambda b_{r+1,\dots,s}(\lambda). \quad (4.10)$$

Beweis. Durch Einsetzen der Definition (4.9) in die Rekursionsformel (4.10) erhält man die Gleichung

$$\sum_{i=r}^s b_i B_{i-r}^{s-r}(\lambda) = (1 - \lambda) \sum_{i=r}^{s-1} b_i B_{i-r}^{s-r-1}(\lambda) + \lambda \sum_{i=r+1}^s b_i B_{i-r-1}^{s-r-1}(\lambda), \quad (4.11)$$

die nun zu beweisen ist. Wegen Eigenschaft 4 von Satz 6 gilt,

$$\begin{aligned} \text{für } i = r : \quad & B_0^{s-r}(\lambda) = (1 - \lambda)B_0^{s-r-1}(\lambda), \\ \text{für } i = s : \quad & B_{s-r}^{s-r}(\lambda) = \lambda B_{s-r-1}^{s-r-1}(\lambda), \end{aligned}$$

und wegen Eigenschaft 10 von Satz 6 folgt

$$\begin{aligned} \text{für } i = r + k, \quad k = 1, \dots, s - r - 1 : \\ B_{i-r}^{s-r} = (1 - \lambda)B_{i-r}^{s-r-1}(\lambda) + \lambda B_{i-r-1}^{s-r-1}(\lambda). \end{aligned}$$

Damit ist (4.11) gezeigt. □

Bemerkung 13. Aus der Definition der $b_{r,\dots,s}$ folgt

1. $b_{0,\dots,n}(\lambda) = x(\lambda)$,
2. $b_{i,i}(\lambda) = b_i$.

Aus diesem Satz folgt die Berechnung des Wertes $x(\lambda)$ nach folgendem Schema:

$$\begin{array}{ccccccc}
 b_0 & & & & & & \\
 b_1 & b_{0,1}(\lambda) & & & & & \\
 b_2 & b_{1,2}(\lambda) & b_{0,1,2}(\lambda) & & & & \\
 b_3 & b_{2,3}(\lambda) & b_{1,2,3}(\lambda) & b_{0,1,2,3}(\lambda) & & & \\
 \vdots & \vdots & \vdots & & \ddots & & \\
 b_n & b_{n-1,n}(\lambda) & b_{n-2,n-1,n}(\lambda) & \cdots & & b_{0,\dots,n}(\lambda) = x(\lambda)
 \end{array}$$

Da wir ein gegebenes Bézier-Polynom an der ebenfalls gegebenen Stelle λ auswerten wollen, sind die Bézier-Punkte b_i , $i = 0, \dots, n$ und somit die erste Spalte dieses Schemas gegeben. Die Elemente der zweiten Spalte lassen sich mit Hilfe des vorhergehenden Satzes berechnen: $b_{0,1} = (1-\lambda)b_0 + \lambda b_1$, $b_{1,2} = (1-\lambda)b_1 + \lambda b_2$, \dots . Auf die gleiche Art berechnet man die restlichen Spalten, bis man bei $b_{0,\dots,n}$ angelangt ist. Dieser Wert entspricht dem gesuchten Wert $x(\lambda)$.

Beispiel und Bild

Für Funktionen, die über einem Definitionsbereich nicht zu „unruhig“ sind, stellen Bézier-Polynome eine gute Möglichkeit der Approximation dar. „Unruhig“ bedeutet dabei, dass nicht viele Extremwerte oder Wendepunkte auftreten. Dann ist es, wie das nächste Beispiel zeigt, auch ohne langes „Herumprobieren“ möglich, ein interpolierendes Bézier-Polynom zu erhalten, sofern man sich geeignete Bedingungen für die Bézier-Punkte überlegt.

Beispiel 15. Es soll der Viertelkreis mit Radius $r = 1$ durch ein kubisches Bézier-Polynom interpoliert werden. Der Anfangspunkt ist $(0, 1)$ und der Endpunkte $(1, 0)$. Damit sind die beiden Bézier-Punkte $b_0 = (0, 1)$ und $b_3 = (1, 0)$ bereits gegeben. Eine sinnvolle Bedingung wäre, dass auch die Tangenten der Bézier-Kurve mit der des Kreises übereinstimmen. b_1 und b_2 sollen also auf der Tangente der Kurve im Punkt b_0 bzw. b_3 liegen. Wenn nun die Tangenten in b_0 und b_3 des Viertelkreises den Tangenten der Bézier-Kurve entsprechen sollen, so sind diese wegen der Symmetrie bis auf einen Parameter ξ gegeben: $b_1 = (\xi, 1)$ und $b_2 = (1, \xi)$. Somit lautet der Ansatz für unser Bézier-Polynom

$$x(\lambda) = \sum_{i=0}^3 b_i B_i^n = \binom{1}{0} (1-\lambda)^3 + 3 \binom{1}{\xi} \lambda (1-\lambda)^2 + 3 \binom{\xi}{1} \lambda^2 (1-\lambda) + \binom{0}{1} \lambda^3. \quad (4.12)$$

ξ lässt sich zum Beispiel bestimmen, indem wir fordern, dass der Punkt $x(0.5)$ auf dem Kreis liegt.

$$x(0.5) = \begin{pmatrix} \cos(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) \end{pmatrix} = \begin{pmatrix} 0.5^3 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0.5^3 \xi \\ 0.5^3 \end{pmatrix} + 3 \begin{pmatrix} 0.5^3 \\ 0.5^3 \xi \end{pmatrix} + \begin{pmatrix} 0 \\ 0.5^3 \end{pmatrix}$$

$$\Rightarrow \xi = 0.5523$$

Die so gewonnene Kurve stellt eine gute Näherung des Viertelkreises dar. In Figur 4.4 ist die Bézier-Kurve dem Kreis (gestrichelte Kurve mit Radius 0.9) gegenübergestellt.

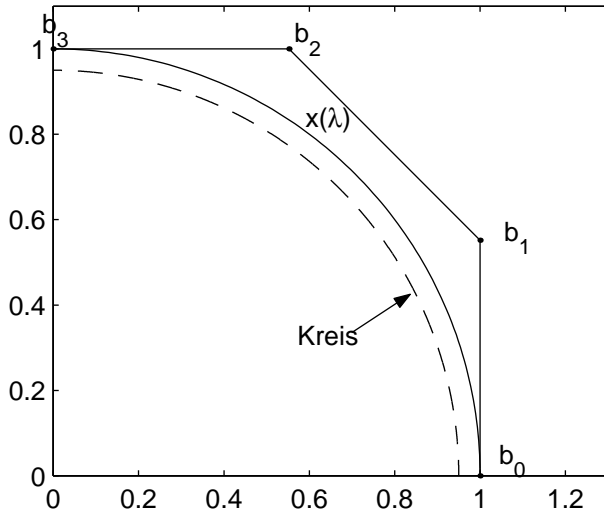


Abbildung 4.4: Bézier-Polynomapproximation für den Viertelkreis.

Bemerkung 14. Wie bereits erwähnt, stellen Bézier-Polynome eine Möglichkeit der Approximation dar. Man definiert b_0 := Anfangspunkt und b_n := Endpunkt und bestimmt durch Vorgeben (und späteres Verändern) der Punkte $b_i, i = 1, \dots, n-1$ den Verlauf der Kurve, wobei der Grad n nach Belieben gewählt werden kann, je nachdem ob man mehr oder weniger Punkte zum Verändern der Funktion haben möchte. Dies ist aber sehr umständlich. Deshalb werden in der Praxis für eine große Anzahl von Interpolationsstellen Bézier-Splines verwendet. Wie bei den kubischen Polynom-Splines in Kapitel 3 wird der Definitionsbereich in Segmente zerlegt, für die jeweils Bézier-Polynome berechnet werden, wobei an den Anschlussstellen wieder bestimmte Bedingungen zu erfüllen sind, die dann den Bézier-Spline definieren.

4.3 Bézier-Splines

Wie bei den Splines erscheint es vorteilhaft, das Intervall in Segmente zu unterteilen und für jedes Intervall ein Bézier-Polynom zu berechnen. Daher nennen wir diese

Kurven *Bézier-Splines*. An den Schnittstellen wollen wir wieder bestimmte Bedingungen verlangen, um Stetigkeit und zweimalige stetige Differenzierbarkeit zu erreichen. In der Praxis haben sich kubische Bézier-Polynome bewährt. Diese werden hier erklärt.

Gegeben seien $m + 1$ Stützpunkte (x_j, y_j) , die durch einen kubischen Bézier-Spline interpoliert werden sollen. Das heißt, dass wir m Bézier-Polynome benötigen, wobei der Ansatz des j -ten

$$x^j(\lambda) = \sum_{i=0}^3 b_i^j B_i^3(\lambda), \quad j = 1, \dots, m \quad (4.13)$$

lautet.

Bemerkung 15. Wir verwenden hier den Index oben, um eine Verwechslung mit der x- oder y-Koordinate zu vermeiden.

Wegen der Interpolationsbedingung muss

$$\begin{aligned} (x_0, y_0) &= b_0^1 & \text{und} \\ (x_j, y_j) &= b_3^j & \text{für alle } j = 1, \dots, m \end{aligned} \quad (4.14)$$

gelten.

Die Stetigkeit des Bézier-Splines ist erfüllt, wenn für alle $j = 1, \dots, m - 1$ gilt

$$b_3^j = b_0^{j+1}. \quad (4.15)$$

Soll nun der Bézier-Spline einmal stetig differenzierbar sein, so muss nach Satz 8 für alle $j = 1, \dots, m - 1$ gelten, dass

$$\dot{x}^j(1) = 3(b_3^j - b_2^j) = \dot{x}^{j+1}(0) = 3(b_1^{j+1} - \underbrace{b_0^{j+1}}_{=b_3^j}), \quad (4.16)$$

woraus folgt, dass

$$2b_3^j = b_1^{j+1} + b_2^j \quad (4.17)$$

gelten muss. Geometrisch interpretiert heisst diese Bedingung nichts anderes, als dass b_3^j die Strecke $\overline{b_1^{j+1} b_2^j}$ halbiert.

Soll darüber hinaus auch noch die zweite Ableitung in den Verbindungsstellen stetig sein, so muss wieder nach Satz 8 gelten

$$\begin{aligned} \ddot{x}^j(1) &= \ddot{x}^{j+1}(0) \\ 6(b_3^j - 2b_2^j + b_1^j) &= 6(b_2^{j+1} - 2b_1^{j+1} + b_0^{j+1}), \end{aligned} \quad (4.18)$$

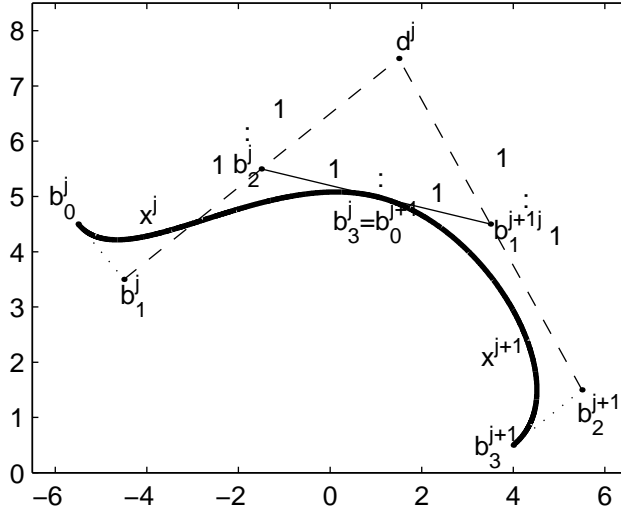


Abbildung 4.5: Geometrische Deutung der Anschlussbedingungen.

was sich mit Hilfe des Hilfspunktes d^j zu

$$2b_2^j - b_1^j = 2b_1^{j+1} - b_2^{j+1} =: d^j \quad (4.19)$$

umformen läßt. Die geometrische Deutung zeigt anschaulich, dass b_2^j bzw. b_1^{j+1} der Halbierungspunkt der Strecke $b_1^j d^j$ bzw. $b_2^{j+1} d^{j+1}$ ist. Die Abbildung 4.5 liefert dazu die Skizze, in der durch '1:1' das Teilungsverhältnis der jeweiligen Strecken angedeutet ist.

Aus (4.19) folgen die Bedingungen

$$\begin{aligned} 2d^j &= 4b_2^j - 2b_1^j, & 2d^j &= 4b_1^{j+1} - 2b_2^{j+1}, \\ d^{j-1} &= 2b_1^j - b_2^j, & d^{j+1} &= 2b_2^{j+1} - b_1^{j+1}. \end{aligned} \quad (4.20)$$

Werden die übereinanderstehenden Gleichungen jeweils addiert, so erhält man

$$2d^j + d^{j-1} = 3b_2^j, \quad 2d^j + d^{j+1} = 3b_1^{j+1}, \quad (4.21)$$

für die wegen (4.17) für $j = 1, \dots, m-1$

$$d^{j-1} + 4d^j + d^{j+1} = 6b_3^j \quad (4.22)$$

gilt. Aus (4.21) folgen die Bedingungen

$$\begin{aligned} 2d^0 + d^1 &= 3b_1^1, \\ d^{m-1} + 2d^m &= 3b_2^m. \end{aligned} \quad (4.23)$$

Somit ist das zu lösende Gleichungssystem gegeben durch

$$\begin{aligned} 2d^0 + d^1 &= 3b_1^1 \\ d^{j-1} + 4d^j + d^{j+1} &= 6b_3^j \quad \text{für } j = 1, \dots, m-1 \\ d^{m-1} + 2d^m &= 3b_2^m. \end{aligned} \quad (4.24)$$

Analog zu den Polynom-Splines sind wieder 2 Bedingungen zu wenig, um das Gleichungssystem zu lösen, da b_1^1 und b_2^m nicht gegeben sind. Je nachdem wie man damit umgeht, gibt es wieder verschiedene Arten von Bézier-Splines.

Bemerkung 16. Offensichtlich ist das Gleichungssystem eindeutig lösbar, da die zugehörige Matrix diagonal dominant ist und somit maximalen Rang besitzt (vgl. Kapitel 3 Splines, Definition 4 und Satz 4).

4.3.1 Arten von Bézier-Splines

Vorgegebene 1. Randableitungen

Gibt man sich die Ableitungen in den Punkten (x_0, y_0) und (x_m, y_m) vor, so erhält man wegen Satz 8 $\dot{x}^1(0) = 3(b_1^1 - b_0^1)$ und $\dot{x}^m(1) = 3(b_3^m - b_2^m)$. Dadurch lassen sich b_1^1 und b_2^m berechnen und in das Gleichungssystem (4.24) einsetzen.

Natürliche kubische Bézier-Splines

Soll die zweite Ableitung in den Randpunkten gleich Null sein, so ist $d^0 = b_0^1$ und $d^m = b_3^m$ zu wählen, denn wegen (4.17) gilt $d^0 = 2b_1^1 - b_2^1$ und nach Satz 8 folgt dann

$$\ddot{x}^1(0) = 3 \cdot 2(b_0^1 - 2b_1^1 + b_2^1) = 6(\underbrace{d^0 - 2b_1^1 + b_2^1}_{=d^0}) = 0. \quad (4.25)$$

Analog gilt genauso

$$\ddot{x}^m(1) = 3 \cdot 2(b_3^m - 2b_2^m + b_1^m) = 6(\underbrace{d^m - 2b_2^m + b_1^m}_{=d^m}) = 0. \quad (4.26)$$

Periodische Bézier-Splines

Die Bedingungen

$$x^1(0) = x^m(1), \quad \dot{x}^1(0) = \dot{x}^m(1), \quad \ddot{x}^1(0) = \ddot{x}^m(1) \quad (4.27)$$

definieren einen periodischen Bézier-Spline.

Periodische Spline sind auch im Anfangs- bzw. Endpunkte $b_0^1 = b_3^m$ 2 mal stetig differenzierbar ist. Daher gelten auch in diesem Punkt die selben Gleichungen (4.22) wie bisher nur für die inneren. Das Gleichungssystem lautet deshalb

$$\begin{pmatrix} 4 & 1 & & & & 1 \\ 1 & 4 & 1 & & & \\ & 1 & 4 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 4 & 1 \\ 1 & & & & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} d^0 \\ d^1 \\ d^2 \\ \vdots \\ d^{m-2} \\ d^{m-1} \end{pmatrix} = 6 \cdot \begin{pmatrix} b_3^m \\ b_3^1 \\ b_3^2 \\ \vdots \\ b_3^{m-2} \\ b_3^{m-1} \end{pmatrix} \quad (4.28)$$

Offensichtlich gilt $d^0 = d^m$ und $b_3^0 = b_3^m$.

4.3.2 Beispiel und Bild

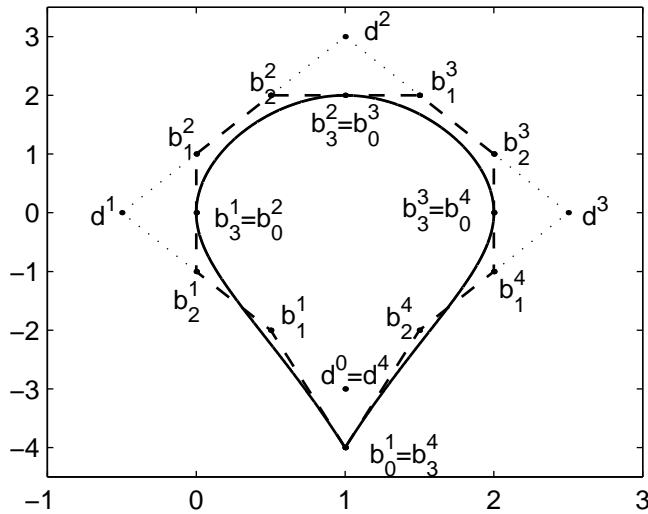


Abbildung 4.6: Ein 2 mal stetig differenzierbarer Bézier-Spline.

Beispiel 16. Wir wollen die Möglichkeit, geschlossene Kurven zu interpolieren, ausnützen und approximieren eine Kurve, die durch die Punkte $(1, -4), (0, 0), (1, 2), (2, 0)$ und $(1, -4)$ gegeben ist. Im Anfangs- bzw. Endpunkt geben wir uns die Ableitungen bzw. die Punkte $b_1^1 = \begin{pmatrix} 0.5 \\ -2 \end{pmatrix}$ und $b_2^4 = \begin{pmatrix} 1.5 \\ -2 \end{pmatrix}$ vor, was wegen Satz 8 quasi dem Vorgeben der Ableitungen in den Anfangs- und Endpunkten entspricht. So erhalten wir einen Bézier-Spline, der an allen anderen Knoten als in $(1, -4)$ 2 mal stetig differenzierbar ist.

4 Bézier-Kurven

Zur Berechnung gehen wir wie folgt vor: Aus den Bedingungen (4.24) erhalten wir das Gleichungssystem

$$\begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} d^0 \\ d^1 \\ d^2 \\ d^3 \\ d^4 \end{pmatrix} = \begin{pmatrix} 3b_1^1 \\ 6b_3^1 \\ 6b_3^2 \\ 6b_3^3 \\ 3b_2^4 \end{pmatrix},$$

in das wir nun die 3 gegebenen und die 2 selbst gewählten Bézier-Punkte einsetzen können. Wir erhalten also 2 Gleichungssysteme, jeweils eines für die x-Koordinate und eines für die y-Koordinate. Um nicht durch noch mehr numerische Indizes die Rechnung unnötig kompliziert zu machen, schreiben wir die Koordinaten des Vektors wie folgt: $d^j = \begin{pmatrix} d_x^j \\ d_y^j \end{pmatrix}$. Die durch Einsetzen erhaltenen Gleichungssysteme lauten

$$\begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix} \begin{pmatrix} d_x^0 \\ d_x^1 \\ d_x^2 \\ d_x^3 \\ d_x^4 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 0 \\ 1 \\ 2 \\ 4.5 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix} \begin{pmatrix} d_y^0 \\ d_y^1 \\ d_y^2 \\ d_y^3 \\ d_y^4 \end{pmatrix} = \begin{pmatrix} -6 \\ 0 \\ 2 \\ 0 \\ -6 \end{pmatrix}.$$

Daraus ergeben sich für die Hilfspunkte d^j die Werte

$$d^0 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \quad d^1 = \begin{pmatrix} -\frac{1}{2} \\ 0 \end{pmatrix}, \quad d^2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad d^3 = \begin{pmatrix} \frac{5}{2} \\ 0 \end{pmatrix}, \quad d^4 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}.$$

Mittels der beiden in (4.21) gegebenen Formeln $2d^j + d^{j-1} = 3b_2^j$ und $2d^j + d^{j+1} = 3b_1^{j+1}$ lassen sich die gesuchten Bézier-Punkte berechnen:

$$b_1^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad b_1^3 = \begin{pmatrix} \frac{3}{2} \\ 2 \end{pmatrix}, \quad b_1^4 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad b_2^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad b_2^2 = \begin{pmatrix} \frac{1}{2} \\ 2 \end{pmatrix}, \quad b_2^3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Die Bézier-Punkte setzen wir für $j = 1, \dots, 4$ in

$$x^j = \sum_{i=0}^3 b_i^j B_i^3(\lambda)$$

ein und erhalten

$$\begin{aligned}
 x^1(\lambda) &= \binom{1}{-4} (1-\lambda)^3 + 3 \binom{0.5}{-2} x(1-\lambda)^2 + 3 \binom{0}{1} x^2(1-\lambda) + \binom{0}{0} x^3, \\
 x^2(\lambda) &= \binom{0}{0} (1-\lambda)^3 + 3 \binom{0}{1} x(1-\lambda)^2 + 3 \binom{0.5}{2} x^2(1-\lambda) + \binom{2}{1} x^3, \\
 x^3(\lambda) &= \binom{2}{1} (1-\lambda)^3 + 3 \binom{1.5}{2} x(1-\lambda)^2 + 3 \binom{2}{1} x^2(1-\lambda) + \binom{2}{0} x^3, \\
 x^4(\lambda) &= \binom{2}{0} (1-\lambda)^3 + 3 \binom{2}{-1} x(1-\lambda)^2 + 3 \binom{1.5}{-2} x^2(1-\lambda) + \binom{1}{-4} x^3.
 \end{aligned}$$

5 Methode der kleinsten Quadrate

Wir gehen diesmal von n gegebenen Punkten (x_i, y_i) , $i = 1, \dots, n$ aus, die eine Funktion f beschreiben sollen. Des weiteren soll die Beziehung zwischen den x - und y -Werten bis auf noch zu bestimmende Parameter p_1, \dots, p_m bekannt sein. Die Ansatzfunktion soll von der Form $f(x) = p_m \varphi_m(x) + p_{m-1} \varphi_{m-1}(x) + \dots + p_2 \varphi_2(x) + p_1 \varphi_1(x)$ sein, wobei die Funktionen $\varphi_1, \dots, \varphi_m$ linear unabhängig sein müssen. Im Folgenden gehen wir von der Annahme $\varphi_k(x) = x^{k-1}$ aus. Diese unorthodoxe Schreibweise eines Polynoms, bei der der Index der Koeffizienten nicht mit dem der Potenzen übereinstimmt, wurde deshalb gewählt, um die später folgende Matrizenschreibweise zu vereinfachen. Genausogut hätte man einen üblichen Polynomansatz verwenden können; dafür wären dann aber die Indizes unangenehmer umzuformen gewesen.

Zum Unterschied zur Polynominterpolation sei die Zahl der gegebenen Punkte n aber größer als der Grad $m-1$ des Polynoms. Geht man realistischerweise davon aus, dass die gegebenen Punkte mit kleinen Fehlern behaftet sind, so ist das Gleichungssystem, das durch Einsetzen der Punkte in die Ansatzfunktion entsteht, nicht lösbar. Die Idee ist nun, dass die Parameter abgeschätzt werden sollen, so dass der Fehler möglichst gering ist. Diese Aufgabe ist auch unter dem Namen *polynomialer Ausgleich* bekannt.

Die dargestellte Situation ist nicht nur theoretischer Natur. In vielen Wissenschaften wie zum Beispiel Physik und Biologie werden Naturgesetze oder Modelle beschrieben, deren Parameter noch bestimmt werden müssen. Bei der rationalen Polynominterpolation (Kapitel 1) gibt es keine von vornherein bestimmte Ansatzfunktion und nicht für jeden Ansatz eine Lösung des Interpolationsproblems. Die Methode der kleinsten Quadrate, die hier vorgestellt wird, hat dagegen den Vorteil, dass immer eine Lösung existiert. Je nachdem, ob die Daten wirklich dem Modell entsprechen, ist die Approximation besser oder schlechter.

Die Zahl der Punkte n sollte (vgl. 5.1) größer sein als die Zahl der zu bestimmenden Parameter m . Ganz allgemein gilt dabei, je größer n desto genauer die Approximation, da statistische Ausreisser nicht so ins Gewicht fallen. Darüber hinaus besteht entweder die Möglichkeit, dass durch theoretische Überlegungen fragliche Punkte entweder ganz ausgeschlossen werden, oder zumindest nicht so stark gewichtet werden wie die restlichen.

Andererseits ist es aber auch möglich, dass die gesuchte Funktion Punkte, die zum Beispiel laut Theorie vorgegeben sind, nicht nur annäherungsweise sondern sogar genau durchläuft. Schließlich sei noch eine Möglichkeit gezeigt, wie für einige Funktionen mit nichtpolynomialem Ansatz der Fehler minimiert werden kann.

5.1 Beschreibung der Methode

Nach [12] seien n Punkte (x_i, y_i) $i = 1, \dots, n$ gegeben, die in die Ansatzfunktion mit m Parametern p_1, \dots, p_m eingesetzt werden, so erhalten wir die n Gleichungen

$$\begin{aligned} x_1^0 p_1 + x_1^1 p_2 + \dots + x_1^{m-1} p_m &= y_1, \\ x_2^0 p_1 + x_2^1 p_2 + \dots + x_2^{m-1} p_m &= y_2, \\ &\vdots \\ x_n^0 p_1 + x_n^1 p_2 + \dots + x_n^{m-1} p_m &= y_n. \end{aligned} \tag{5.1}$$

Diese Gleichungen wollen wir in Hinblick auf eine Matrixschreibweise in das folgende Gleichungssystem umschreiben:

$$\begin{aligned} a_{11} p_1 + a_{12} p_2 + \dots + a_{1m} p_m &= y_1, \\ a_{21} p_1 + a_{22} p_2 + \dots + a_{2m} p_m &= y_2, \\ &\vdots \\ a_{n1} p_1 + a_{n2} p_2 + \dots + a_{nm} p_m &= y_n, \end{aligned} \tag{5.2}$$

wobei die Werte x_i^{j-1} durch $a_{ij} \in \mathbb{R}$ ersetzt worden sind. Somit sind nur mehr die Parameterwerte p_i , $i = 1, \dots, m$ unbekannt.

Sind gleich viele Parameter wie Punkte gegeben, ist also $n = m$, so ist kein Ausgleich notwendig, da das Gleichungssystem eindeutig lösbar ist.

Sind weniger Punkte als Parameter gegeben, ist also $n < m$, so kann man $k := m - n$ Parameter frei wählen. Für beliebige k Parameter können nun im Prinzip beliebige Werte eingesetzt werden. Das restliche Gleichungssystem mit $m - k = n$ Unbekannten und n Gleichungen ist wieder ohne Ausgleich eindeutig lösbar.

Der dritte und letzte Fall, dass nämlich mehr Punkte als Parameter gegeben sind, also $m < n$, ist der, mit dem wir uns in diesem Kapitel beschäftigen wollen. Hier existiert nämlich im Allgemeinen keine Lösung des Gleichungssystems (5.2). Dies ist zum Beispiel die Norm für Daten, die aus Messreihen stammen und in der Regel mit Fehlern behaftet sind. Die Idee ist nun Parameter zu finden, so dass zumindest der Fehler $r_i := y_i - (a_{i1} p_1 + a_{i2} p_2 + \dots + a_{im} p_m)$ möglichst klein ist. Die folgende Definition erleichtert die Schreibweise.

Definition 6.

$$r_i := y - (a_{i1}p_1 + a_{i2}p_2 + \cdots + a_{im}p_m). \quad (5.3)$$

Die Werte r_i heissen *Residuen*.

Die Methode der kleinsten Quadrate besagt nun, dass die Summe der Quadrate der einzelnen Residuen möglichst gering werden sollen. Es soll also

$$S(p_1, p_2, \dots, p_m) := r_1^2 + r_2^2 + \cdots + r_n^2 = \sum_{i=1}^n r_i^2 = \min! \quad (5.4)$$

gelten. Die Minimalbedingungen dafür lauten

$$\begin{aligned} \frac{\partial}{\partial p_1} S(p_1, p_2, \dots, p_m) &= 0, \\ \frac{\partial}{\partial p_2} S(p_1, p_2, \dots, p_m) &= 0, \\ &\vdots \\ \frac{\partial}{\partial p_m} S(p_1, p_2, \dots, p_m) &= 0. \end{aligned} \quad (5.5)$$

Nun muss nur noch das Gleichungssystem (5.5) gelöst werden, um die angenäherten Parameter p_1, p_2, \dots, p_m zu erhalten. Dabei gibt $S(p_1, p_2, \dots, p_m) := r_1^2 + r_2^2 + \cdots + r_n^2$ einen Wert für die Abweichung der Daten von der berechneten Funktion an. Für eine weitere Analyse muss aber auf die Regressionsanalyse der angewandten Statistik verwiesen werden.

5.2 Berechnung in Matrizenschreibweise

Das Gleichungssystem (5.2) lässt sich einfach in eine Matrixgleichung umformen. Seien y und p die Spaltenvektoren $y = (y_1, y_2, \dots, y_n)^T$ und $p = (p_1, p_2, \dots, p_m)^T$ und A die $n \times m$ -Matrix mit den Elementen a_{ij} , so gilt

$$Ap \approx y. \quad (5.6)$$

Mit dem Residuenspaltenvektor $r = (r_1, \dots, r_n)^T$ gilt

$$r = y - Ap. \quad (5.7)$$

Für die Summe der Quadrate der Fehler gilt

$$\begin{aligned} S(p) &= S(p_1, p_2, \dots, p_m) = r_1^2 + r_2^2 + \cdots + r_n^2 = \|r\|^2 = r^T r = (y - Ap)^T (y - Ap) = \\ &= (y^T - p^T A^T)(y - Ap) = y^T y - y^T Ap - p^T A^T y + p^T A^T Ap \end{aligned}$$

und da $y^T A p$ ein Skalar ist, gilt weiter

$$= y^T y - 2p^T A^T y + p^T A^T A p. \quad (5.8)$$

Null ist eine untere Schranke von $S(p)$, da ein negativer Fehler nicht zugelassen wird. Andererseits ist die Funktion nach oben nicht beschränkt. Der Fehler kann nämlich beliebig groß werden. Aus der Definition ist ersichtlich, dass nur ein Extremum, nämlich ein Minimum existiert. Um dieses eindeutige Minimum zu berechnen, setzen wir die jeweiligen Ableitung gleich Null, sodass gilt

$$\frac{\partial}{\partial p_i} S(p) = 0, \quad i = 1, \dots, m. \quad (5.9)$$

Daraus folgt

$$2A^T A p - 2A^T y = 0. \quad (5.10)$$

Somit lässt sich der Parametervektor

$$p = (A^T A)^{-1} A^T y \quad (5.11)$$

berechnen.

Bemerkung 17. Was noch fehlt ist der Beweis, dass $(A^T A)^{-1}$ überhaupt existiert. Dies möchte ich in der folgenden Überlegung noch nachholen:

Wir können davon ausgehen, dass die Matrix A maximalen Rang, nämlich $Rg(A) = m$, hat, da die Funktionen x^0, \dots, x^{m-1} bzw. im allgemeinen Fall die Funktionen $\varphi^1, \dots, \varphi^m$ linear unabhängig sind. Zu zeigen ist nun $Rg(A^T A) = m$. Da die Matrix $A^T A$ symmetrisch ist, müssen wir noch beweisen, dass sie auch positiv definit ist, denn dann ist auch gezeigt, dass ihr Rang gleich m ist. Eine Matrix C ist genau dann positiv definit, wenn

$$z^T C z > 0, \quad \text{für } z \in \mathbb{R}^m / \{0\}$$

gilt und $x^T C x = 0$ nur für $x = 0$ folgt. Nun gilt aber

$$z^T A^T A z = \underbrace{(Az)^T}_{=a^T} \underbrace{Az}_{=a} = a^T a, \quad a \in \mathbb{R}^n.$$

$a^T a = 0$ ist aber nur für $a = 0$ erfüllt. Daraus folgt $z = 0$, weil die Spaltenvektoren von A linear unabhängig sind. Somit gilt $z^T A^T A z = 0$ nur für $z = 0$. Das heißt $Rg(A^T A) = m$ und somit ist $A^T A$ invertierbar.

5.3 Gewichtete Methode der kleinsten Quadrate

Wir gehen von einem Ausgleichsproblem aus, das durch

$$r = y - Ap \quad (5.12)$$

gegeben ist. Man kann nun die einzelnen Residuen r_i mit Gewichten $w_i > 0$ werten. Die zu minimierende Funktion lautet dann

$$S_W(p) := \sum_{i=1}^n w_i r_i^2 = r^T W r \quad (5.13)$$

mit der Diagonalmatrix der Gewichte

$$W := \begin{pmatrix} w_1 & & & \\ & w_2 & & 0 \\ & & \ddots & \\ & 0 & & w_{n-1} & \\ & & & & w_n \end{pmatrix}. \quad (5.14)$$

Somit ergibt sich für die zu minimierende Funktion durch analoge Rechnung

$$S_W(p) := r^T W r = y^T W y - 2p^T A^T W y + p^T A^T W A p \quad (5.15)$$

und für das zu lösende Gleichungssystem

$$A^T W A p = A^T W y, \quad (5.16)$$

woraus sich der Parametervektor p durch

$$p = (A^T W A)^{-1} A^T W y \quad (5.17)$$

berechnen lässt.

Bemerkung 18. Offensichtlich ist die Methode der kleinsten Quadrate ein Spezialfall der gewichteten Methode für $w_1 = w_2 = \dots = w_n$. Weiters gilt, je größer w_i im Vergleich zu den restlichen Gewichten, desto mehr Gewicht hat dieser Punkt und desto kleiner muss r_i werden. Die Idee ist also, zweifelhafte Punkte mit Gewichten kleiner 1, „normale“ Punkte mit Gewichten gleich 1 und ziemlich sichere Punkte mit Gewichten größer 1 zu bewerten. Die Gewichtung liegt im Ermessen des Anwenders. Ein Grund für eine Gewichtung eines Punktes ist zum Beispiel dann gegeben, wenn bei Wiederholung der Messung für den Wert x_i genau derselbe Wert y_i wieder auftritt.

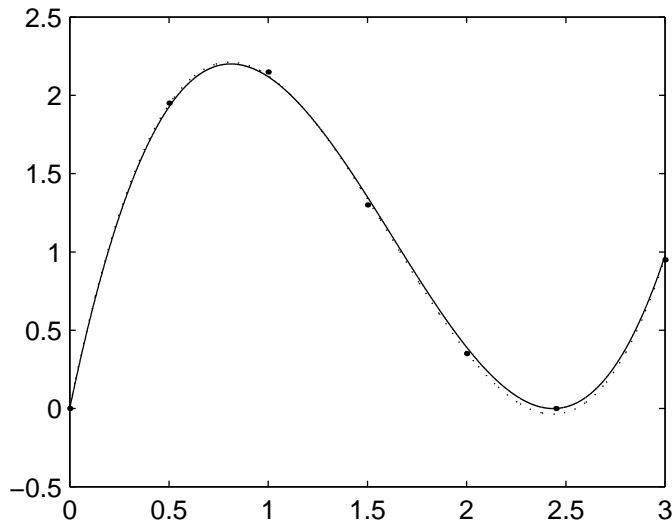


Abbildung 5.1: Interpolation der beiden Nullstellen und Approximation der restlichen Stützstellen mit Hilfe der Methode der kleinsten Quadrate.

5.4 Interpolative Approximation mittels der Methode der kleinsten Quadrate

Wir gehen von folgender leicht geänderter Situation aus: Gegeben sind $n+k$ Punkte, die durch eine Funktion mit $m+k$ Parametern approximiert werden sollen. Für die Punkte (x_i, y_i) , $i = n+1, \dots, n+k$ gilt, dass sie mit der gesuchten Funktion genau inzidieren sollen, während die restlichen Punkte (x_j, y_j) , $j = 1, \dots, n$ nach der bereits bekannten Methode der kleinsten Quadrate anzunähern sind.

Die Idee besteht darin, die Punkte (x_i, y_i) , $i = n+1, \dots, n+k$ in die Ansatzfunktion einzusetzen, die Parameterwerte durch andere auszudrücken und diese zu substituieren. Durch die k gegebenen Punkte erhält man k Gleichungen, die in die Ansatzfunktion eingesetzt werden, so dass k Parameter eliminiert werden. Dadurch hat sich nun die Zahl der Parameter von $m+k$ auf m reduziert. Mit der neuen Ansatzfunktion werden wie bereits beschrieben die Parameter p_1, \dots, p_m ausgerechnet. Die restlichen Parameter p_{m+1}, \dots, p_{m+k} können nun durch die anfangs aufgestellten Gleichungen berechnet werden. Ich möchte die Methode anhand des folgenden Beispiels verdeutlichen.

Beispiel 17. Es seien die Punkte $(0, 0)$, $(0.5, 1.95)$, $(1, 2.15)$, $(1.5, 1.3)$, $(2, 0.35)$, $(2.45, 0)$ und $(3, 0.95)$ gegeben, die durch ein kubisches Polynom zu approximieren sind, allerdings mit der Zusatzaufgabe, dass die Approximationsfunktion genau die selben beiden Nullstellen besitzen soll. Aus dem Punkt $(0, 0)$ folgt eindeutig für unsere Ansatzfunktion $P(x) = p_4x^3 + p_3x^2 + p_2x + p_1$, dass $p_1 = 0$. Setzt man den zweiten Punkt $(2.45, 0)$ in

den Ansatz ein, so erhält man

$$p_2 = -p_4 2.45^2 - p_3 2.45, \quad (5.18)$$

was eingesetzt in die neue Ansatzfunktion

$$\tilde{P}(x) = (x^3 - 2.45^2 x)p_4 + (x^2 - 2.45^2 x)p_3 \quad (5.19)$$

ergibt. Diese Funktion ist nun nach dem beschriebenen Schema zu approximieren. Wir berechnen das zu minimierende Gleichungssystem $A \cdot p = y$, das lautet:

$$\begin{pmatrix} -2.8763 & -0.9750 \\ -5.0025 & -1.4500 \\ -5.6288 & -1.4250 \\ -4.0050 & -0.9000 \\ 8.9925 & 1.6500 \end{pmatrix} \begin{pmatrix} p_4 \\ p_3 \end{pmatrix} = \begin{pmatrix} 1.95 \\ 2.15 \\ 1.30 \\ 0.35 \\ 0.95 \end{pmatrix}. \quad (5.20)$$

Die beiden exakt zu durchlaufenden Nullstellen gehen in dieses Gleichungssystem nicht mehr ein. Wegen (5.17) ergeben sich die Werte $p_4 = 1.0267$ und $p_3 = -5.0038$ aus der Gleichung $p = (A^T A)^{-1} A^T y$. p_2 lässt sich nach (5.18) durch $p_2 = -p_4 2.45^2 - p_3 2.45 = 6.0965$ berechnen. Somit lautet das kubische Approximationspolynom

$$P(x) = 1.0267x^3 - 5.0038x^2 + 6.0965x.$$

Das Ergebnis ist in Abbildung 5.1 zu sehen. Die berechnete Funktion interpoliert die Nullstellen und approximiert die restlichen.

Zum Vergleich ist hier auch die Approximationsfunktion, die alle Punkte nach der Methode der kleinsten Quadrate approximiert, eingezeichnet (punktiert). Diese gibt für die zweite Nullstelle einen zu kleinen Wert an. Als Grundlage der Funktion f wurde das Polynom $x^3 - 2\sqrt{6}x^2 + 6x$ verwendet, dessen Funktionswerte abgesehen von den Nullstellen geringfügig nach oben oder unten verändert wurden. Um die Rechnung nicht durch lange Zahlen unnötig unübersichtlich werden zu lassen, wurde die zweite Nullstelle auf 2 Dezimalstellen gerundet.

5.5 Allgemeinere Ansatzfunktionen

Die Methode der kleinsten Quadrate wurde bisher nur für polynomiale Ansatzfunktionen beschrieben. Wie bereits in der Einleitung erwähnt, ist sie aber für alle Funktionen $\varphi_1, \dots, \varphi_m$, die linear unabhängig sind, anwendbar. An dieser Stelle soll noch auf die Möglichkeit der Variablentransformation aufmerksam gemacht werden, die allgemeine Ansatzfunktion wieder auf die spezielle polynomiale Form bringt und dann

5 Methode der kleinsten Quadrate

mit obigen Verfahren lösbar ist. Möchten wir zum Beispiel die Parameter a und b der Funktion $y = ae^{bx}$ wissen, müssen wir die Variablen so transformieren, dass die neue Funktion $v = \alpha + \beta u$ lautet. Zu diesem Zweck logarithmieren wir die Ausgangsfunktion und erhalten

$$v = \ln y = \ln(ae^{bx}) = \ln a + bx. \quad (5.21)$$

Daraus folgt für die Variablen u und v :

$$\begin{aligned} v &= \ln y \\ u &= x. \end{aligned} \quad (5.22)$$

Die gegebenen Punkte (x_i, y_i) werden nun transformiert zu $(x_i, \ln y_i) = (u_i, v_i)$ und können nun in die lineare Funktion $v = \alpha + \beta u$ zur Erstellung des Gleichungssystems eingesetzt werden. Sind α und β berechnet, so sind die eigentlich gesuchten Parameter a und b gegeben durch

$$\begin{aligned} a &= e^\alpha \\ b &= \beta. \end{aligned} \quad (5.23)$$

Es folgt nun eine aus [12] und [6] zusammengestellte Tabelle, die für einige Funktionen die entsprechenden Transformationen angibt.

Funktion $y = f(x)$	Transformation
$y = ax^b$	$u = \ln x, v = \ln y$
$y = ae^{bx}$	$u = x, v = \ln y$
$y = \frac{a}{x+b}$	$u = x, v = \frac{1}{y}$
$y = \frac{1}{ax+b}$	$u = x, v = \frac{1}{y}$
$y = \frac{ax}{x+b}$	$u = \frac{1}{x}, v = \frac{1}{y}$
$y = \frac{1}{ax^2+bx+c}$	$u = x, v = \frac{1}{y}$
$y = \frac{x}{ax^2+bx+c}$	$u = x, v = \frac{x}{y}$
$y = a + \frac{b}{x} + \frac{c}{x^2}$	$u = \frac{1}{x}, v = y$
$y = a \ln x + b$	$u = \ln x, v = y$
$y = \frac{1}{(ax+b)^2}$	$u = x, v = \frac{1}{y^{\frac{1}{2}}}$

5.6 Beispiele und Bilder

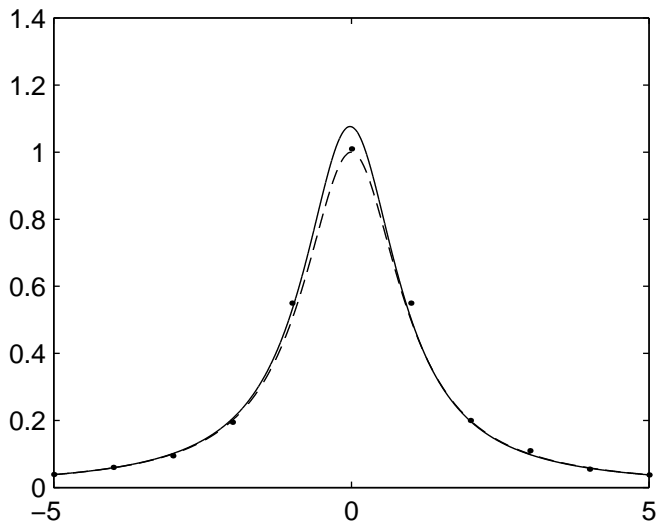


Abbildung 5.2: Approximation von $\frac{1}{1+x^2}$ bei geringen Messfehlern.

Beispiel 18. Die Graphen in den Abbildung 5.2 und 5.3 zeigen beide die Approximation der Funktion $f(x) = \frac{1}{1+x^2}$. Der Ansatz lautet bei beiden $\frac{1}{ax^2+bx+c}$, nur dass, wie aus der folgenden Tabelle ersichtlich ist, die Fehler in den Funktionswerten in der zweiten Abbildung größer sind:

x-Werte	$f(x)$ auf 4 Dezimalen gerundet	Funktionswerte für Abb. 5.2	Funktionswerte für Abb. 5.3
-5	0.0385	0.039	0.044
-4	0.0588	0.06	0.068
-3	0.1	0.095	0.89
-2	0.2	0.195	0.17
-1	0.5	0.55	0.59
0	1	1.01	1.04
1	0.5	0.55	0.37
2	0.2	0.2	0.2
3	0.1	0.11	0.115
4	0.0588	0.055	0.05
5	0.0385	0.038	0.03

Nach Transformation der Punkte von (x, y) zu $(x, \frac{1}{y})$ wenden wir die Methode der kleins-

ten Quadrate an und erhalten als Lösung dann die Approximationspolynome

$$P(x) = \frac{1}{1.0087x^2 + 0.0443x + 0.9295},$$

$$\tilde{P}(x) = \frac{1}{1.0661x^2 + 0.5984x + 0.8788}.$$

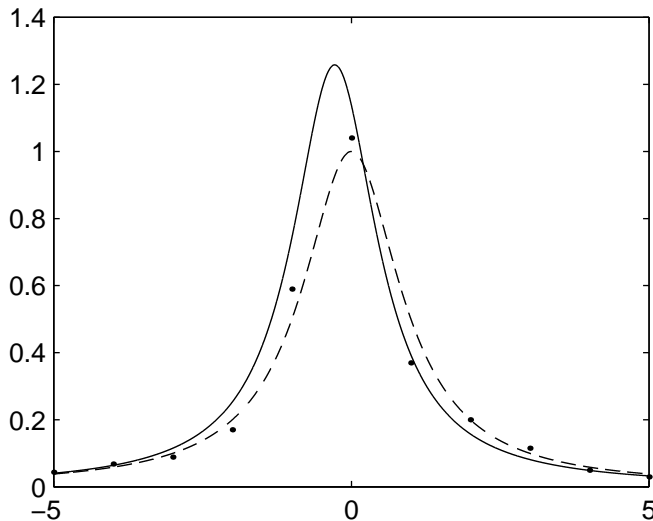


Abbildung 5.3: Approximation von $\frac{1}{1+x^2}$ bei größeren Messfehlern.

Die Funktion P dargestellt in Abbildung 5.2 beschreibt eine gute Annäherung. In Abbildung 5.3 kann man gut erkennen, wie sich noch größere Fehler auswirken. Die Approximationsfunktion ist noch weiter von der eigentlichen entfernt. Zusätzlich ist das Maximum deutlich größer als bei der ersten Approximation. Hinzu kommt, dass die Extremstelle eindeutig zu weit links ist. Das hängt vor allem damit zusammen, dass die zu approximierenden Punkte nicht symmetrisch liegen. Während die gegebenen Punkte von P eher symmetrisch liegen, kann speziell für die Punkte $(-1, 0.59)$ und $(1, 0.37)$ nicht einmal mehr annähernd von Symmetrie gesprochen werden und deshalb ist die Extremstelle verglichen mit der eigentlichen Funktion verschoben.

6 Fourier-Analyse

Möchte man eine periodische Funktion approximieren, so ist es einleuchtend, Linearkombinationen von periodischen Funktionen zu verwenden. Die Funktionen $\sin(x)$ und $\cos(x)$ bieten sich dabei natürlich an. Die daraus resultierenden Fourier-Polynome lassen sich sowohl zur Approximation als auch zur Interpolation verwenden.

Bevor wir auf die diskrete Fourier Analyse eingehen, zitieren wir noch die Ergebnisse für die Approximation von gegebenen Funktionen, da die diskrete Fourier Analyse davon ausgeht. Weiters lässt sich so die Analogie der Berechnung der Koeffizienten gut erkennen.

6.1 Ergebnisse aus der Fourier-Analysis

Die folgenden Ergebnisse werden als bekannt vorausgesetzt und daher nicht hergeleitet, auch wenn sie Ausgangspunkt für unsere weiteren Überlegungen sind. Für die Herleitung sei zum Beispiel auf [9] verwiesen.

Ist eine stetige Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit der Periode 2π gegeben, so lässt sie sich durch ein trigonometrisches Polynom der Form

$$g_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^n \{a_k \cos(kx) + b_k \sin(kx)\} \quad (6.1)$$

nach der Methode der kleinsten Quadrate approximieren, sodass für den Fehler gilt

$$F(x) = \|g_n(x) - f(x)\|_2 = \left[\int_{-\pi}^{\pi} (g_n(x) - f(x))^2 dx \right]^{1/2} = Min!. \quad (6.2)$$

Daraus ergeben sich für die sogenannten *Fourierkoeffizienten* a_k und b_k die Gleichungen

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx, & k = 0, 1, \dots, n, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx, & k = 1, \dots, n. \end{aligned} \quad (6.3)$$

Für $n \rightarrow \infty$ ergibt sich die unendliche *Fourierreihe*

$$g(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} \{a_k \cos(kx) + b_k \sin(kx)\}. \quad (6.4)$$

Wir wollen aber nicht eine gegebene Funktion approximieren, sondern eine durch eine Wertetabelle gegebene Funktion.

6.2 Diskrete Fourier-Analyse

Gegeben sei eine Funktion f mit der Periode 2π durch $N + 1$ äquidistante Stützstellen $(x_j, f(x_j))$. Sei h der Abstand der Stützstellen, so gilt

$$h = \frac{2\pi}{N}, \quad x_j = h \cdot j = \frac{2\pi}{N} \cdot j, \quad j = 0, 1, \dots, N. \quad (6.5)$$

Bemerkung 19. War bis jetzt immer $n + 1$ bzw. n die Anzahl der gegebenen Punkte, so verwenden wir jetzt $N + 1$, weil dies in der Fourier-Analyse so üblich ist. n wird nämlich für den Grad des Fourierpolynoms verwendet.

Es bietet sich an, die Fourierkoeffizienten, die durch Integrale zu berechnen sind, numerisch zu approximieren. Wir verwenden dazu die *Sehnentrapezregel*.

Sehnentrapezregel

Es soll das Integral $\int_a^b f(x) dx$ angenähert werden, wobei das Intervall durch $N + 1$ Stützstellen unterteilt ist, so dass $a = x_0 < x_1 < \dots < x_N = b$. Die Idee besteht darin, das im allgemeinen krumme Kurvenstück durch ein lineares zu ersetzen, also durch die Gerade

$$y(x) = \frac{(f(x_{i+1}) - f(x_i))(x - x_i)}{x_{i+1} - x_i} + f(x_i). \quad (6.6)$$

Da

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{N-1}}^{x_N} f(x)dx \quad (6.7)$$

gilt, wird also jedes Integral durch

$$\begin{aligned} \int_{x_i}^{x_{i+1}} f(x)dx &\approx \int_{x_i}^{x_{i+1}} y(x) = \int_{x_i}^{x_{i+1}} \frac{(f(x_{i+1}) - f(x_i))(x - x_i)}{x_{i+1} - x_i} + f(x_i) = \cdots = \\ &= \frac{1}{2}(f(x_{i+1}) + f(x_i))(x_{i+1} - x_i) \end{aligned} \quad (6.8)$$

angenähert, was durch Ausmultiplizieren und Vereinfachen leicht einzusehen ist.

Annäherung der Fourierkoeffizienten

Wird also die Sehnentrapezregel auf die Integrale der Fourierkoeffizienten (6.3) angewendet, so erhält man

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx \approx \\ &\approx \frac{1}{\pi} \frac{1}{2} \sum_{j=0}^{N-1} \left(f(x_j) \cos(kx_j) + f(x_{j+1}) \cos(kx_{j+1}) \right) \underbrace{(x_{j+1} - x_j)}_{=\frac{2\pi}{N}} = \\ &= \frac{1}{\pi} \frac{1}{2} \frac{2\pi}{N} \left(f(x_0) \cos(kx_0) + 2 \sum_{j=1}^{N-1} f(x_j) \cos(kx_j) + f(x_N) \cos(kx_N) \right) \end{aligned} \quad (6.9)$$

und da $f(x) \cos(kx)$ periodisch ist, gilt $f(x_0) \cos(kx_0) = f(x_N) \cos(kx_N)$ und somit gilt weiter

$$= \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(kx_j). \quad (6.10)$$

Durch analoge Rechnung erhalten wir auch eine Approximation für b_k , so dass somit die angenäherten Fourierkoeffizienten durch

$$\begin{aligned} a_k^* &= \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(kx_j), & k &= 0, 1, \dots \\ b_k^* &= \frac{2}{N} \sum_{j=1}^N f(x_j) \sin(kx_j), & k &= 1, \dots \end{aligned} \quad (6.11)$$

gegeben sind.

Eigenschaften der trigonometrischen Funktionen

Satz 10. Für die äquidistanten Stützstellen $x_j = h \cdot j = \frac{2\pi}{N}j$ mit $h = \frac{2\pi}{N}$ gilt

$$\begin{aligned} \sum_{j=1}^N \cos(kx_j) &= \begin{cases} 0, & \text{falls } k/N \notin \mathbb{Z} \\ N, & \text{falls } k/N \in \mathbb{Z} \end{cases} \\ \sum_{j=1}^N \sin(kx_j) &= 0, \quad \text{für alle } k \in \mathbb{Z}. \end{aligned} \quad (6.12)$$

Beweis. Wir untersuchen die komplexe Summe

$$\begin{aligned} S &= \sum_{j=1}^N [\cos(kx_j) + i \sin(kx_j)] = \sum_{j=1}^N e^{ikx_j} = e^{ikh} + e^{2ikh} + e^{3ikh} + \dots + e^{Nikh} = \\ &= e^{ikh} \left(1 + e^{ikh} + e^{2ikh} + \dots + e^{(N-1)ikh} \right). \end{aligned} \quad (6.13)$$

Diese Summe ist eine geometrische Reihe mit dem Quotienten $q = e^{ikh} = e^{ik2\pi/N}$. Für $k/N \in \mathbb{Z}$ ist $q = 1$ und somit folgt $S = N$.

Für $k/N \notin \mathbb{Z}$ erhalten wir mittels der Summenformel für die geometrische Reihe

$$S = e^{ikh} \cdot \frac{q^N - 1}{q - 1} = e^{ikh} \frac{e^{ik2\pi} - 1}{e^{ikh} - 1} = 0. \quad (6.14)$$

Dabei ergibt der Realteil von S die erste, der Imaginärteil die zweite Gleichung.

□

Satz 11. Für die äquidistanten Stützstellen $x_j = h \cdot j = \frac{2\pi}{N} \cdot j$ mit $h = \frac{2\pi}{N}$ gilt für $k, l \in \mathbb{N}$

$$\begin{aligned} \sum_{j=1}^N \cos(kx_j) \cos(lx_j) &= \begin{cases} 0, & \text{falls } \frac{k+l}{N} \notin \mathbb{Z} \text{ und } \frac{k-l}{N} \notin \mathbb{Z} \\ \frac{N}{2}, & \text{falls entweder } \frac{k+l}{N} \in \mathbb{Z} \text{ oder } \frac{k-l}{N} \in \mathbb{Z} \\ N, & \text{falls } \frac{k+l}{N} \in \mathbb{Z} \text{ und } \frac{k-l}{N} \in \mathbb{Z} \end{cases} \\ \sum_{j=1}^N \sin(kx_j) \sin(lx_j) &= \begin{cases} 0, & \begin{cases} \text{falls } \frac{k+l}{N} \notin \mathbb{Z} \text{ und } \frac{k-l}{N} \notin \mathbb{Z} \\ \text{oder } \frac{k+l}{N} \in \mathbb{Z} \text{ und } \frac{k-l}{N} \in \mathbb{Z} \end{cases} \\ -\frac{N}{2}, & \text{falls } \frac{k+l}{N} \in \mathbb{Z} \text{ und } \frac{k-l}{N} \notin \mathbb{Z} \\ \frac{N}{2}, & \text{falls } \frac{k+l}{N} \notin \mathbb{Z} \text{ und } \frac{k-l}{N} \in \mathbb{Z} \end{cases} \\ \sum_{j=1}^N \cos(kx_j) \sin(lx_j) &= 0, \quad \text{für } k, l \in \mathbb{N} \end{aligned} \quad (6.15)$$

Beweis. Es gelten die folgenden trigonometrischen Identitäten

$$\begin{aligned}\cos(kx_j) \cdot \cos(lx_j) &= \frac{1}{2}[\cos(k+l)x_j + \cos(k-l)x_j], \\ \sin(kx_j) \cdot \sin(lx_j) &= \frac{1}{2}[\cos(k-l)x_j - \cos(k+l)x_j], \\ \cos(kx_j) \cdot \sin(lx_j) &= \frac{1}{2}[\sin(k+l)x_j - \sin(k-l)x_j],\end{aligned}\tag{6.16}$$

aus denen wegen (6.12) die Beziehungen des Satzes folgen.

□

6.2.1 Interpolation und Approximation durch Fourier-Polynome

Interpolation durch Fourier-Polynome

Sind N Stützpunkte $(x_j, f(x_j))$, $j = 1, 2, \dots, N$ gegeben, die durch ein Fourier-Polynom

$$g_n(x) := \frac{1}{2}a_0^* + \sum_{k=1}^n \{a_k^* \cos(kx) + b_k^* \sin(kx)\} \tag{6.17}$$

interpoliert werden, so sind die Koeffizienten durch

$$\begin{aligned}a_k^* &:= \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(kx_j), & k = 0, 1, \dots \\ b_k^* &:= \frac{2}{N} \sum_{j=1}^N f(x_j) \sin(kx_j), & k = 1, \dots\end{aligned}\tag{6.18}$$

eindeutig bestimmt.

Satz 12. *Ist die Stützstellenzahl $N = 2n$ gerade, so ist das Fourier-Polynom eindeutig gegeben durch*

$$g_n(x) := \frac{1}{2}a_0^* + \sum_{k=1}^{n-1} \{a_k^* \cos(kx) + b_k^* \sin(kx)\} + \frac{a_n^*}{2} \cos(nx). \tag{6.19}$$

Ist die Stützstellenanzahl ungerade $N = 2n + 1$, so ist das Fourier-Polynom eindeutig gegeben durch

$$g_n(x) := \frac{1}{2}a_0^* + \sum_{k=1}^n \{a_k^* \cos(kx) + b_k^* \sin(kx)\}. \tag{6.20}$$

Beweis. Der Satz wird für eine gerade Stützstellenanzahl $N = 2n$ bewiesen. Für ungerade n verläuft der Beweis analog. Das Interpolationspolynom ist durch N Punkte festgelegt, wobei $f(x_0) = f(x_N)$ gilt. Durch Einsetzen erhält man ein lineares Gleichungssystem mit N Gleichungen und N Unbekannten.

$$g_n(x_j) = \frac{1}{2}\alpha_0 + \sum_{k=1}^{n-1} \{\alpha_k \cos(kx_j) + \beta_k \sin(kx_j)\} + \frac{\alpha_n}{2} \cos(nx_j) = f(x_j), \quad j = 1, \dots, N. \quad (6.21)$$

Die Spalten sind wegen (6.15) paarweise orthogonal und von Null verschieden. Daher hat die Matrix des Gleichungssystems maximalen Rang und das Gleichungssystem ist eindeutig lösbar.

Zur Berechnung der α_l wird das Gleichungssystem mit $\cos(lx_j)$ multipliziert, wobei für ein fixes l , $0 \leq l \leq n$ gelten soll. Werden alle Gleichungen für $j = 1, \dots, N$ anschließend summiert, erhält man ausgehend von

$$g_n(x_j) = \frac{1}{2}\alpha_0 + \sum_{k=1}^{n-1} \{\alpha_k \cos(kx_j) + \beta_k \sin(kx_j)\} + \frac{\alpha_n}{2} \cos(nx_j) = f(x_j) \quad j = 1, \dots, N \quad (6.22)$$

die Gleichung

$$\begin{aligned} & \frac{1}{2}\alpha_0 \sum_{j=1}^N \cos(lx_j) + \sum_{k=1}^{n-1} \left(\alpha_k \sum_{j=1}^N \cos(lx_j) \cos(kx_j) + \beta_k \sum_{j=1}^N \cos(lx_j) \sin(kx_j) \right) + \\ & + \frac{\alpha_n}{2} \sum_{j=1}^N \cos(lx_j) \cos(nx_j) = \sum_{j=1}^N \cos(lx_j) f(x_j). \end{aligned} \quad (6.23)$$

Die linke Seite der Gleichung vereinfacht sich wegen (6.12) und (6.15) zu $\frac{N}{2}\alpha_l$, sodass gilt

$$\alpha_l = \frac{2}{N} \sum_{j=1}^n \cos(lx_j) f(x_j), \quad (6.24)$$

was aber genau der Definition von a_l^* entspricht.

Um β_l zu bestimmen, geht man analog vor, nur dass man das Gleichungssystem (6.22) mit $\sin(lx_j)$ multiplizieren muss.

□

Approximation durch Fourier-Polynome

Gegeben seien N Punkte, die durch ein Fourier-Polynom g_n n -ten Grades approximiert werden sollen. Dabei ist die Zahl der gesuchten Koeffizienten kleiner als die Zahl der Stützstellen, also gilt $2n < N$. Möchte man nun die Koeffizienten mittels der Methode der kleinsten Quadrate approximieren, so erhält man

$$g_n(x) := \frac{1}{2}a_0^* + \sum_{k=1}^n \{a_k^* \cos(kx) + b_k^* \sin(kx)\} \quad (6.25)$$

wieder mit den Koeffizienten

$$\begin{aligned} a_k^* &:= \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(kx_j), & k = 0, 1, \dots, n, \\ b_k^* &:= \frac{2}{N} \sum_{j=1}^N f(x_j) \sin(kx_j), & k = 1, \dots, n. \end{aligned} \quad (6.26)$$

Satz 13. *Das Fourier-Polynom*

$$g_n(x) := \frac{1}{2}a_0^* + \sum_{k=1}^n \{a_k^* \cos(kx) + b_k^* \sin(kx)\} \quad (6.27)$$

mit den Koeffizienten (6.18) approximiert die Funktion $f(x)$ entlang der Punkte $(x_j, f(x_j))$ so, dass die Summe der Quadrate der Abweichungen minimal ist.

Beweis. Wir gehen vom Fourier-Polynom

$$g_n(x) = \frac{1}{2}\alpha_0 + \sum_{k=1}^n \{\alpha_k \cos(kx) + \beta_k \sin(kx)\} \quad (6.28)$$

aus, dessen Koeffizienten $\alpha_0, \alpha_k, \beta_k, k = 1, \dots, n$ noch zu bestimmen sind, sodass

$$F(\alpha_0, \alpha_k, \beta_k) := \sum_{j=1}^N \left[\frac{1}{2}\alpha_0 + \sum_{k=1}^n \{\alpha_k \cos(kx_j) + \beta_k \sin(kx_j)\} - f(x_j) \right]^2 \quad (6.29)$$

minimal wird. Die Koeffizienten $\alpha_0, \alpha_k, \beta_k$ lassen sich aus den $2n + 1$ Bedingungen

$\frac{\partial F}{\partial \alpha_0} = 0$, $\frac{\partial F}{\partial \alpha_k} = 0$, $\frac{\partial F}{\partial \beta_k} = 0$ berechnen.

$$\begin{aligned}
 \frac{\partial F}{\partial \alpha_0} &= 2 \cdot \frac{1}{2} \sum_{j=1}^N \left[\frac{1}{2} \alpha_0 + \sum_{k=1}^n \{ \alpha_k \cos(kx_j) + \beta_k \sin(kx_j) \} - f(x_j) \right] = \\
 &= \sum_{j=1}^N \left[\frac{1}{2} \alpha_0 - f(x_j) \right] + \underbrace{\sum_{k=1}^n \sum_{j=1}^N \{ \alpha_k \cos(kx_j) + \beta_k \sin(kx_j) \}}_{=0 \text{ wegen (6.12)}} = \\
 &= \frac{1}{2} N \alpha_0 - \sum_{j=1}^N f(x_j) = 0 \\
 &\rightarrow \alpha_0 = \frac{2}{N} \sum_{j=1}^N f(x_j).
 \end{aligned} \tag{6.30}$$

Für $0 < l \leq n$ gilt

$$\begin{aligned}
 \frac{\partial F}{\partial \alpha_l} &= 2 \sum_{j=1}^N \cos(lx_j) \left[\frac{1}{2} \alpha_0 + \sum_{k=1}^n \{ \alpha_k \cos(kx_j) + \beta_k \sin(kx_j) \} - f(x_j) \right] = \\
 &= 2 \left(\alpha_l \frac{N}{2} - \sum_{j=1}^N f(x_j) \cos(lx_j) \right) = 0 \\
 &\rightarrow a_l = \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(lx_j).
 \end{aligned} \tag{6.31}$$

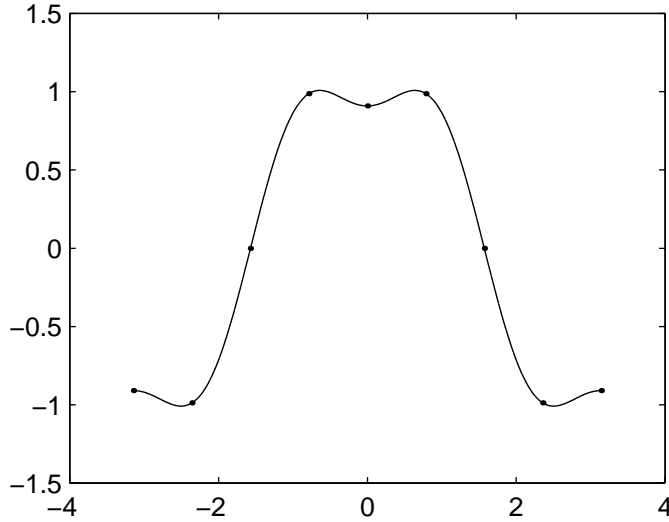
Analog ergibt sich $\beta_l = \frac{2}{N} \sum_{j=1}^N f(x_j) \sin(lx_j)$. Wir vergleichen nun die berechneten Koeffizienten α_0 , α_l und β_l mit den Definitionen der Koeffizienten a_0^* , a_l^* und b_l^* für $l = 1, 2, \dots, n$ und sehen, dass sie übereinstimmen. Damit ist der Satz bewiesen.

□

Bemerkung 20. Möchte man eine Funktion mit Periode $\tau \neq 2\pi$ approximieren, so wendet man die Transformation $x \rightarrow \bar{x} = \frac{2\pi}{\tau}x$ an und erhält eine Funktion $F(\bar{x}) = F(\frac{2\pi}{\tau}x)$ mit Periode 2π .

Bemerkung 21. Ist eine Funktion $f(x)$ ungerade, so gilt $a_k^* = 0$; ist sie gerade so, gilt $b_k^* = 0$ für alle k .

6.2.2 Beispiele und Bilder

Abbildung 6.1: $\sin(2 \cos(x))$ approximiert durch ein Fourier-Polynom 4. Grades.

Beispiel 19. Interpoliert man die Funktion $f = \sin(2(\cos(x)))$ mit der Fourier-Methode, so erhält man bereits für ein Fourier-Polynom 4. Grades, also für nur 8 Stützstellen, ein sehr gutes Ergebnis (vgl. Abb. 6.1). Wir benötigen folgende Angaben:

$$\begin{array}{c|c|c|c|c|c|c|c} -\frac{3}{4}\pi & -\frac{\pi}{2} & -\frac{\pi}{4} & 0 & \frac{\pi}{4} & \frac{\pi}{2} & \frac{3}{4}\pi & \pi \\ \hline -\sin\sqrt{2} & 0 & \sin\sqrt{2} & \sin 2 & \sin\sqrt{2} & 0 & -\sin\sqrt{2} & -\sin 2 \end{array}.$$

Zusätzlich wissen wir, dass auf Grund der Periodizität $f(-\pi) = f(\pi) = -\sin 2$ gilt. Nun lassen sich die Koeffizienten wie folgt berechnen:

$$a_0^* = \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(0 \cdot x_j) = \frac{1}{4}(-\sin\sqrt{2} + \sin\sqrt{2} + \sin 2 + \sin\sqrt{2} - \sin\sqrt{2} - \sin 2) = 0$$

$$\begin{aligned} a_1^* &= \frac{2}{N} \sum_{j=1}^N f(x_j) \cos(1 \cdot x_j) = \frac{1}{4}(-\sin\sqrt{2} \cdot (-\sqrt{\frac{1}{2}}) + \sin\sqrt{2} \cdot \sqrt{\frac{1}{2}} + \sin 2 + \sin\sqrt{2} \cdot \sqrt{\frac{1}{2}} - \\ &\quad - \sin\sqrt{2} \cdot (-\sqrt{\frac{1}{2}}) + \sin 2) = \frac{\sin\sqrt{2} \cdot \sqrt{2} + \sin 2}{2} \approx 1.1531. \end{aligned}$$

Für die restlichen Koeffizienten geht man analog vor und erhält

$$\begin{aligned} a_2^* &= 0, \\ a_3^* &= \frac{-\sin\sqrt{2} \cdot \sqrt{2} + \sin 2}{2} \approx -0.2438, \\ a_4^* &= 0. \end{aligned}$$

Wie in Bemerkung 21 bereits erwähnt, sind alle b_k^* wegen der Symmetrie $f(-x) = f(x)$ gleich Null. Somit ergibt sich das Fourierpolynom

$$g_4(x) = 1.1531 \cos(x) - 0.2438 \cos(3x).$$

In diesem Maßstab ist zwischen der approximierten und der eigentlichen Funktion kein Unterschied zu erkennen. Zum Vergleich ist $\sin(2 \cos(x))$ in Abbildung 7.4 angegeben.

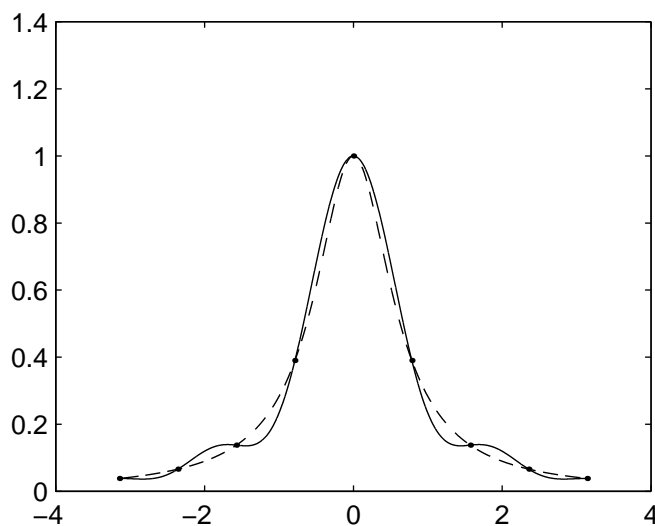


Abbildung 6.2: $\frac{1}{1+x^2}$ approximiert durch ein Fourier-Polynom 4. Grades.

Beispiel 20. Fourierinterpolation ist aber nicht nur für periodische Funktionen geeignet, sondern auch für solche, deren Funktionswerte an den Rändern übereinstimmen. Dies ist zum Beispiel für $f(x) = \frac{1}{1+x^2}$ der Fall. Sie entspricht der Funktion $\bar{f}(\bar{x}) = \frac{1}{1+(\frac{5}{\pi}\bar{x})^2}$ für $\bar{x} \in [-\pi, \pi]$. Aus der Wertetabelle

$-\frac{3}{4}\pi$	$-\frac{\pi}{2}$	$-\frac{\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3}{4}\pi$	π
0.0664	0.1379	0.3902	1	0.3902	0.1379	0.0664	0.0385

lassen sich die Fourierkoeffizienten berechnen, sodass das zugehörige Fourierpolynom

$$g_4(\bar{x}) = 0.2784 + 0.3549 \cos(\bar{x}) + 0.1907 \cos(2\bar{x}) + 0.1259 \cos(3\bar{x}) + 0.0501 \cos(4\bar{x})$$

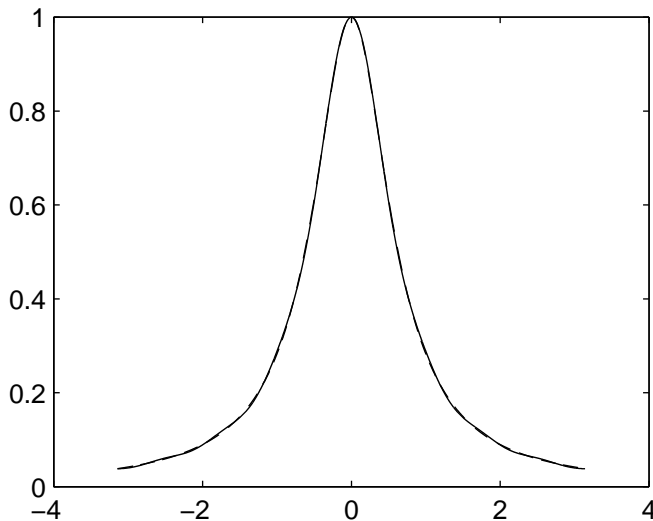


Abbildung 6.3: $\frac{1}{1+x^2}$ approximiert durch ein Fourier-Polynom 8. Grades.

lautet. Dieses ist in Abbildung 6.2 gezeichnet. Falls man mit dieser Näherung noch nicht zufrieden ist, so erhält man zum Beispiel durch Verdoppelung der Stützstellenanzahl eine erhebliche Verbesserung. Wir nehmen die Werte

$-\frac{7\pi}{8}$	$-\frac{5\pi}{8}$	$-\frac{3\pi}{8}$	$-\frac{\pi}{8}$	$\frac{\pi}{8}$	$\frac{3\pi}{8}$	$\frac{5\pi}{8}$	$\frac{7\pi}{8}$
0.0497	0.0929	0.2215	0.7191	0.7191	0.2215	0.0929	0.0497

noch hinzu und erhalten das in Abbildung 6.3 dargestellte Fourierpolynom

$$g_8(\bar{x}) = 0.2746 + 0.3444 \cos(\bar{x}) + 0.1757 \cos(2\bar{x}) + 0.9730 \cos(3\bar{x}) + 0.0501 \cos(4\bar{x}) + \\ + 0.0286 \cos(5\bar{x}) + 0.0150 \cos(6\bar{x}) + 0.0105 \cos(7\bar{x}) + 0.0038 \cos(8\bar{x}).$$

In diesem Maßstab ist nur mehr wenig Unterschied zwischen Approximation und Original festzustellen. Zum Vergleich ist auch die Ausgangsfunktion eingezeichnet. Zu bedenken ist, dass das Fourier-Polynom wegen seiner Periodizität an den Stellen $-\pi$ und π ein Minimum annimmt, während $\frac{1}{1+x^2}$ an denselben Stellen streng monoton steigend bzw. fallend ist. g_4 lässt diese Tatsache noch recht gut erkennen, während dies für g_8 kaum mehr auffällt.

Beispiel 21. Funktionen mit Schwingungen sind für die Fourier-Interpolation kein Problem. Vorsichtiger muss man aber mit Funktionen sein, die eine kleine Krümmung haben, denn Fourierpolynome tendieren dazu, etwas zu schwingen. Allerdings ist das Problem verglichen mit der Polynominterpolation einfach zu lösen. Je mehr Interpolationspunkte verwendet werden, desto mehr Schwingungen entstehen zwar, aber desto kleiner fallen sie auch aus. Als Extrembeispiel mit Krümmung Null soll die Funktion $f(x) = |x|$

betrachtet werden (Abbildung 6.4). Die Wertetabelle für g_4 lautet

$-\frac{3}{4}\pi$	$-\frac{\pi}{2}$	$-\frac{\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3}{4}\pi$	π
$\frac{3}{4}\pi$	$\frac{\pi}{2}$	$\frac{\pi}{4}$	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3}{4}\pi$	π

Für g_8 wird sie durch die Punkte

$-\frac{7}{8}\pi$	$-\frac{5}{8}\pi$	$-\frac{3}{8}\pi$	$-\frac{1}{8}\pi$	$\frac{1}{8}\pi$	$\frac{3}{8}\pi$	$\frac{5}{8}\pi$	$\frac{7}{8}\pi$
$\frac{7}{8}\pi$	$\frac{5}{8}\pi$	$\frac{3}{8}\pi$	$\frac{1}{8}\pi$	$\frac{1}{8}\pi$	$\frac{3}{8}\pi$	$\frac{5}{8}\pi$	$\frac{7}{8}\pi$

ergänzt. Daraus lassen sich die Koeffizienten berechnen, die eingesetzt die Fourierpolynome

$$g_4(x) = 1.5708 - 1.3408 \cos(x) - 0.2300 \cos(3x),$$

$$g_8(x) = 1.5708 - 1.2897 \cos(x) - 0.1590 \cos(3x) - 0.0710 \cos(5x) - 0.0510 \cos(7x)$$

ergeben.

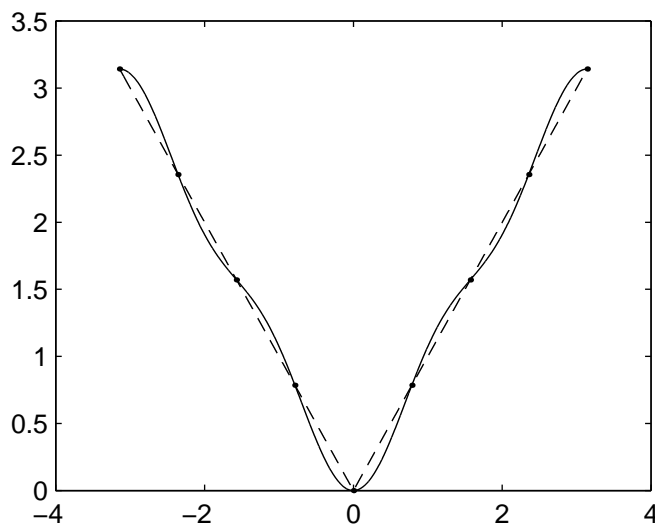


Abbildung 6.4: $|x|$ approximiert durch ein Fourier-Polynom 4. Grades.

Das Fourier-Polynom 4. Grades hat größere Amplituden als das Polynom 8. Grades. Weiters fällt auf, dass die Approximation im Knick natürlich schlechter ist als sonst. Das ist darauf zurückzuführen, dass das Fourier-Polynom \mathcal{C}^∞ ist, $|x|$ aber nicht. Auch die Approximation in diesem Punkt verbessert sich mit Erhöhung der Stützstellen.

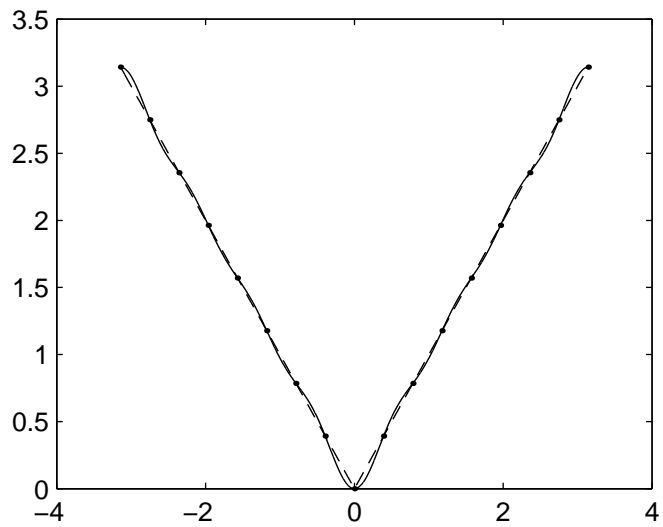


Abbildung 6.5: $|x|$ approximiert durch ein Fourier-Polynom 8. Grades.

7 Approximation mittels Taylorformel

Gleich zu Beginn muss festgehalten werden, dass der Bereich der Taylorapproximation das eigentliche Thema dieser Arbeit nur tangiert. Für die Taylorapproximation muss zwar eine Wertetabelle gegeben sein, aber diese besteht nicht wie bisher aus einer Reihe von Stützpunkten, sondern aus den ersten k Ableitungen in einem Punkt m samt dem Funktionswert an dieser Stelle. Aus diesem Grund wird die Taylorapproximation hier zwar erwähnt, aber nicht im Detail besprochen. Es soll genügen, relevante Sätze zu zitieren und einige Beispiele zu betrachten.

7.1 Beschreibung der Methode

Dieses Verfahren geht auf Taylorreihen zurück und den aus der Analysis bekannten Satz [5], dass sich jede Potenzreihe $p : (m-r, m+r) \rightarrow \mathbb{R}$, $p(x) = a_0 + a_1(x-m) + a_2(x-m)^2 + a_3(x-m)^3 + \dots$ mit Mittelpunkt m und Konvergenzradius r , für den $0 < r \leq \infty$ gelten muss, auf $(m-r, m+r)$ durch

$$p(x) = p(m) + p'(m)(x-m) + \frac{p''(m)}{2!}(x-m)^2 + \frac{p'''(m)}{3!}(x-m)^3 + \dots \quad (7.1)$$

darstellen lässt.

Für das Approximationsverfahren bedeutet dies, dass möglichst viele Ableitungen in einem Punkt gegeben sein sollten. Sind der Funktionswert $f(m)$ und die k Ableitungen $f'(m)$, $f''(m)$, \dots , $f^{(k)}(m)$ gegeben, so bezeichnet

$$t_k(x) := f(m) + f'(m)(x-m) + \frac{f''(m)}{2!}(x-m)^2 + \dots + \frac{f^{(k)}(m)}{k!}(x-m)^k \quad (7.2)$$

die Taylorreihenentwicklung bis zum k -ten Glied.

Für eine Polynomfunktion bedeutet dieses Verfahren, dass, sofern alle Ableitungen bekannt sind und keine Mess- oder Rechenfehler das Ergebnis verzerren, diese eindeutig gegeben ist.

Es stellt sich nun die Frage nach dem Fehler für andere Funktionen bei dieser Methode. Sind $k + 1$ Ableitungen gegeben, so kann dieser für t_k und ein bestimmtes $x_0 \in (m - r, m + r)$ nach [5] durch

$$f(x_0) - t_k(x_0) = \frac{1}{k!} \int_m^{x_0} f^{(k+1)}(z)(x_0 - z)^k dz \quad (7.3)$$

angegeben werden, der sich nach dem Restglied nach Lagrange zu

$$f(x_0) - t_k(x_0) = \frac{f^{(k+1)}(\xi(x_0))}{(k+1)!} (x_0 - m)^{k+1} \quad (7.4)$$

umformen lässt. ξ ist dabei ein von x_0 abhängiger Wert zwischen m und x_0 .

Für die Praxis bedeutet dies, dass im Entwicklungspunkt m alle k Ableitungen übereinstimmen und dass der Fehler zunehmend größer wird, je weiter man sich von m entfernt. Daraus folgt, dass diese Methode für ein Intervall $I = (m - \epsilon_1, m + \epsilon_2)$, $\epsilon_1, \epsilon_2 > 0$ brauchbar ist.

Bemerkung 22. Ist eine Funktion f nun k mal differenzierbar, so besteht entweder die Möglichkeit t_{k-1} zu berechnen und zu versuchen den Fehler abzuschätzen, oder, wenn man keinen Wert auf eine Abschätzung legt, diese zu kompliziert oder gar nicht möglich ist, t_k zu berechnen.

7.2 Beispiele und Bilder

Beispiel 22. Beginnen wir damit, die Funktion $f = x^2 \cdot |x|$ an der Stelle $(2, 8)$ mittels Taylorformel zu entwickeln. Aus $f(2) = 8$, $f'(2) = 12$, $f''(2) = 12$, $f'''(2) = 6$ und der Tatsache, dass alle höheren Ableitungen im Punkt 2 gleich Null sind, ergibt sich wegen (7.2) eindeutig das Taylorpolynom $t_3(x) = 8 + 12(x - 2) + 6(x - 2)^2 + (x - 2)^3$.

Für positive x entspricht f dem Polynom x^3 und dem Taylorpolynom t_4 , da Polynome durch eine Taylorentwicklung bei genügend großem Grad exakt approximiert werden. Die Schwäche der Methode offenbart sich bei Betrachtung der negativen x -Werte (vgl. Abb. 7.2). Die Betragsfunktion $|x|$ ist in sämtlichen Ableitungen an der Stelle 2 nicht von der Funktion x zu unterscheiden. Deshalb teilen sich Taylorpolynom und f im Punkt $(0, 0)$. Während das Taylorpolynom der Funktion x^3 für negative x entspricht, beträgt der y -Wert von f gleich $-x^3$. Eine Verbesserung der Approximation für negative x -Werte gibt es nicht, da die Möglichkeiten der Taylormethode schon ausgeschöpft sind. Ab der 4. Ableitung sind nämlich sämtliche höheren Ableitungen gleich Null und damit ist eine Erhöhung des Grades nicht mehr möglich.

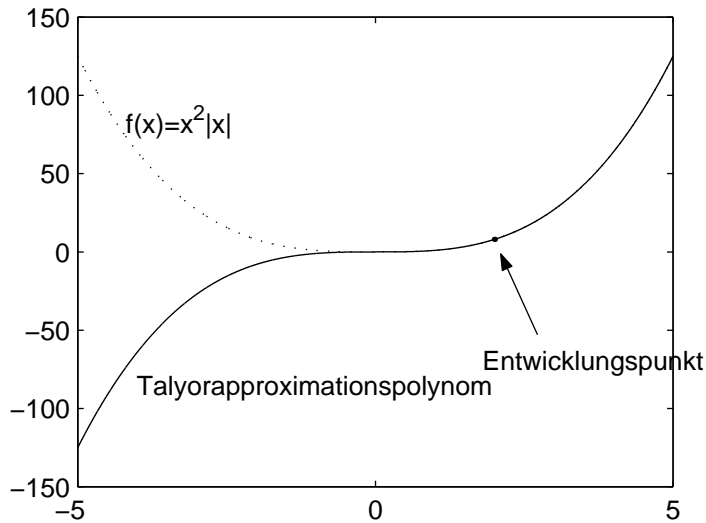


Abbildung 7.1: Approximation von $f(x) = x^2 \cdot |x|$ mittels eines Taylorpolynoms höchstmöglichen Grades.

Beispiel 23. Wie die Taylorentwicklung für $f = \frac{1}{1+x^2}$ aussieht ist in Abbildung 7.2 durch die Polynome t_4 , t_8 und t_{16} gezeigt. Sind die folgende Werte gegeben

$$\begin{aligned} f(0) &= 1! & f^{(9)}(0) &= 0 \\ f'(0) &= 0 & f^{(10)}(0) &= -10! \\ f''(0) &= -2! & f^{(11)}(0) &= 0 \\ f'''(0) &= 0 & f^{(12)}(0) &= 12! \\ f^{(4)}(0) &= 4! & f^{(13)}(0) &= 0 \\ f^{(5)}(0) &= 0 & f^{(14)}(0) &= -14! \\ f^{(6)}(0) &= -6! & f^{(15)}(0) &= 0 \\ f^{(7)}(0) &= 0 & f^{(16)}(0) &= 16 \\ f^{(8)}(0) &= 8! \end{aligned}$$

so lauten die Taylorpolynome:

$$\begin{aligned} t_4(x) &= 1 - x^2 + x^4 \\ t_8(x) &= 1 - x^2 + x^4 - x^6 + x^8 \\ t_{16}(x) &= 1 - x^2 + x^4 - x^6 + x^8 - x^{10} + x^{12} - x^{14} + x^{16} \end{aligned}$$

Es stellt sich heraus, dass alle 3 Polynome im Bereich um den Entwicklungspunkt $(0, 1)$ eine gute Approximation ohne unerwünschte Schwingungen darstellen. Je höher der Grad des Polynoms, desto größer ist das Intervall $I = (0 - \epsilon_1, 0 + \epsilon_2)$, für das das Taylorpolynom eine gute Approximation darstellt. Allerdings ist auch gut zu sehen, dass je höher der Grad von t_k ist, desto steiler ist der Anstieg des Approximationspolynoms ausserhalb des Intervalls I . Der Funktionswert von $t_{16}(\pm 5)$ liegt in der Größenordnung von 10^{11} ! Abbildung 7.2 zeigt noch einmal diese Funktion aber in einem anderen Maßstab.

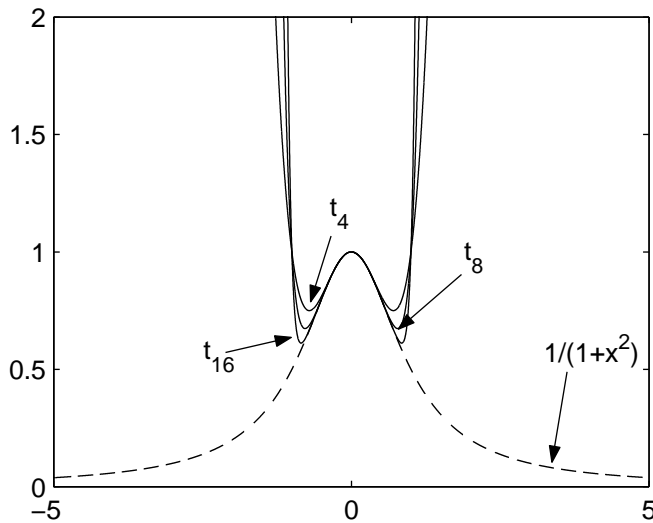


Abbildung 7.2: Der Bereich der „guten“ Approximation vergrößert sich für Taylorpolynome mit dem Erhöhen des Grades nur langsam; zusätzlich wächst der Fehler ausserhalb dieses Bereichs.

Beispiel 24. Abschließend möchte ich noch die Taylorformel auf die Funktion $f(x) = \sin(2 \cos(x))$ anwenden (siehe Abb. 7.4). In diesem Beispiel geht es einerseits darum, eine periodische Funktion zu approximieren, andererseits soll auch beleuchtet werden, inwieweit die Position des Entwicklungspunktes einen Einfluss auf das Approximationsverhalten von t_k hat. Es sind wieder die Ableitungen gegeben:

$$\begin{aligned}
 f(0) &= \sin(2) & f^{(9)}(0) &= 0 \\
 f'(0) &= 0 & f^{(10)}(0) &= -12602 \cos(2) - 49380 \sin(2) \\
 f''(0) &= -2 \cos(2) & f^{(11)}(0) &= 0 \\
 f'''(0) &= 0 & f^{(12)}(0) &= 1494242 \cos(2) + 383988 \sin(2) \\
 f^{(4)}(0) &= 2 \cos(2) - 12 \sin(2) & f^{(13)}(0) &= 0 \\
 f^{(5)}(0) &= 0 & f^{(14)}(0) &= 41337660 \sin(2) - 41681642 \sin(2) \\
 f^{(6)}(0) &= 118 \cos(2) + 60 \sin(2) & f^{(15)}(0) &= 0 \\
 f^{(7)}(0) &= 0 & f^{(16)}(0) &= -5763749438 \cos(2) - 2663125932 \sin(2) \\
 f^{(8)}(0) &= 1428 \sin(2) - 1678 \cos(2)
 \end{aligned}$$

Das Taylorpolynom t_k ist dann durch

$$t_k(x) = \sum_{i=1}^k \frac{f^{(i)}(0)}{i!} (x - m)^i \quad (7.5)$$

gegeben.

Anhand dieses Beispiels ist zu erkennen, dass die Periodizität für die Taylorapproximation keine Vorteile bringt. Wie bereits in den vorhergehenden Beispielen existiert

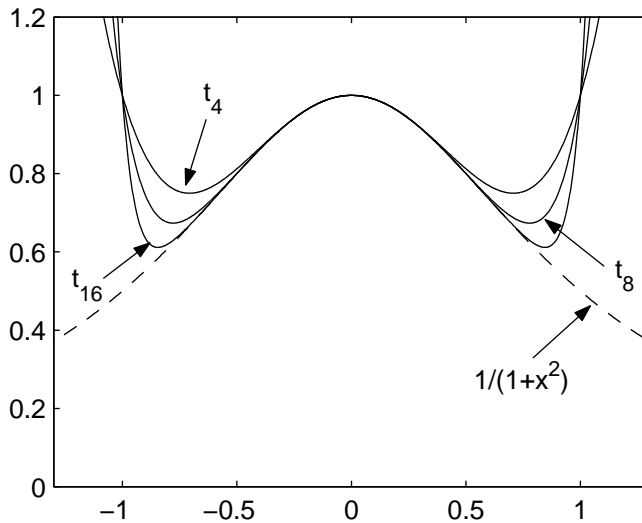


Abbildung 7.3: Taylorpolynome hohen wie niedrigen Grades schwingen nicht im Bereich um den Entwicklungspunkt und bieten auch schon ab einem niedrigen Grad eine passable Approximation in einem gewissen Bereich.

ein Intervall, für das die Taylorpolynome eine gute Approximation bieten. Für t_{16} wird sogar ungefähr die Hälfte des Intervalls $(-\pi, \pi)$ gut approximiert. Das hat weniger mit der Periodizität der zu approximierenden Funktion zu tun, denn das Taylorpolynom selbst ist ja auch nicht periodisch, als mit dem kleinen Definitionsintervall.

Interessanter wird es schon, wenn man die Abbildungen 7.4 und 7.5 bezüglich ihres Entwicklungspunktes vergleicht. Wir geben uns die Ableitungen in Punkt $(\frac{\pi}{3}, f(\frac{\pi}{3}))$ vor und erhalten

$$\begin{aligned} f\left(\frac{\pi}{3}\right) &= \sin(1) & f^{(5)}\left(\frac{\pi}{3}\right) &= 45 \sin(1) - 23\sqrt{3} \cos(1) \\ f'\left(\frac{\pi}{3}\right) &= -\sqrt{3} \cos(1) & f^{(6)}\left(\frac{\pi}{3}\right) &= -346 \cos(1) - 105 \sin(1) \\ f''\left(\frac{\pi}{3}\right) &= -\cos(1) - 3 \sin(1) & f^{(7)}\left(\frac{\pi}{3}\right) &= 91\sqrt{3} \cos(1) - 882\sqrt{3} \sin(1) \\ f'''\left(\frac{\pi}{3}\right) &= 4\sqrt{3} \cos(1) - 3\sqrt{3} \sin(1) & f^{(8)}\left(\frac{\pi}{3}\right) &= 7225 \cos(1) - 819 \sin(1) \\ f^{(4)}\left(\frac{\pi}{3}\right) &= 19 \cos(1) + 18 \sin(1) \end{aligned}$$

Daraus lässt sich das auf 4 Stellen gerundete Taylorpolynom

$$\begin{aligned} t_8(x) &= 0.8415 - 0.9358\left(x - \frac{\pi}{3}\right) - 1.5324\left(x - \frac{\pi}{3}\right)^2 - 0.1048\left(x - \frac{\pi}{3}\right)^3 + \\ &\quad + 1.0588\left(x - \frac{\pi}{3}\right)^4 + 0.3516\left(x - \frac{\pi}{3}\right)^5 - 0.3824\left(x - \frac{\pi}{3}\right)^6 - \\ &\quad - 0.2382\left(x - \frac{\pi}{3}\right)^7 + 0.0797\left(x - \frac{\pi}{3}\right)^8 \end{aligned}$$

berechnen. Der Entwicklungspunkt $(0, f(0))$ stellt sich eindeutig als besser geeignet heraus als der Punkt $(\frac{\pi}{3}, f(\frac{\pi}{3}))$. Der Grund dafür besteht darin, dass die erste Entwicklung

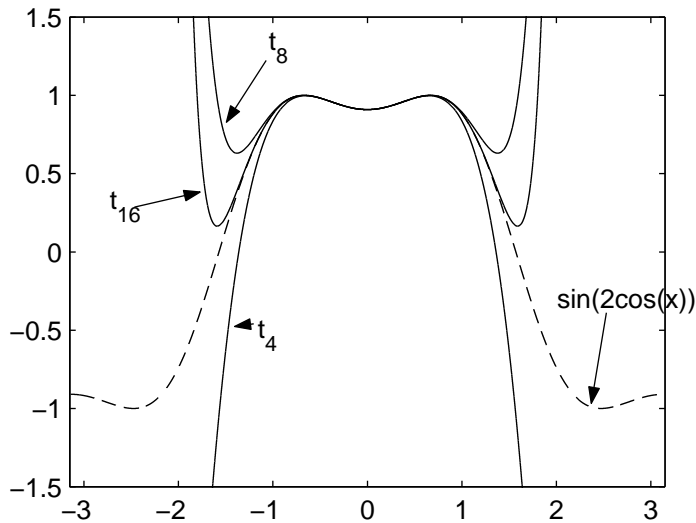


Abbildung 7.4: Taylorapproximation einer periodischen Funktion im Symmetriepunkt.

im Symmetriepunkt der Funktion f stattfindet. Dadurch „erkennt“ die Taylormethode, dass es sich um eine gerade Funktion handeln muss und entsprechend sind auch alle Koeffizienten ungerader Potenzen gleich 0, weil alle ungeraden Ableitungen gleich Null sind. Zwar ist das Intervall, für das das Taylorpolynom den Tendenzen von f folgt, noch ähnlich groß wie im vorhergehenden Beispiel, aber das Intervall I , für das das Taylorpolynom eine wirklich gute Approximation bietet, ist deutlich kleiner geworden.

Bemerkung 23. In allen Beispielen ist zu erkennen, dass die Intervalle guter Approximation $I_k, I_{k+1}, I_{k+2}, \dots$ für die entsprechenden Funktionen $t_k, t_{k+1}, t_{k+2}, \dots$ für wachsendes k immer langsamer größer werden.

Bemerkung 24. Für eine entsprechend oft differenzierbare Funktion f lässt sich der maximale Fehler für ein bestimmtes Intervall I bei vorgegebenem Grad von t_k berechnen, wenn man $\max_{x \in I} f^{(k+1)}(x)$ abschätzen kann [2].

Beispiel 25. Die Funktion $f(x) = \sin(x)$ lässt sich zum Beispiel leicht abschätzen. Entwickelt man f für $m = 0$ durch ein Taylorpolynom 4. Grades im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$, so erhält man: $t_4(x) = x - \frac{1}{6}x^3$. Wegen $f^{(5)}(x) = \sin(x) \leq 1$ und $(\frac{\pi}{2})^5 \leq 0.96$ beträgt der Fehler $f(x) - t_4(x) = \frac{1}{5!}f^{(5)}(\xi(x))(x - m)^5 \leq 0.08$. Abbildung 7.6 zeigt die Graphen im Vergleich. Für eine bessere Darstellung der Funktion $\sin(x)$ ist diese auf dem Intervall $[-\pi, \pi]$ gezeichnet (strichliert). t_4 ist im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$ als Linie gezeichnet, ausserhalb punktiert.

Bemerkung 25. Möchte man eine Funktion über einem größeren Intervall approximieren, so ist es ratsam, mehrere Taylorpolynome mit verschiedenen Entwicklungspunkten aneinanderzuhängen.

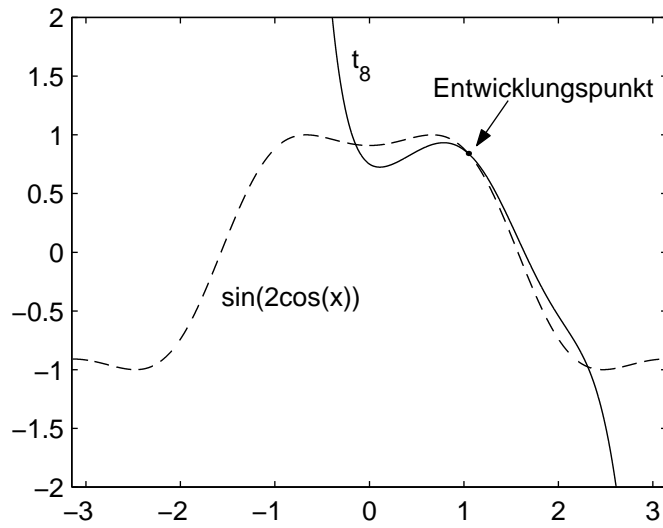


Abbildung 7.5: Taylorapproximation einer periodischen Funktion an einer beliebigen, nicht speziellen Stelle, z.B.: $m = \left(\frac{\pi}{3}, f\left(\frac{\pi}{3}\right)\right)$.

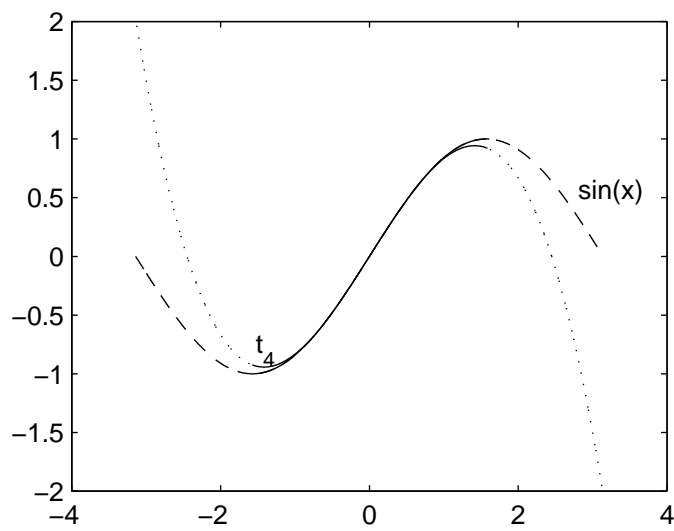


Abbildung 7.6: Taylorapproximation von $\sin(x)$.

8 Zusammenfassung

Die vorliegende Diplomarbeit beschäftigt sich mit der Interpolation und Approximation von eindimensionalen Funktionen auf \mathbb{R} , die anhand einer Tabelle gegeben sind. Die beschriebenen Methoden lassen sich nach dem Verwendungszweck unterscheiden.

Interpolationspolynome sind für eine Annäherung von gemessenen Werten nicht so gut geeignet, da Messfehler, speziell bei einer großen Anzahl von Stützstellen, unangenehmerweise das Polynom besonders an den Rändern zum „Schwingen“ bringen. Selbst Tschebyscheff-Polynome, die die Oszillation möglichst minimieren, sind noch Schwingungen unterworfen, auch wenn sie weniger stark ausfallen. Dafür sind die Interpolationspunkte nicht frei wählbar, was bei einer Erhöhung des Grades bedeutet, dass alle Stützstellen von neuem berechnet werden müssen. Für exakte Werte sind die Methoden von Lagrange, Newton und Tschebyscheff aber durchaus zu verwenden. Speziell sollte man sich über die Gestalt der Funktion Gedanken machen. Sind Asymptoten oder Polstellen zu vermuten oder sogar gegeben, so sind rationale Polynome besser zur Approximation geeignet.

Für eine Interpolation von gemessenen Werten ist die Spline-Interpolation vorzuziehen. Splines oszillieren nicht und sind für beliebige Stützstellen berechenbar und ohne unerwünschte „Nebenwirkungen“ verwendbar. Sind viele Stützstellen gegeben, so können bereits lineare Splines eine gute Approximation bieten. Darüberhinaus haben sie auch Eigenschaften, die Splines höherer Ordnung nicht mehr aufweisen können: Positivität, Monotonie und Konvexität im Sinne von Kapitel 3.1.1. Normalerweise entsprechen kubische Splines aber mehr der Vorstellung, da sie glatt sind. Für die verschiedenen Anwendungen gibt es unterschiedliche Ansätze, da kubische Splines nur bis auf 2 Werte eindeutig bestimmt sind. Es besteht die Möglichkeit die ersten, zweiten oder dritten Ableitungen in den Randpunkten frei zu wählen. Darüberhinaus können periodische Funktionen auch durch periodische Splines interpoliert werden. Diese sind eindeutig bestimmt. Weiters können mittels parametrischen Splines auch Kurven mit Doppelpunkten interpoliert werden.

Auch wenn durch Aneinanderhängen von Splines Kurven mit Knicken erzeugt werden können, so sind doch Bézier-Kurven in diesem Fall besser geeignet. Bei Bézier-

Polynomen sind Anfangs- und Endpunkt wählbar und der Kurvenverlauf ist durch beliebig viele Bèzier-Punkte bestimmbar. Für den Fall, dass bestimmte Punkte interpoliert werden sollen, ist es einfacher, Bèzier-Splines zu verwenden. Kubische Bèzier-Polynome interpolieren die Intervalle zwischen den gegebenen Punkten. Durch Wählen der Bèzier-Punkte lassen sich nicht nur die Punkte auf eine bestimmte Weise interpolieren, sondern es besteht auch die Möglichkeit, Knicke an den Knoten zu erzeugen. Ebenso können die Bèzier-Polynome auch so aneinandergesetzt werden, dass der Spline an den Knoten ein oder zwei mal stetig differenzierbar ist. Auch für Bèzier-Splines gibt es eine Variante für periodische oder geschlossene Kurven. Bèzier-Splines sind speziell für das Modellieren von Kurven brauchbar, da durch Versetzen von einzelnen Bèzier-Punkten, die Kurve einfach verändert werden kann.

Für die Interpolation von periodischen Funktionen sind Fourier-Polynome besonders geeignet. Durch die benötigten äquidistanten Stützstellen ist man zwar etwas eingeschränkt, aber da sie eine gute Approximation liefern, sind sie dennoch auch für durch Messung erhaltene Wertetabellen anwendbar. Darüberhinaus sind Fourier-Polynome wegen ihrer Eigenschaft nach der Methode der kleinsten Quadrate zu approximieren sehr von Vorteil.

Die Methode der kleinsten Quadrate stellt ein gutes Instrument dar, bis auf Parameter noch zu bestimmende Funktionen zu berechnen. Dabei bestehen auch die Möglichkeiten Punkte zu gewichten oder beliebige Punkte zu interpolieren und die restlichen zu approximieren, was zum Beispiel, wenn ein Punkt laut Theorie bereits feststeht, von Vorteil ist. Diese Methode überzeugt besonders bei geeigneter Wahl der Ansatzfunktion. Es stehen nicht nur Polynome, die Exponentialfunktion und rationale Polynome zum Approximieren zur Verfügung, sondern alle jene Funktionen $\varphi_1, \dots, \varphi_m$, die linear unabhängig sind.

Schließlich ist noch die Taylorentwicklung zu erwähnen, die für die Approximation von Funktionen gedacht ist, von denen man n Ableitungen in einem Punkt kennt. Sie ist in einem Intervall um den Entwicklungspunkt ausgezeichnet, wobei sich der Fehler aber rasch vergrößert, wenn man sich davon entfernt. Ausserhalb dieses Intervalls kann nicht mehr von Approximation der Funktion gesprochen werden. Ist $f^{(n+1)}$ beschränkt, so lässt sich auch der Fehler abschätzen, was sehr von Vorteil ist. Allerdings ist für etwas komplizierte Funktionen eine sinnvolle Abschätzung meist nicht einfach oder manchmal auch unmöglich.

Literaturverzeichnis

- [1] Becker, Jürgen, Wolfhart Haacke und Rudolf Nabert. *Numerische Mathematik für Ingenieure*. Stuttgart: B.G. Teubner, 1985.
- [2] Böhm, Wolfgang, Günther Gose und Jürgen Kahmann. *Methoden der numerischen Mathematik*. Braunschweig: Vieweg & Sohn, 1985.
- [3] Deuffhard, Peter und Andreas Hohmann. *Numerische Mathematik*. Berlin: Walter de Gruyter, 1991.
- [4] Engeln-Müllges, Gisela und Fritz Reutter. *Numerische Mathematik für Ingenieure*. Mannheim: Wissenschaftsverlag, 1988.
- [5] Liedl, Roman und Kristian Kuhnert. *Analysis in einer Variablen*. Mannheim: Wissenschaftsverlag, 1992.
- [6] Mathews, John H. *Numerical Methods for Mathematics, Science, and Engineering*. London: Prentice-Hall International, 1992.
- [7] Oevel, Walter. *Einführung in die Numerische Mathematik*. Heidelberg: Spektrum Akademischer Verlag, 1996.
- [8] Rutishauser, Heinz. *Vorlesung über Numerische Mathematik*. Basel: Birkhäuser Verlag, 1976.
- [9] Schwarz, Hans Rudolf. *Numerische Mathematik*. Stuttgart: B.G. Teubner, 1997.
- [10] Scheid, Francis. *Numerische Analysis*. London: McGraw-Hill Book Company, 1991.
- [11] Späth, Helmuth und Jörg Meier. *Eindimensionale Spline-Interpolation-Algorithmen*. München: R. Oldenbourg, 1990.
- [12] Weller, Friedrich. *Numerische Mathematik für Ingenieure und Naturwissenschaftler*. Braunschweig: Vieweg & Sohn, 1996.
- [13] Werner, Jochen. *Numerische Mathematik 1*. Wiesbaden: Vieweg & Sohn, 1992.

Lebenslauf

Name: Robert Morandell

Geburtsdatum: 9. November 1974

Geburtsort: Innsbruck, Österreich

Staatsbürgerschaft: Österreich

Ausbildung: 1981-1985 Volksschule
1985-1993 Gymnasium
1993-2000 Lehramtstudium der Mathematik, Anglistik und
Amerikanistik an der Leopold-Franzens-Universität Innsbruck
Seit 1997 Instrumental- und Gesangspädagogik-Studium
mit Hauptfach Gitarre am Tiroler Landeskonservatorium