

Die Konstruktion der Tschebyscheff-Approximierenden bei der Anpassung mit Exponentialsummen

D. BRAESS

*Westfälische Wilhelms-Universität, Institut für Numerische und Instrumentelle Mathematik,
44 Münster, Germany*

Communicated by E. W. Cheney

Received August 14, 1969

Für die Konstruktion bester Approximationen bei der Anpassung mit Exponentialsummen wird ein Algorithmus vom Remes-Typ entworfen. Dieser läßt sich in der einfachen Fassung nur für positive Exponentialsummen anwenden; im allgemeinen Fall muß man dafür sorgen, daß die Iterationsfolge nicht gegen lokale Minima konvergiert. Bei der Wahl der Ausgangsvektoren wird die Vorzeichenstruktur in der Familie der Exponentialsummen berücksichtigt.

EINLEITUNG

Für die Berechnung der besten Exponentialapproximation im Tschebyscheffschen Sinne wird ein numerisches Verfahren diskutiert. Es wird aufgebaut auf einem Basisalgorithmus, der sich als Newtonsches Verfahren oder als Gradientenverfahren im weiteren Sinne interpretieren läßt. Er wird so formuliert, daß die gemeinsamen Eigenschaften mit einem Algorithmus für die rationale Approximation von Werner [9] hervortreten. Wegen der Struktur der Exponentialsummen werden jedoch bestimmte Erweiterungen notwendig. Andererseits ergeben sich Vereinfachungen, durch die Übertragung von Ideen, wie sie vom Verfasser schon bei anderen Minimierungsaufgaben herangezogen wurden [3].

Für die Herleitung der Iterationsvorschrift wird das Approximationsproblem als Minimumproblem aufgefaßt. Zum Aufsuchen der Minimallösung bei positiven Exponentialsummen ist die Iteration uneingeschränkt geeignet. Dagegen ist die Konvergenz bei allgemeinen Exponentialsummen nicht immer gewährleistet. Wie nämlich in einer früheren Arbeit gezeigt wurde, gibt es bei dem Extremalproblem im allgemeinen *lokale Minima* [2]. Aus den Ergebnissen der Arbeit kann man schließen, daß die Faktoren α_n gemäß der Darstellung der Exponentialsummen

$$E(a, x) = \sum_{n=1}^N \alpha_n e^{\lambda_n x},$$
$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N,$$

schon in der Ausgangsnäherung die gleichen Vorzeichen haben sollten wie in der Minimallösung. Da die Vorzeichenverteilung für die Lösung nicht immer von vorn herein vollständig bekannt ist, wird häufig die Iteration mit mehreren Ausgangsnäherungen erforderlich. Daß nicht eine einzige Iterationsfolge ausreicht, ist andererseits verständlich, weil bei der Exponentialapproximation mehrere Minimallösungen existieren können [1]. Diese Arbeit schließt direkt an die Untersuchungen des Verfassers über die Eindeutigkeitsfragen [1] und die Diskussion der Struktur der Exponentialsummen [2] an. Die dort eingeführten Bezeichnungen und Formeln werden auch hier benutzt. Die Abschnitte sind weiter numeriert, und Zitate mit (1.1) bis (11.4) beziehen sich auf Formeln in jenen Arbeiten. Das gleiche gilt für die Sätze 1 bis 12.

12. DER BASISALGORITHMUS

Bei der Exponentialapproximation betrachtet man die reellen Funktionen, die sich in der Form

$$E(a, x) = \sum_{n=1}^l P_n(x) e^{\lambda_n x}, \quad P_n(x) = \sum_{\nu=0}^{M_n-1} \gamma_{n\nu} x^\nu,$$

mit

$$\sum_{n=1}^l M_n = k \leq N$$

darstellen lassen und bezeichnet die Menge mit V_N . Die Funktionen, bei denen die Polynome nur den Grad 0 haben, definieren die Teilmenge V_N^0 . Die Funktion $f(x)$ sei in einem reellen Intervall X definiert und dort stetig. Die beste Approximation minimiert das Funktional,

$$\Phi(a) = \|f - E[a]\| = \sup_{x \in X} |f(x) - E(a, x)|. \quad (12.1)$$

Durch die im folgenden beschriebene Iteration wird eine monotone Folge erzeugt

$$\Phi(a^0) \geq \Phi(a^1) \geq \Phi(a^2) \geq \dots$$

Für die numerische Behandlung seien die Parameter Vektoren $a = \{a_i\}$ im R^{2N} . Unabhängig davon, ob für die Summen die obige Darstellung oder die aus Abschnitt 9 verwendet wird, sind N Komponenten von a lineare Faktoren und die übrigen Komponenten bedeuten Frequenzen. Die Dimension der Tangentialmannigfaltigkeit ist $d = d(a) = N + l(a)$. Die Parameter

a_i seien so durchnummeriert, daß sich die ersten N auf die Faktoren beziehen und die folgenden $l(a)$ auf verschiedene Frequenzen λ_n . Dann bilden die Funktionen

$$u_i(x) = \frac{\partial}{\partial a_i} E(a, x), \quad i = 1, 2, \dots, d, \quad (12.2)$$

die Basis eines linearen Haarschen Systems. Zur Abkürzung schreiben wir auch mit einem beliebigen Vektor $\Delta a \in R^d$

$$(\Delta a, u(x)) = \sum_{i=1}^d \Delta a_i \cdot u_i(x). \quad (12.3)$$

Es sei $E[a] \in V_N$ nicht Minimallösung für $f(x)$ bez. V_N . Dann wird die Fehlerkurve $\epsilon(x) = f(x) - E(a, x)$ zunächst mit Funktionen der Form (12.2) approximiert, wobei die Zahlen Δa_i ($i = 1, 2, \dots, d$) freie Parameter sind. Diese lineare Tschebyscheff-Approximation wird nun im allgemeinen nicht über dem Intervall X durchgeführt. Um den Rechenaufwand möglichst niedrig zu halten, wird sie auf einer diskreten Teilmenge von X ausgeführt, sofern das nach der im folgenden beschriebenen Fallunterscheidung möglich ist. Es wird eine Punktmenge konstruiert, die aus mindestens $d + 1$ verschiedenen Punkten besteht. Außerdem wird die Anzahl der Punkte durch eine vorgegebene Zahl $j_0 \geq 2N + 1$ nach oben beschränkt. Ferner sei $\zeta \in (0, \frac{1}{2})$ eine kleine positive Zahl.

Fall 1(F 1).

Die Menge

$$\{x \in X : |\epsilon(x)| \leq \|\epsilon\| (1 - \zeta)\} \quad (12.4)$$

bestehe aus nicht mehr als j_0 Teilintervallen, und in jedem solchen Teilintervall sei das Maximum von $|\epsilon(x)|$ eindeutig bestimmt. X_1 sei die Menge der Maximalpunkte.

Ferner zerlege man das Grundintervall so in Teilintervalle (Vorzeichenintervalle), daß $\epsilon(x)$ in jedem dieser Intervalle das Vorzeichen nicht wechselt, in benachbarten Intervallen jedoch verschiedene Vorzeichen besitzt (vgl. [9]). Wenn mehr als j_0 Vorzeichenintervalle vorliegen, genügt eine Auswahl von j_0 Intervallen. Aus jedem greift man einen Punkt heraus. Günstig ist es, Extrempunkte zu verwenden. Die Menge dieser Punkte werde mit X_2 bezeichnet.

Schließlich sei X_3 eine beliebige Menge von $d + 1$ verschiedenen Punkten aus X . Nun können wir aus X_2 eine Teilmenge X_2' und aus X_3 eine Teilmenge X_3' so auswählen, daß die Vereinigung,

$$X_0 = X_1 \cup X_2' \cup X_3',$$

mindestens $d + 1$ und höchstens j_0 verschiedene Punkte enthält. (Für die Praxis heißt das: Wenn X_1 weniger als $d + 1$ Punkte enthält, ergänze man durch Punkte aus verschiedenen Vorzeichenintervallen und—sofern notwendig—durch beliebige weitere Punkte.)

Fall 2 (F2).

Die bei der Definition von X_1 unter F1 genannten Voraussetzungen seien nicht erfüllt, dann wird

$$X_0 = X$$

gesetzt.

Die Iterationsvorschrift lautet: Zu $a = a^0$ berechne man die nächste Näherung nach folgendem Schema:

1. Führe die Fallunterscheidung aus, und bestimme zugleich X_0 .

2. Suche die beste Tschebyscheffapproximation für die Fehlerkurve $\epsilon(x)$ über X_0 bez. des linearen Haarschen Systems mit $d = d(a)$ freien Parametern Δa_i :

$$\max_{x \in X_0} |\epsilon(x) - (\Delta a, u(x))| = \min!$$

Die Lösung des linearisierten Problems wird mit $e(x) = e^0(x)$ bezeichnet und mit $\Delta \epsilon$ die Verbesserung

$$\begin{aligned} \Delta \epsilon &= \|\epsilon^0\| - \max_{x \in X_0} |\epsilon^0(x) - e^0(x)| \\ &= \|f - E[a^0]\| - \max_{x \in X_0} |E(a^0, x) + e^0(x) - f(x)|. \end{aligned} \quad (12.5)$$

3. Setze $t = 1$ und bestimme durch sukzessives Halbieren von t einen Wert $t = t^0 > 0$, so daß für das in (12.1) definierte Funktional,

$$\Phi(a + t\Delta a) \leq \Phi(a) - \frac{1}{3}t\Delta \epsilon, \quad (12.6)$$

gilt. Dann setze die Iteration mit $a = a^{0+1} = a^0 + t\Delta a$ fort. Beendet wird die Iteration, sobald sich wegen der Rundungen keine Verbesserung mehr ergibt.

Zusatzregel. Falls F1 vorliegt und im dritten Teilschritt die Relation (12.6) nur mit $t < \zeta$ erfüllt werden kann, wiederhole den Iterationsschritt unter Anwendung der Vorschrift zu (F2), d.h. mit $X_0 = X$.

Im Gegensatz zu dem Algorithmus von Werner lösen wir bei jedem normalen Schritt kein lineares Gleichungssystem¹, sondern ein diskretes

¹ Für rationale Funktionen reduziert sich diese Aufgabe auf eine einfache Interpolationsaufgabe.

lineares Tschebyscheffproblem. Letzteres ist nur wenig aufwendiger als die Lösung eines linearen Gleichungssystems [5, 8]. Vorteilhaft ist unsere Methode vor allem, wenn in einem Vorzeichenintervall mehrere Extrema von $|\epsilon(x)|$ liegen oder wenn mehr als $d(a) + 1$ Extrema mit ungefähr gleichem Betrag auftreten. Ferner sei betont, daß die Iteration weitgehend formuliert wurde, ohne daß auf die spezielle Struktur der Exponentialsummen Bezug genommen wurde. Dies konnte letzten Endes deswegen geschehen, weil im dritten Teilschritt nicht nur $\Phi(a + t\Delta a) \leq \Phi(a)$, sondern die schärfere Relation (12.6) gefordert wurde vgl. [3]).

13. KONVERGENZSATZ FÜR DAS NEWTONSCHE VERFAHREN

Nach den Ausführungen des Abschnitts 11 können wir nicht erwarten, daß die Iteration immer zur Minimallösung führt. Vielmehr gilt der schwächere

SATZ 13. *Bei der Iteration konvergiere eine Teilfolge gegen*

$$E[a^*] \in V_N^0 - V_{N-1}^0.$$

Dann konvergiert die ganze Folge, und $E[a^]$ ist Minimallösung bez. V_N .*

Dem Beweis schicken wir zwei Hilfssätze voraus, wobei wir zunächst zeigen, wann die Iteration schrittweise bessere Approximation erzeugt.

Hilfssatz 6. Die Länge der Alternante von $E[a]$ sei kleiner als $d(a) + 1$. Dann existiert eine positive Zahl t_0 , so daß die Relation (12.6) für alle $t \in (0, t_0)$ erfüllt ist.

Beweis. Zunächst zeigen wir die Behauptung für (F2). Wegen der Voraussetzung über die Alternante liefert das linearisierte Problem ein positives $\Delta\epsilon$. Wegen der Differenzierbarkeit der Exponentialsummen gilt für hinreichend kleine t

$$\|E[a + t\Delta a] - E[a] - t(\Delta a, \nabla E(a, x))\| \leq \frac{t}{6} \Delta\epsilon. \quad (13.1)$$

Mit Hilfe der Dreiecksungleichung folgt daraus:

$$\begin{aligned} \|E[a + t\Delta a] - f\| &\leq \|E[a] + t(\Delta a, \nabla E) - f\| + \frac{t}{6} \Delta\epsilon \\ &\leq t \|E[a] + (\Delta a, \nabla E) - f\| + (1 - t) \|E[a] - f\| + \frac{t}{6} \Delta\epsilon \\ &\leq \|\epsilon\| - \frac{t}{3} \Delta\epsilon. \end{aligned}$$

Für F1 sei der Beweis nur skizziert (da die Beweisidee die gleiche ist wie beim Hilfssatz 5.1 in [9]). Für die Punkte $x_i \in X_0$, für die $\epsilon(x_i) > \|\epsilon\| - \frac{1}{8}\Delta\epsilon$ ist, gilt $e^\rho(x_i) \leq -\frac{5}{8}\Delta\epsilon$. Wegen der Stetigkeit läßt sich um jeden dieser Punkte ein Intervall abgrenzen, so daß in diesem

$$e^\rho(x) \leq -\frac{3}{8}\Delta\epsilon$$

und so nach (13.1) für hinreichend kleine t in diesen Intervallen auch

$$|E(a + t\Delta a, x) - f(x)| \leq \|\epsilon\| - \frac{1}{8}\Delta\epsilon$$

gilt. Das gleiche läßt sich ebenso für Intervallumgebungen der Punkte mit $\epsilon(x_i) < -\|\epsilon\| + \frac{1}{8}\Delta\epsilon$ erreichen.

In den übrigen Punkten von X ist für $t = 0$: $|(E(a + t\Delta a) - f)| < \|\epsilon\|$ und diese Relation ist auch für hinreichend kleine t erfüllt (vgl. den Beweis des zitierten Hilfssatzes).

Hilfssatz 7. Eine Folge $E[a^\rho]$ strebe gleichmäßig gegen

$$E[\tilde{a}] \in V_N^0 - V_{N-1}^0.$$

Dann streben die laut F2 berechneten Lösungen des linearisierten Problems $e^\rho(x)$ gleichmäßig gegen die zu $E[\tilde{a}]$ entsprechend konstruierte Funktion $\tilde{e}[x]$.

Beweis. Abgesehen von endlich vielen Indizes gilt $E[a^\rho] \in V_N^0 - V_{N-1}^0$, und die Funktionen $e^\rho(x)$ lassen sich in der Form

$$e^\rho(x) = \sum_{n=1}^N e^{\lambda_n^\rho x} [\beta_n^\rho + \gamma_n^\rho x], \quad (13.2)$$

darstellen, wobei die Frequenzen λ_n^ρ zum Spektrum von $E[a^\rho]$ gehören.

Die Folge (13.2) ist in V_{2N} enthalten, und wegen $\|e^\rho - \epsilon^\rho\| \leq \|\epsilon^\rho\|$ gilt $\|e^\rho\| \leq 2\|\epsilon^\rho\| \leq 2\|f\| + 2\|E[a^\rho]\|$. Da die e^ρ in einer beschränkten Teilmenge von V_{2N} liegen und die Frequenzen beschränkt sind, konvergiert eine Teilfolge [2, 7]. Aus der Stetigkeit der Norm und der Dreiecksrelation folgt nun für diese Teilfolge

$$\lim_{\rho \rightarrow \infty} \|e^\rho + E[a^\rho] - f\| \leq \|\tilde{e} + E[\tilde{a}] - f\|.$$

Wegen der Eindeutigkeit der linearen Approximation ist

$$\lim_{\rho \rightarrow \infty} \|e^\rho + E[a^\rho] - f\| = \|\tilde{e} + E[\tilde{a}] - f\|.$$

In der Teilfolge konvergiert e^ρ gegen \tilde{e} . Da dies auch für jede andere Teilfolge gilt, ist die Behauptung richtig.

Der indirekte Beweis des Konvergenzsatzes verläuft nun ähnlich wie bei Werner [9].

Beweis zu Satz 13. Die Grenzfunktion $E[a^*]$ sei nicht Minimallösung. Dann ist die zu $E[a^*]$ gemäß F2 berechnete Änderung $\Delta\epsilon^* > 0$. Aus Hilfsatz 7 folgt unmittelbar, daß für die Teilfolge—abgesehen von endlich vielen Ausnahmen— $\Delta\epsilon^p \geq \frac{1}{2}\Delta\epsilon^*$ gilt, sofern $\Delta\epsilon^p$ gemäß F2 ermittelt wird. Diese Relation gilt dann auch für F1, da in letzterem Fall die Approximation nur auf einer Teilmenge betrachtet wird. Wir unterscheiden nun, ob in der genannten Teilfolge endlich oft F2 angewandt wird oder nicht.

Zunächst sei angenommen, es gebe eine gegen $E[a^*]$ konvergente Teilfolge, bei der immer gemäß F2 verfahren wird. Wegen der Konvergenz der $E[a^p]$ und der Δa^p läßt sich nun ein $t_1 > 0$ angeben, so daß mit Ausnahme endlich vieler Glieder die Relation (13.1) für $0 \leq t \leq t_1$ richtig ist und so $t^p \geq \frac{1}{2}t_1$ wird:

$$\Phi(a^p) - \Phi(a^{p+1}) \geq \frac{1}{3} t^p \cdot \Delta\epsilon^p \geq \frac{1}{3} \cdot \frac{t_1}{2} \cdot \frac{\Delta\epsilon^*}{2}.$$

Das steht jedoch im Widerspruch zur Tatsache, daß $\Phi(a^p)$ monoton gegen $\Phi(a^*)$ strebt.

Es sei nun angenommen, es gebe eine gegen $E[a^*]$ konvergente Teilfolge, bei der gemäß F1 verfahren wird. Auf Grund der Zusatzregel ist

$$\Phi(a^p) - \Phi(a^{p+1}) \geq \frac{1}{3} \zeta \cdot \Delta\epsilon^p \geq \frac{1}{3} \zeta \cdot \frac{\Delta\epsilon^*}{2}.$$

Das ist wie oben ein Widerspruch zur Monotonie des Funktionals.

14. AUSBAU DES ALGORITHMUS, GRENZEN DES VERFAHREN

Betrachten wir zunächst die Konstruktion der besten positiven Exponentialsumme zu einer gegebenen Funktion $f(x) \in C(X)$. Um die Voraussetzungen des Konvergenzsatzes immer erfüllen zu können, gehen wir induktiv vor.

Die Minimallösung \hat{E} bez. V_{N-1}^+ sei bekannt². Wir brauchen uns nur mit dem Fall zu befassen, daß \hat{E} nicht schon Minimallösung bez. V_N^+ ist. Nach

² $V_0 = V_0^+$ enthält als einziges Element die identisch verschwindende Exponentialsumme $E(a, x) = 0$.

dem Kriterium in Satz 4 können wir also annehmen, daß \hat{E} eine positive Alternante der Länge $2N - 1$ besitzt. Dann liegen alle Exponentialsummen aus V_N , die f besser als \hat{E} approximieren, in V_N^+ .

Für die Iterationsfolge benutzen wir die Parameter in der Darstellung (14.1) und setzen

$$\left. \begin{aligned} \alpha_n^0 &= \hat{\alpha}_n, & \lambda_n^0 &= \hat{\lambda}_n \\ \alpha_N^0 &= 0, & \lambda_N^0 &\neq \hat{\lambda}_n \end{aligned} \right\} \quad (n = 1, 2, \dots, N - 1). \quad (14.1)$$

λ_N^0 ist eine beliebige Zahl, die nicht zum Spektrum von $E[\hat{a}]$ gehört. Die Dimension der Tangentialmannigfaltigkeit ist $d(a^0) = 2N - 1$, und so bringt der erste Schritt eine Verbesserung mit sich.

Die ganze Folge verläuft für $\rho \geq 1$ in der kompakten Menge

$$\{E[a] : E[a] \in V_N^+ - V_{N-1}^+, \|E[a] - f\| \leq M_1\}, \quad (14.2)$$

$$\text{mit } M_1 = \|E[a^1] - f\| < \inf_{E[a] \in V_{N-1}} \|E[a] - f\|.$$

Da die Folge also in V_N^0 enthalten ist, kann man durchweg mit der Darstellung (8.1) rechnen. Es gibt eine konvergente Teilfolge. Wegen der Abgeschlossenheit von V_N^+ liegt auch die Grenzfunktion der Teilfolge in der Menge (14.2) und erfüllt die Voraussetzungen des Konvergenzsatzes.

Es lassen sich also nacheinander die besten Approximationen bez. V_1^+ , V_2^+ , V_3^+ , ... bestimmen. Dieses induktive Vorgehen ist in der Praxis (vor allem bei der Analyse von Meßdaten) häufig auch vom Problem her zweckmäßig. Denn oft ist es schwer, von vorn herein zu erkennen, wie viele Frequenzen man aus vorgegebenen Daten auflösen kann. Für die rationale Approximation hat Werner auf die Vorteile des Fortschreitens in der Ordnung hingewiesen.

Bei der Approximation ohne die Vorzeichenbeschränkung sind die Voraussetzungen des Konvergenzsatzes nicht immer erfüllt, wenn die beste Exponentialsumme sowohl positive als auch negative Faktoren enthält. Dies zeigen wir, indem wir eine Arbeitshypothese einführen, die später durch eine schwächere Forderung ersetzt wird.

Es sei $a^{\rho+1} = a^\rho + t^\rho \Delta a^\rho$. Durch die Relation (12.6) wird erreicht, daß dann

$$\Phi(a^\rho + t^\rho \Delta a^\rho) < \Phi(a^\rho) \quad (14.3)$$

für $t = t^\rho$ gilt. Nach dem Hilfssatz 6 ist (14.3) auch für kleine positive t richtig. Wie man aus dem Beweis von Satz 13 erkennt, ist es prinzipiell möglich, t^ρ so zu wählen, daß (12.6) für alle $t \in [0, t^\rho]$ erfüllt ist. Offensichtlich ändert diese "starke Monotonieforderung" nicht die Konvergenzsätze.

Bei starker Monotonie enthält die Menge,

$$\{E[a] : E[a] \in V_N, \|f - E[a]\| \leq \Phi(a^1)\}, \quad (14.4)$$

nicht nur Elemente $E[a^p]$ der Folge, sondern auch einen verbindenden Bogen. Die Folge läuft also innerhalb einer Zusammenhangskomponente von (14.2). Wenn $E[a^1]$ eine bessere Approximation als die Funktionen aus V_{N-1} ist, dann liegt die Folge innerhalb einer Zusammenhangskomponente von $V_N - V_{N-1}$, nach Satz 10 also innerhalb einer Vorzeichenklasse. Also besitzen alle Exponentialsummen die gleiche Vorzeichenverteilung wie $E[a^1]$. Wie beim Beweis der Existenz mehrerer lokaler Minima deutlich wurde, gibt es aber meistens Exponentialsummen mit verschiedenen Vorzeichenverteilungen, die bessere Approximationen als die Summen aus V_{N-1} sind. Also gibt es auch Folgen mit verschiedenen Grenzfunktionen.

Betrachten wir den Fall, daß wir formal genauso wie bei der Konstruktion der besten positiven Summen vorgehen und die Iteration mit der Minimallösung bez. V_{N-1} beginnen, wobei eine neue Frequenz mit verschwindendem Faktor hinzugefügt wird. Wenn es nach den obigen Überlegungen überhaupt sinnvoll ist, an diesem Schema festzuhalten, muß das Verfahren mindestens eine Forderung erfüllen, die durch folgende Erweiterung des Algorithmus charakterisiert wird.

Man startet die Iteration wie beschrieben. Bei der Iteration sorgt man dafür, daß sich die Vorzeichenverteilung nicht ändert. Im allgemeinen ist dies durch die Forderung (12.6) von selbst erfüllt, andernfalls verhindert man die Änderung durch Verkleinern des Dämpfungsparameters $t = t^p$. Sobald man feststellt, daß zwei oder mehrere Frequenzen zusammenlaufen, wiederholt man die Iteration mit einem anderen Anfangswert für die zusätzliche Frequenz. Dabei ist für λ_N eine solche Zahl einzusetzen, daß man eine Folge mit einer anderen Vorzeichenverteilung erhält (s. Abschnitt 11). Die Zahl der Möglichkeiten ist durch (11.4) gegeben. Es ist nicht gesichert, daß das Verfahren immer zum Ziel führt, denn es ist für $N > 3$ durchaus fraglich, ob nicht mehr lokale Minima existieren, als in Abschnitt 11 genannt wurden. Bei den Beispielen, die wir durchgerechnet haben, hat das Verfahren immer die Minimallösung geliefert.

15. BEISPIELE

Die behandelten Beispiele sollen die Ergebnisse der Arbeit beleuchten, (obgleich sie für Anwendungen uninteressant sind). Als Beispiel für ein in der Praxis auftretendes Problem wird die Anpassung von Meßdaten aus einem Reaktorexperiment beschrieben.

Um den Rechenaufwand bei der Durchführung der Iteration in Grenzen zu halten, wurden die Funktionen diskretisiert. Der Abstand der Stützstellen wurde 0.01 oder kleiner gewählt.

In V_1 wurde die Iteration mit der besten konstanten Funktion

$$\frac{1}{2} \max_{x \in X} f(x) + \frac{1}{2} \min_{x \in X} f(x),$$

gestartet. Für die Konstruktion der Minimallösung bez. V_N gingen wir, wie beschrieben, von der besten Approximation $E[\hat{a}]$, bez. V_{N-1} aus. Die $l(\hat{a})$ Frequenzen teilen die reelle Zahlengerade in $l(\hat{a}) + 1$ Intervalle. Die zusätzlichen Frequenzen wurden nacheinander in diese Intervalle gelegt und die Iteration gestartet. In dem Normalfall $k = N - 1$ wurden die Werte

$$\frac{1}{2}(\hat{\lambda}_{i+1} + \hat{\lambda}_i), \quad \text{für } i = 1, 2, \dots, l(\hat{a}) - 1,$$

sowie

$$\hat{\lambda}_1 - \bar{\lambda}, \quad \hat{\lambda}_{l(\hat{a})} + \bar{\lambda} \quad \text{mit} \quad \bar{\lambda} = \frac{3}{\max_{x \in X} (x) - \min_{x \in X} (x)}$$

gewählt. Sobald die Iteration bei einer Folge auf eine Exponentialsumme mit einer Alternante der Länge $2N + 1$ führt, wird die Prozedur abgebrochen.

Die Zahlenangaben für die Parameter sind durchweg gerundete Größen.

A. Approximation von $f(x) = 1/(1+x)$ in $[0, 1]$

Die beste Approximation für $1/(1+x)$ bez. V_N ist für alle N eine positive Exponentialsumme.

Beweis. Sei $E(a^*, x) \in V_N$ Minimallösung. Die mit $1+x$ multiplizierte Fehlerkurve,

$$h(x) = 1 - (1+x) \cdot E(a^*, x),$$

ist in $V_{k^*+l^*+1}$ enthalten, hat also höchstens $k^* + l^*$ Nullstellen. Da $\epsilon(x)$ nach dem Alternantenkriterium mindestens $N + l^*$ Nullstellen hat, ist $k^* = N$. Die Behauptung ist für V_1 offensichtlich richtig und sei für $N - 1$ schon bewiesen. Damit $h(x) \in V_{2N+1}$ $2N$ Nullstellen hat, müssen alle Frequenzen negativ sein. Da $\lim_{x \rightarrow \infty} h(x) = 1$ ist, ist nach einem schon mehrfach benutzten Schluß die Alternante positiv, und laut Satz 4 ist die Minimallösung bez. V_{N+1} eine positive Exponentialsumme.

Durch eine grobe Abschätzung läßt sich außerdem zeigen, daß für große N nicht alle Frequenzen beschränkt bleiben. Sei $E(x) = \sum \alpha_n e^{\lambda_n x}$ Minimallösung bez. V_N , und es sei Γ eine Schranke für die Frequenzen,

$$|\lambda_n| \leq \Gamma, \quad n = 1, 2, \dots, N.$$

Da alle Frequenzen negativ sind, haben in den Ableitungen $E^{(m)}(x)$ alle Terme das gleiche Vorzeichen, und es gilt

$$|E^{(m)}(x)| \leq \Gamma |E^{(m-1)}(x)|.$$

Für die Ableitungen der Hilfsfunktion $h(x)$

$$h^{(m)}(x) = -(1+x)E^{(m)}(x) - mE^{(m-1)}(x)$$

folgt deshalb aus der Existenz einer Nullstelle ξ

$$m |E^{(m-1)}(\xi)| = |(1+\xi)E^{(m)}(\xi)| \leq |1+\xi| \cdot \Gamma |E^{(m-1)}(\xi)|$$

$$\Gamma \geq \frac{m}{|1+\xi|}.$$

Da h zwischen den Alternantenpunkten insgesamt $2N$ Nullstellen besitzt, hat $h^{(2N-1)}$ im Intervall $[0, 1]$ eine Nullstelle. Daraus folgt

$$\Gamma \geq \frac{2N-1}{2} = N - \frac{1}{2}.$$

An den in der Tabelle aufgeführten Minimallösungen kann man die Trennungseigenschaft der Frequenzen gut ablesen (Tab. I).

B. Approximation von $f(x) = \sqrt{x}$ in $[0, 1]$

Die Approximation von \sqrt{x} liefert für $N = 1$ eine positive Summe, führt aber schon bei $N = 2$ aus V_N^+ heraus. Bis $N = 4$ wurde das im letzten Abschnitt beschriebene Verfahren erfolgreich erprobt.

In V_2 kommt man zu einem lokalen Minimum, wenn man beim Start der Iteration $\lambda_2 > 1.751$ wählt, und mit $\lambda_2 < 1.75$ läuft die Iteration zur Minimallösung. Es zeigt sich analog, daß man in V_3 mit $\lambda_3 < 0.8$ starten muß, und in V_4 mit $\lambda_4 < 0.6$.

Die Alternantenpunkte liegen in der Umgebung des Nullpunktes dichter als im übrigen Intervall, so daß man bei der Diskretisierung für kleine x das Raster feiner wählen muß.

C. Approximation von $f(x) = \operatorname{tg} x$ in $[-\pi/4, +\pi/4]$

Es zeigt sich, daß die besten Approximationen für $N \leq 5$ eindeutig und deshalb antisymmetrisch sind. Der Grad ist dann eine gerade Zahl. Obwohl für $N = 2$ und $N = 4$ mehr als $2N + 1$ Alternantenpunkte vorliegen, konvergiert die Iteration gut.

TABELLE I

N	Minimallösungen für $f(x)$ bez. V_N	Länge und Vorzeichen der Alternante	Minimal abstand
$f(x) = 1/(1+x)$ in $[0, 1]$			
1	$0.977 \cdot e^{-0.715x}$	3 +	$2.13 \cdot 10^{-2}$
2	$0.714 \cdot e^{-0.407x} + 0.286 \cdot e^{-2.443x}$	5 +	$2.08 \cdot 10^{-4}$
3	$0.0459 \cdot e^{-4.507x} + 0.349 \cdot e^{-1.601x}$ $+ 0.560 \cdot e^{-0.287x}$	7 +	$1.83 \cdot 10^{-6}$
4	$0.00572 \cdot e^{-6.740x} + 0.1074 \cdot e^{-3.187x}$ $+ 0.4264 \cdot e^{-1.212x} + 0.4605 \cdot e^{-0.223x}$	9 +	$1.54 \cdot 10^{-8}$
5	$6.16 \cdot 10^{-4} \cdot e^{-9.078x} + 0.0211 \cdot e^{-4.996x} + 0.159 \cdot$ $e^{-2.506x} + 0.428 \cdot e^{-0.978x} + 0.391 \cdot e^{-0.182x}$	11 +	$1.26 \cdot 10^{-10}$
$f(x) = \operatorname{tg} x$ in $[-\pi/4, +\pi/4]$			
1	0	2 +	1
2.3	$0.607 \sinh 1.619x$	6 +	$2.96 \cdot 10^{-3}$
4.5	$0.794 \sinh 1.232x + 0.0425 \sinh 4.972x$	10 +	$7.01 \cdot 10^{-6}$
$f(x) = \sqrt{x}$ in $[0, 1]$			
1	$0.21 \cdot e^{1.751x}$	3 -	0.21
2	$-0.431 \cdot e^{-9.47x} + 0.461 \cdot e^{0.804x}$	5 -	0.030
3	$-0.0982 \cdot e^{-142.8x} - 0.448 \cdot e^{-4.72x} + 0.555 \cdot e^{0.601x}$	7 -	0.0081
4	$-0.0348 \cdot e^{-1128x} - 0.111 \cdot e^{-45.7x} - 0.474 \cdot e^{-3.38x}$ $+ 0.622 \cdot e^{0.493x}$		
Reaktordaten			
1	$814.7 \cdot e^{-0.229x}$	3 +	45.2
2	$200.5 \cdot e^{-1.36x} + 654.6 \cdot e^{-0.162x}$	5 +	4.87
3	$152.8 \cdot e^{-2.327x} + 556.7 \cdot e^{-0.216x} + 151.4 \cdot e^{-0.0868x}$	7 -	0.944

D. Anpassung von Reaktordaten

Die Werte der untenstehenden Tabelle sind Meßwerte, die sich beim Abschalten eines Reaktors ergaben. Es sollte festgestellt werden, wie viele Frequenzen mit nur positiven Faktoren herausgezogen werden können. Man erhält für $N \leq 3$ positive Summen. Die beste Approximation bez. V_3 hat eine negative Alternante der Länge 7, so daß bessere Approximationen notwendigerweise auch negative Faktoren enthalten.

Das Iterationsverfahren wurde außerdem noch bei verschiedenen Funktionen, z.B. bei der Gauß'schen Fehlerfunktion, der Gammafunktion sowie

TABELLE II

x	$f(x)$	x	$f(x)$
0	860	9	148
1	603	10	128
2	491	11	109
3	407	12	96
4	341	13	82
5	288	14	71
6	242	15	62
7	205		
8	174		

bei anderen trigonometrischen und rationalen Funktionen bis $N = 6$ untersucht. Im großen und ganzen zeigt sich folgendes Bild.

Für $N = 1$ und $N = 2$ ergeben sich keinerlei numerische Schwierigkeiten. Für $N \geq 3$ ist eine Modifikation zur Konvergenzbeschleunigung zweckmäßig, die als allgemeine Methode in [4] beschrieben wird. Für $N \geq 5$ zeigen sich aber meistens größere Schwierigkeiten. Das gilt auch für positive Exponentialsummen. Häufig werden die Frequenzen sehr groß (wie z.B. bei $f(x) = \sqrt{x}$), oder einzelne Amplituden tragen zur Summe nur wenig bei und lassen sich deshalb schlecht auflösen (z.B. ist bei $f(x) = 1/(1+x)$ und $N = 5$ in der Minimallösung der erste Term im ganzen Intervall kleiner als $1/1000$ der Summe).

LITERATUR

1. D. BRAESS, Approximation mit Exponentialsummen, *Computing* **2** (1967), 309–321.
2. D. BRAESS, Über die Vorzeichenstruktur der Exponentialsummen, *J. Approximation Theory* **3** (1970), 101–113.
3. D. BRAESS, Über Dämpfung bei Minimalisierungsverfahren, *Computing* **1** (1966), 264–272.
4. D. BRAESS, Eine Möglichkeit zur Konvergenzbeschleunigung bei Iterationsverfahren für bestimmte nichtlineare Probleme, *Numer. Math.* **14** (1970), 468–475.
5. G. MEINARDUS, "Approximation von Funktionen und ihre numerische Behandlung," Berlin/Göttingen/Heidelberg/New York, 1964.
6. J. R. RICE, Chebyshev Approximation by Exponentials, *SIAM J. Appl. Math.* **10** (1962), 149–161.
7. E. SCHMIDT, "Normalität und Stetigkeit bei der Tschebyscheff-Approximation mit Exponentialsummen," Dissertation, Münster, 1968.
8. E. STIEFEL, Über diskrete und lineare Tschebyscheff-Approximation, *Numer. Math.* **1** (1959), 1–28.
9. H. WERNER, Die konstruktive Ermittlung der Tschebyscheff-Approximierenden im Bereich der rationalen Funktionen, *Arch. Rational Mech. Anal.* **11** (1962), 368–384.