§4 Lokale Minima

Es sei hier wieder $I \subseteq \mathbb{R}$ ein kompaktes Intervall und $f \in C(I)$; weiter gelte $N \in \mathbb{N}$.

Definition 4.1

Eine Funktion $E_o \in V_N$ $(E_o \in V_N^o)$ heißt ein lokales Minimum für f in V_N (V_N^o), wenn es eine Umgebung U von E_o in V_N (V_N^o) gibt, so daß für alle $E \in U$ erfüllt ist:

$$||f-E_o||_I \leqq ||f-E||_I$$

Es wird gezeigt, daß im allgemeinen bei der Approximation von f durch Exponentialsummen mit der Existenz lokaler Minima nach Definition 4.1 zu rechnen ist; dies ist von großer Bedeutung für die Konstruktion einer Minimallösung durch Iterationsverfahren.
Einfache Verhältnisse liegen in V_N^o vor:

Satz 4.1 (Werner, [21])

Jedes lokale Minimum für f in V_N^o ist die beste Approximation an f bzgl. V_N^o auf I.
Beweis:

Es sei $E(a,x) = \sum\limits_{i=1}^{N} a_i e^{t_i x} \in V_N^o$ mit $\mathrm{grd}(E(a))=K$ ein lokales Minimum für f. Wählt man $\varepsilon>o$ hinreichend klein, dann ist E(a) nach Voraussetzung beste Approximation an f auf I bzgl. $V:=\{\sum\limits_{i=1}^{N} b_i e^{s_i x} \in V_N^o \mid |a_i-b_i|<\varepsilon>|s_i-t_i|$ für $1 \leqq i \leqq K$,

$$|b_i|<\varepsilon>|s_i| \text{ für } K<i\leqq N\};$$

V erfüllt die Voraussetzungen von Hilfssatz 3.2 und der Gradientenraum von E(a) als Element von V hat die Dimension Dimension N+K. Damit besitzt f-E(a) in I eine Alternante der Länge N+K+1 und nach Korollar 3.1 ist E(a) die (eindeutig bestimmte) beste Approximation an f bzgl. V_N^o.

Aus Satz 4.1 folgt mit Korollar 3.1 unmittelbar

Korollar 4.1

Ist $E(a) \in V_N^o$ ein lokales Minimum für f in V_N, dann ist E(a) die beste Approximation für f bzgl. V_N.

Damit ist gezeigt, daß lokale Minima für f in V_N, die nicht zugleich Minimallösungen sind, nur in $V_N - V_N^o$ enthalten sein können.

Hilfssatz 4.1 (Braess, [2])

Es sei E(a) mit grd(E(a))=K auf I Minimallösung für f bzgl. V_{N-1}^o, nicht aber bzgl. V_N^o. Die Menge

$$\{t_i \mid t_i \in \mathbb{R},\ K+1 \leqq i \leqq N,\ t_i < t_{i+1}\}$$

sei so gegeben, daß sie mit dem Spektrum von E(a) kein Element gemeinsam hat.

Behauptung:

In jeder Umgebung von E(a) in V_N gibt es ein $E(b) \in V_K^o$ und es gibt $a_i \in \mathbb{R}$, $K+1 \leqq i \leqq N$, so daß für $E(\bar{a},x) = E(b,x) + \sum_{i=K+1}^{N} a_i e^{t_i x}$

gilt: $\quad ||f-E(a)||_I > ||f-E(\bar{a})||_I$.

Beweis: (Man vergleiche hierzu den Beweis von Satz 8 und 9 in [11] und Satz 2.1o in [24])

Man betrachte die Menge

$$V := \{E(d) \in V_N \mid E(d,x) = \sum_{i=1}^{N} d_i e^{s_i x},\ (d_1,\dots,d_N,s_1,\dots,s_K) \in \mathbb{R}^{N+K},$$

$$s_i = t_i \text{ für } K < i \leqq N\}.$$

Die Elemente von V werden mit v(p) bezeichnet, wobei der Parametervektor $p \in \mathbb{R}^{N+K}$ dem Element $v(p) \in V$ zugeordnet ist durch:

$$p = (d_1,\dots,d_N,s_1,\dots s_K) \in \mathbb{R}^{N+K} \longleftrightarrow$$

$$\longleftrightarrow \quad v(p,x) = \sum_{i=1}^{K} d_i e^{s_i x} + \sum_{i=K+1}^{N} d_i e^{t_i x} \in V$$

E(a) ist ein Element von V und der dazugehörige Parametervektor sei wieder mit a bezeichnet: $\quad E(a) = v(a) \in V$.

Der Gradientenraum W von v(a) hat die Dimension N+K; nach Hilfssatz 3.2 ist daher v(a) nicht Minimallösung für f bzgl. V, da F=f-v(a) in diesem Fall, im Widerspruch zur Voraussetzung über E(a), eine Alternante der Länge N+K+1 hätte. Die Punkte $x_i \in I$, $1 \leqq i \leqq N+K$, seien Alternantenpunkte einer Alternante der Länge N+K von F in I, die nach Voraussetzung existiert; D sei die Menge der Extremalpunkte von F in I:

$$D=\{x\in I \mid |F(x)|=||F||_I\}.$$

Aus $F\in C(I)$ folgt die Existenz von N+K-1 Nullstellen $z_1<z_2<\ldots<z_{N+K-1}$ von F in I, so daß F auf $D\cap[z_i,z_{i+1}]$ für $o\leqq i\leqq N+K-1$ das Vorzeichen nicht wechselt; dabei sei $z_o=q_1$, $z_{N+K}=q_2$ mit $I=[q_1,q_2]$.

Da W die Haarsche Bedingung erfüllt, gibt es mit $d_1=\min D$ (D ist kompakt) ein $b\in\mathbb{R}^{N+K}$, so daß gilt:

$$(b,\operatorname{grad}(v(a,d_1))) = \operatorname{sign}(F(d_1))$$
$$(b,\operatorname{grad}(v(a,z_i))) = o \qquad 1\leqq i\leqq N+K-1$$

Nach Bemerkung 3.2 hat (b,grad(v(a))) in I genau die N+K-1 Nullstellen z_i, $1\leqq i\leqq N+K-1$, und wechselt dort sein Vorzeichen. Daraus folgt wie im Hilfssatz 3.2 :

$$\min_{x\in D}(F(x)\cdot(b,\operatorname{grad}(v(a,x)))\ ) >o \tag{4.1}$$

Es wird nun benutzt, daß v(a) Fréchet-differenzierbar ist:

$$||v(a+r)-v(a)-(r,\operatorname{grad}(v(a)))||_I = o(||r||),$$

wobei $r\in\mathbb{R}^{N+K}$ und $||r||$ die euklidische Norm von r ist.

Da D kompakt ist, gibt es eine offene Umgebung U von D in I, so daß wegen (4.1) für ein c>o, $c\in\mathbb{R}$, gilt:

$$F(x)(b,\operatorname{grad}(v(a,x)))\geqq 2c>o \qquad x\in U .$$

Für $x\in U$ folgt mit $t\in\mathbb{R}$, t>o

$$F(x)(v(a+tb,x)-v(a,x))=$$
$$=F(x)t(b,\operatorname{grad}(va,x))) + F(x)(v(a+tb,x)-v(a,x)-(tb,\operatorname{grad}(v(a,x))))$$
$$\geqq 2ct-||F||_I||v(a+tb)-v(a)-(tb,\operatorname{grad}(v(a)))||_I$$

Daher gibt es ein $T_1>o$, so daß für $t\in(o,T_1]$ die Ungleichung

$$F(x)(v(a+tb,x)-v(a,x))\geqq ct$$

für alle $x\in U$ gilt.

Die Menge I-U ist abgeschlossen in I und für $x\in I-U$ gilt

$$|F(x)|<||F||_I \ ;$$

daraus folgt $h:=||F||_I - ||F||_{I-U} > o$.

Die Dreiecksungleichung ergibt für $t\in\mathbb{R}$

$$||v(a+tb)-v(a)||_I \leqq ||(tb,grad(v(a)))||_I +$$
$$+||v(a+tb)-v(a)-(tb,grad(v(a)))||_I \ ;$$

daher gibt es ein $T_2>o$, so daß für $t\in(o,T_2]$ gilt:

$$||v(a+tb)-v(a)||_I \leqq 2t||(b,grad(v(a)))||_I \ .$$

Für $T\in\mathbb{R}$ gelte

$$o<T<\min\{T_1,T_2,\frac{c}{4||(b,grad(v(a)))||_I^2},\frac{h}{4||(b,grad(v(a)))||_I}\}$$

Für $x\in U$ erhält man also

$$|f(x)-v(a+Tb,x)|^2 =$$
$$=|f(x)-v(a,x)|^2-2F(x)(v(a+Tb,x)-v(a,x))+|v(a,x)-v(a+Tb,x)|^2\leqq$$
$$\leqq||F||_I^2 -2cT +4T^2||(b,grad(v(a)))||_I^2 \leqq ||F||_I^2 -2cT +cT <$$
$$<||f-v(a)||_I^2$$

Für $x\in I-U$ gilt die Abschätzung

$$|f(x)-v(a+Tb,x)|\leqq|F(x)|+|v(a,x)-v(a+Tb,x)| \leqq$$
$$\leqq||F||_I-||F||_I+||F||_{I-U}+||v(a)-v(a+Tb)||_I \leqq$$
$$\leqq||F||_I -h+2T||(b,grad(v(a)))||_I \leqq ||F||_I -h+\frac{h}{2}<||f-v(a)||_I$$

Also gilt für $x\in I$

$$|f(x)-v(a+Tb,x)|< ||f-v(a)||_I = ||f-E(a)||_I \ .$$

Da $T>o$ beliebig klein gewählt werden kann, ist die Behauptung nach Definition von V gezeigt.

Satz 4.2 (Braess, [2])

$E(a,x)=\sum_{i=1}^{K} a_i e^{t_i x}$ mit dem Grad K und $S(a)=\mathrm{sign}(E(a))$ sei

Minimallösung für f bzgl. V^o_{N-1}, nicht aber bzgl. V_N.
In I habe $f-E(a)$ eine positive (negative) Alternante der Länge N+K, deren Alternantenpunkte mit x_i, $1\leq i\leq N+K$, bezeichnet werden. $S=(S_1,\ldots,S_N)$ sei ein Vorzeichenvektor mit N Komponenten, so daß erfüllt ist:

1. Durch Streichen von N-K Komponenten S_{i_k}, $1\leq k\leq N-K$, in S erhält man S(a) und es gilt $i_{k-1}<i_k$.
2. $S_{i_{N-K}}=+1$ $(S_{i_{N-K}}=-1)$ und $\mathrm{sign}(S_{i_{k-1}})=-\mathrm{sign}(S_{i_k})$, $1<k\leq N-K$.

Behauptung:
$V_N(S)$ enthält eine bessere Approximation als E(a).

Beweis:

(4.2) Die Menge $T=\{s_i \mid K+1\leq i\leq N,\ s_i<s_j \text{ für } i<j\}\subseteq \mathbb{R}$ sei so gegeben, daß T und das Spektrum von E(a) disjunkt sind.

Ordnet man die Frequenzen von E(a) und die Elemente von T der Größe nach, so erhält man eine aufsteigende Kette:

$$s'_1<s'_2<s'_3<\ldots<s'_N\ ;$$

für die Elemente von T gelte hierbei:

$$s'_{i_k}=s_{K+k}\quad,\ 1\leq k\leq N-K\ . \tag{4.3}$$

Nach Hilfssatz 4.1 gibt es in V_N eine bessere Approximation

$E(c,x)=E(b,x)+\sum_{i=K+1}^{N} c_i e^{s_i x}$ mit $E(b,x)=\sum_{i=1}^{K} b_i e^{s_i x}$, so daß gilt:

(4.4)
1. $\mathrm{sign}(E(a)) = \mathrm{sign}(E(b))$
2. $|s_i-t_i|<|s_i-t|$ für alle $t\in T$ und für alle i mit $1\leq i\leq K$.

Dies ergibt sich aus der Konstruktion der besseren Approximation E(c) nach Hilfssatz 4.1 . Nach einem mehrfach benutzten Schluß besitzt $E:=E(c)-E(a)\not\equiv o$ in (x_1,x_{N+K}) mindestens N+K-1 Nullstellen und $S(E)=\mathrm{sign}(E)$ hat nach Satz 1.3 daher mindestens N+K-1 Vorzeichenwechsel;

(4.5) S(E) besteht also aus genau N+K Komponenten, die abwechselnd positiv und negativ sind.

Nach Satz 1.1 hat E genau N+K-1 reelle Nullstellen. Ist die Alternante der Voraussetzung positiv (negativ), so gilt

$$E(x_{N+K})>o \quad (E(x_{N+K})<o).$$

Man beachte hierbei, daß nach Voraussetzung alle Alternanten von f-E(a) in I mit der Länge N+K positiv (negativ) sind, falls eine Alternante dieser Länge positiv (negativ) ist.
Aus (4.5) folgt weiter $c_i \neq o$ für $K+1 \leqq i \leqq N$.
Wie im ersten Teil des Beweises zu Satz 3.4 erhält man damit

für $s_N > t_K$: $\operatorname{sign}(c_N)=+1 \quad (\operatorname{sign}(c_N)=-1)$,

$$\text{für } s_N<t_K : \begin{cases} \operatorname{sign}(b_K)=+1 & (\operatorname{sign}(b_K)=-1) \quad \text{falls } s_K>t_K \text{ gilt} \\ \operatorname{sign}(a_K)=-1 & (\operatorname{sign}(a_K)=+1) \quad \text{falls } s_K<t_K \text{ gilt.} \end{cases}$$

Die (N+K)-te Komponente von S(E) ist damit positiv (negativ).
In (s_i, s_{i+1}) liegt für $K+1 \leqq i \leqq N-1$ stets eine gerade Anzahl von Elementen aus dem Spektrum von E(a)+E(b); dies folgt aus (4.4) . Mit (4.5) erhält man also für $K+1 \leqq i \leqq N$ $\operatorname{sign}(c_i)=-\operatorname{sign}(c_{i+1})$ und $\operatorname{sign}(c_N)=+1$ $(\operatorname{sign}(c_N)=-1)$.
Damit ist die Behauptung gezeigt, denn $E(c) \in V_N(\bar{S})$ ist eine bessere Approximation an f und wegen (4.3) erfüllt $\bar{S}=\operatorname{sign}(E(c))$ die Voraussetzung des Satzes.

<u>Bemerkung 4.1</u> (Braess, [2])
Für K=N-1 läßt sich die Zahl der verschiedenen Vorzeichenvektoren S, die man nach der in Satz 4.2 verwandten Konstruktion erhalten kann, genau bestimmen:
Es liege eine positive Alternante von f-E(a) vor und S(a) habe genau k^- negative Komponenten; die entsprechenden Frequenzen von E(a) teilen $\mathbb{R}$ in k^-+1 Intervalle. Die Menge T besteht hier nur aus dem Element s_N und die Wahl von s_N bestimmt, wie oben gezeigt, den Vorzeichenvektor S der besseren Approximation E(c). Daher gibt es also k^-+1 verschiedene Vorzeichenklassen, die bessere Approximationen enthalten.
Ist die betrachtete Alternante negativ, so ergeben sich analog k^++1 Vorzeichenklassen, wenn S(a) genau k^+ positive Komponenten enthält.

<u>Satz 4.3</u> (Braess, [2])

$E(a) \in V_{N-1}^{o}$ mit $grd(E(a))=K \neq o$ und dem Vorzeichenvektor $S(a)=(S_1,\ldots,S_K)$ sei die Minimallösung bzgl. V_{N-1} für f auf I. Erfüllt jede Minimallösung E für f bzgl. V_N die Beziehung $E \notin V_{N-1} \cup V_N^{+} \cup V_N^{-}$, dann gibt es in V_N mindestens zwei lokale Minima für f.

Beweis:

Es sei wieder $F=f-E(a)$. Nach Voraussetzung besitzt F in I eine Alternante der Länge N+K und falls diese positiv (negativ) ist, sind alle Alternanten von F in I mit dieser Länge positiv (negativ).

In V_N werden zwei verschiedene Vorzeichenklassen angegeben, die bessere Approximationen als E(a) enthalten:

1.Fall: K<N-1

Es sei s=+1, falls F in I eine positive Alternante der Länge N+K besitzt, und s=-1, falls diese negativ ist.

Die N-K Komponenten von

$$((-1)^{N-K-1}s,\ldots,(-1)^{1}s,(-1)^{o}s)$$

lassen sich auf mindestens zwei Arten in S(a) einfügen, so daß man die verschiedenen Vorzeichenvektoren

$$(S_1,(-1)^{N-K-1}s,\ldots) \quad \text{und}$$

$$((-1)^{N-K-1}s,(-1)^{N-K-2}s,S_1,\ldots) \quad \text{erhält.}$$

2.Fall: K=N-1

F habe in I eine positive Alternante der Länge N+K.

Hieraus folgt $E(a) \notin V_{N-1}^{+}$: Denn mit $E(a) \in V_{N-1}^{+}$ erhält man durch Einfügen von s=+1 in S(a) stets einen Vorzeichenvektor mit N positiven Komponenten, so daß nach Satz 4.2 in $V_N^{+}-V_{N-1}$ eine bessere Approximation als E(a) enthalten ist; da f bzgl. V_N^{+} eine beste Approximation auf I besitzt und diese nach Satz 3.2 Element von $V_N^{+}-V_{N-1}$ und auch beste Approximation bzgl. V_N ist, erhält man einen Widerspruch zur Voraussetzung.

Es gibt also ein $i \in \{1,\ldots,N-1\}$ mit $S_i=-1$.

Durch Einfügen von s=+1 in S(a) erhält man die Vorzeichenvektoren $(s_1,\ldots,s_{i-1},-1,s_{i+1},\ldots,s_{N-1},+1)$ und

$(s_1,\ldots,s_{i-1},+1,-1,\ldots,s_{N-1})$.

Hat F eine negative Alternante der Länge N+K in I, so erhält man analog $E(a)\notin V_{N-1}^-$; es gibt also ein $i\in\{1,\ldots,N-1\}$ mit $s_i=+1$ und man erhält durch Einfügen von s=-1 entsprechend die Vorzeichenvektoren

$(s_1,\ldots,s_{i-1},+1,s_{i+1},\ldots,s_{N-1},-1)$ und

$(s_1,\ldots,s_{i-1},-1,+1,\ldots,s_{N-1})$.

Es gibt somit für K=N-1 und K<N-1 nach Satz 4.2 mindestens zwei verschiedene Vorzeichenklassen in V_N, die bessere Approximationen an f als E(a) enthalten.

Nun wird gezeigt, daß es bezüglich jeder der soeben konstruierten Vorzeichenklassen eine beste Approximation an f gibt:

Es sei also $V:=V_N(S)$ so gegeben, daß V eine bessere Approximation an f enthält als E(a); nach Voraussetzung hat diese den Grad N. Die Folge $(h_m)_{m\in\mathbb{N}}$ sei eine Minimalfolge in V; es gilt also $\lim\limits_{m\to\infty}||f-h_m||_I = \inf\limits_{v\in V}||f-v||_I$ und es kann o.B.d.A. $grd(h_m)=N$ für $m\in\mathbb{N}$ angenommen werden.

Nach Satz 2.3 gibt es eine Teilfolge in $(h_m)_{m\in\mathbb{N}}$, die auf jedem abgeschlossenen Teilintervall von I gleichmäßig konvergiert und die Grenzfunktion ist ein Element von V nach Korollar 2.5; daraus folgt mit $f\in C(I)$ wie im Beweis zu Satz 2.4 die Existenz einer besten Approximation h bzgl. V .

Nach Voraussetzung gilt $||f-E(a)||_I > ||f-h||_I$ und mit $0<\varepsilon=||f-E(a)||_I - ||f-h||_I$ gilt für $h_1\in V_N$ mit $||h-h_1||_I<\varepsilon$

$||f-h_1||_I \leq ||f-h||_I + ||h-h_1||_I < ||f-E(a)||_I$.

Mit Korollar 2.7 gibt es also eine Umgebung U von h in V_N mit $U\subseteq V_N(S)-V_{N-1}$, so daß gilt:

$$||f-h_1||_I \geq ||f-h||_I \qquad h_1\in U .$$

Dies war zu zeigen.