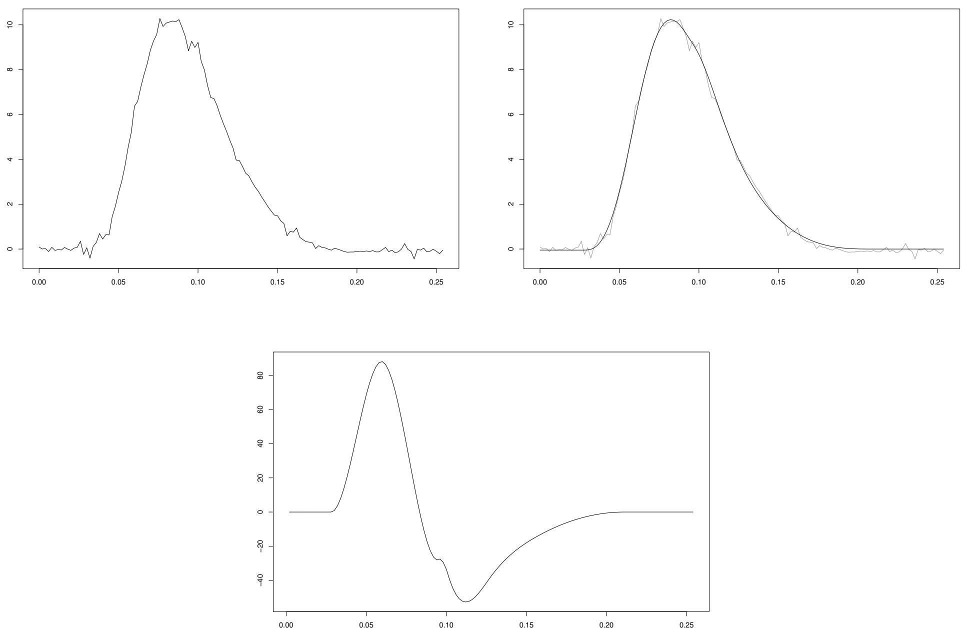
Glatte nichtparametrische Regression

unter formerhaltenden Bedingungen



Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

vom Fachbereich Mathematik und Informatik

der Universität Essen genehmigte

DISSERTATION

von

Ali Majidi aus Rasht (Iran)

Tag der mündlichen Prüfung: 30. Mai 2003

Gutachter: Prof. Dr. P.L. Davies (Universität Essen-Duisburg)

Prof. Dr. Enno Mammen (Universität Heidelberg)

Ali Majidi

Universität Essen

45117 Essen

ali.majidi@uni-essen.de

Inhaltsverzeichnis

1 Einleitung **1**

2 Regression und Extremwertkontrolle **9**

2.1 Problemstellung . . . 9

2.2 Parametrische Ansätze . . . 10

2.3 Nichtparametrische Ansätze . . . 11

2.3.1 Kernschätzer . . . 11

2.3.2 Strafterme . . . 12

2.3.3 Basisfunktion-Methoden . . . 16

2.3.4 Zusammenfassung . . . 16

2.4 Extremwertkontrolle und Residuenbasierte Regression . . . 18

2.5 Zusammenfassung . . . 24

3 Das Glättungsproblem **25**

3.1 Problemstellung . . . 26

3.2 Splines in der nichtparametrischen Regression . . . 27

3.3 Diskrete Nebenbedingungen . . . 35

3.4 Berechnung auf äquidistanten Gittern und QP . . . 38

3.5 Zusammenfassung . . . 42

4 Quadratisches Programmieren **43**

4.1 Problemstellung . . . 43

4.2 Aktive-Menge-Methoden . . . 45
4.2.1 Implementierung einer Aktive-Menge-Methode . . . 45
4.2.2 Aktive-Menge Methoden für den stetigen Fall . . . 59
4.3 Gradient-Projektion-Methoden . . . 62
4.3.1 Implementierung einer Gradient-Projektion-Methode . . . 64
4.3.2 QSOR und Multigrid . . . 65
4.3.3 Multigrid-QSOR . . . 66
4.4 Zusammenfassung . . . 66

5 Rauscherkennungskriterium und Multiresolution **73**
5.1 Nichtparametrische Regression, Residuen und Optimierung . . . 74
5.2 Asymptotische Konvergenz . . . 79
5.3 Berechnung . . . 88
5.3.1 Strafterm-Algorithmen und LANCELOT . . . 89
5.4 Schätzen von Ableitungen . . . 96

6 Anwendung auf die Dichteschätzung **99**
6.1 Problemstellung . . . 99
6.2 Zusammenfassung . . . 107

7 Weiterführende Gedanken **109**

Kapitel 1

Einleitung

You see things; and you say "Why?"

But I dream things that never were; and I say "Why not?"

GEORGE BERNARD SHAW

Eine häufig gestellte Frage in den meisten (empirischen) Wissenschaften ist die Frage der Interpretation von Beobachtungen. Eine Reihe von Beobachtungen, die z.B. an verschiedenen Zeitpunkten, bzw. an verschiedenen Orten gemacht wurden, können mittels zweier Komponenten als

$$(t_1, y_1), (t_2, y_2), \ldots, (t_n, y_n),$$

verstanden und, z.B. in einer Graphik (wie in Abbildung 1.1), visualisiert werden. In vielen Fällen stellt man jedoch fest, dass die zugrunde liegende Struktur des beobachteten Vorgangs nicht klar erkennbar und damit nicht interpretierbar ist, da die Daten durch (Mess-)Fehler *verrauscht* worden sind. Die *Regression* ist ein Begriff in der Statistik, mit dem man das Entrauschen der Daten, in der obigen Situation beschreibt. In dieser Situation verwendet man häufig die Modellierung

$$\textit{Daten} = \textit{Signal} + \textit{Rauschen},$$

um den Vorgang zu beschreiben. Gesucht ist das Signal, das wir als eine Funktion f bezeich-

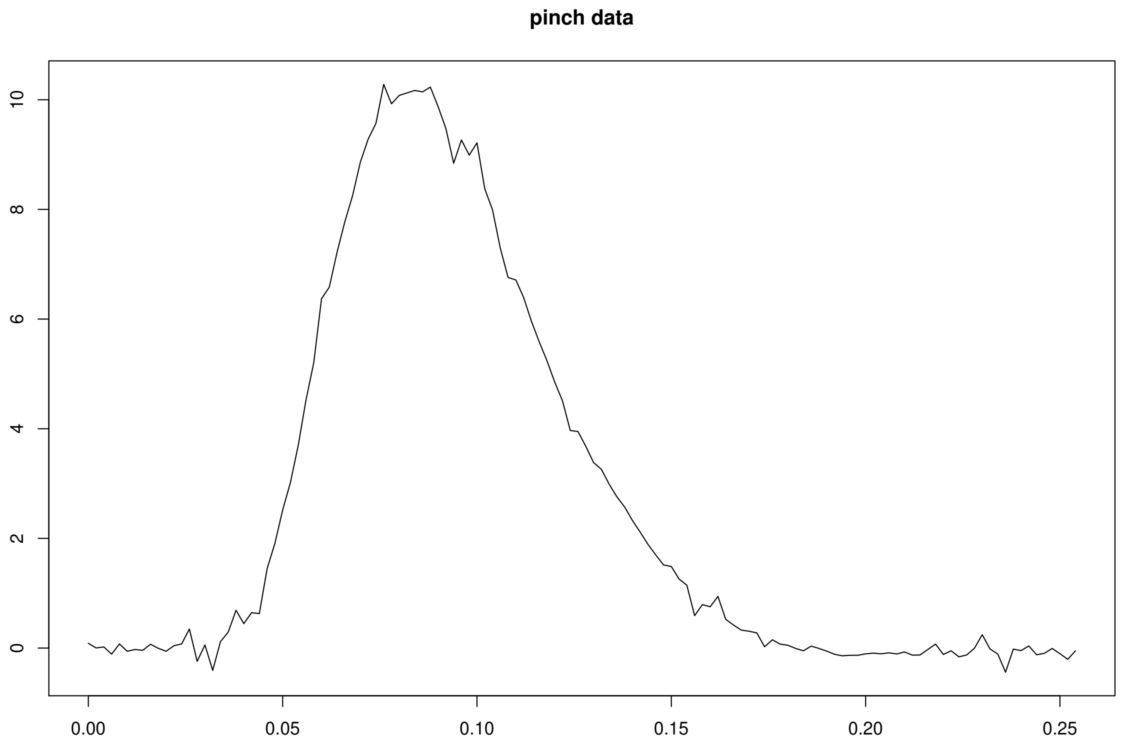


Abbildung 1.1: Die Daten stammen aus einem biomechanischen Versuch, der dazu dienen soll, die Neurophysiologie der Muskelgruppen zwischen dem Zeigefinger und dem Daumen zu verstehen. Dabei wurde auf ein Messgerät zwischen dem Daumen und dem Zeigefinger eine Kraft ausgeübt, die kontinuierlich bis 10 Newton erhöht wurde (vgl. [55] und [56]).

nen (siehe Abbildung 1.2 und Abbildung 1.3). In dieser Arbeit werden nur eindimensionale Daten betrachtet und verschiedene Methoden der Regression vorgestellt. Dabei sagt man häufig, dass das Signal *einfach* und das Rauschen *komplex* sei. Diese Bezeichnungen stammen aus der menschlichen Eigenschaft, strukturlose Konstrukte als komplex einzustufen. Genauer formuliert, soll das Signal die Struktur des Datensatzes und das Rauschen, den strukturlosen oder chaotischen Anteil des Datensatzes beschreiben. Diese Begriffe sind alle unpräzise und sollen hier, im eindimensionalen Fall, genauer erläutert werden. Dabei lassen wir uns von der "elementaren Kurvendiskussion" leiten. Soll eine eindimensionale (differenzierbare) Funktion diskutiert werden, so genügt es die Extremwertstellen und das Krümmungsverhalten der Funktion zu bestimmen, um eine ausreichende Skizze des Graphen zu erstellen. Dabei gilt: Die Skizze wird einfacher, je weniger Extremstellen die Funktion hat und je einfacher das Krümmungsverhalten ist. Kurven mit einem einfachen Krümmungsverhalten werden als *glatt* wahrgenommen. Somit heißt unsere Forderung nach der Einfachheit des Signals, dass das Signal eine glatte Funktion mit wenigen Extremwerten sein soll, die unserer Modellierung entspricht. Im folgenden werden diese Begrifflichkeiten genauer beschrieben und damit messbar gemacht.

Diese Arbeit schließt sich an die Arbeiten von Dr. Martin Löwendick [42] und Dr. Lars Metzner [50] an. Es werden vier Testfunktionen benutzt, um die Eigenschaften der Methoden zu visualisieren. Die Datensätze (vgl. Abbildung 1.5, Abbildung 1.4) wurden zum ersten Mal von [18] vorgestellt und werden häufig in der Literatur benutzt. Sie stellen sehr unterschiedliche Herausforderungen an einer Regressionsmethode und sind daher zum Testen und Vergleichen von Regressionsmethoden geeignet.

Die Algorithmen werden mittels eines "Pseudo-Codes" beschrieben, das eine Implementierung der Methoden in einer höheren Programmiersprache z.B. **R** oder **MATLAB** erleichtern soll. Die Programme, die in dieser Arbeit benutzt werden, wurden hauptsächlich in der Programmiersprache **C** implementiert.

Im zweiten Kapitel wird zunächst eine grobe Einführung in die Problematik der Regression gegeben. Einige nichtparametrische Methoden u.a. die Run-Methode und die Taut-String-Methode werden vorgestellt (vgl.[15]).

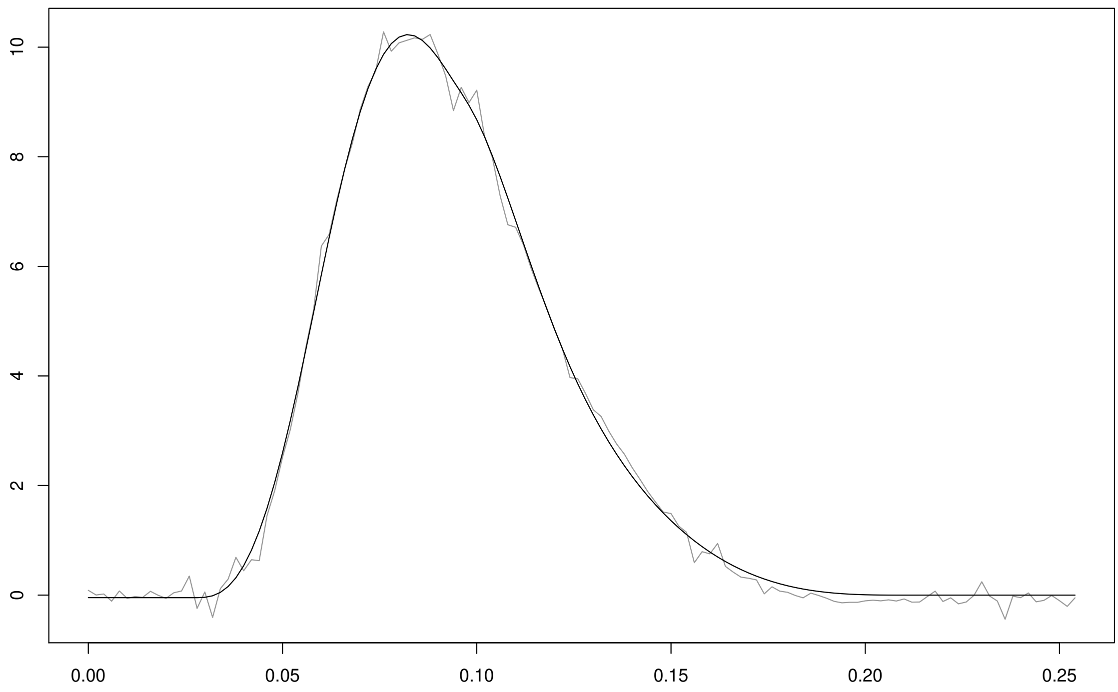


Abbildung 1.2: Die Regressionsfunktion für den Pinch-Datensatz ermittelt mittels glatter residuenbasierter Nichtparametrischer Regression. Die Daten sind in Grau eingezeichnet. Offensichtlich wird die Struktur des Datensatzes mittels einer glatten Funktion wiedergegeben. Die Regressionsfunktion hat, wie erwartet, genau ein Maximum.

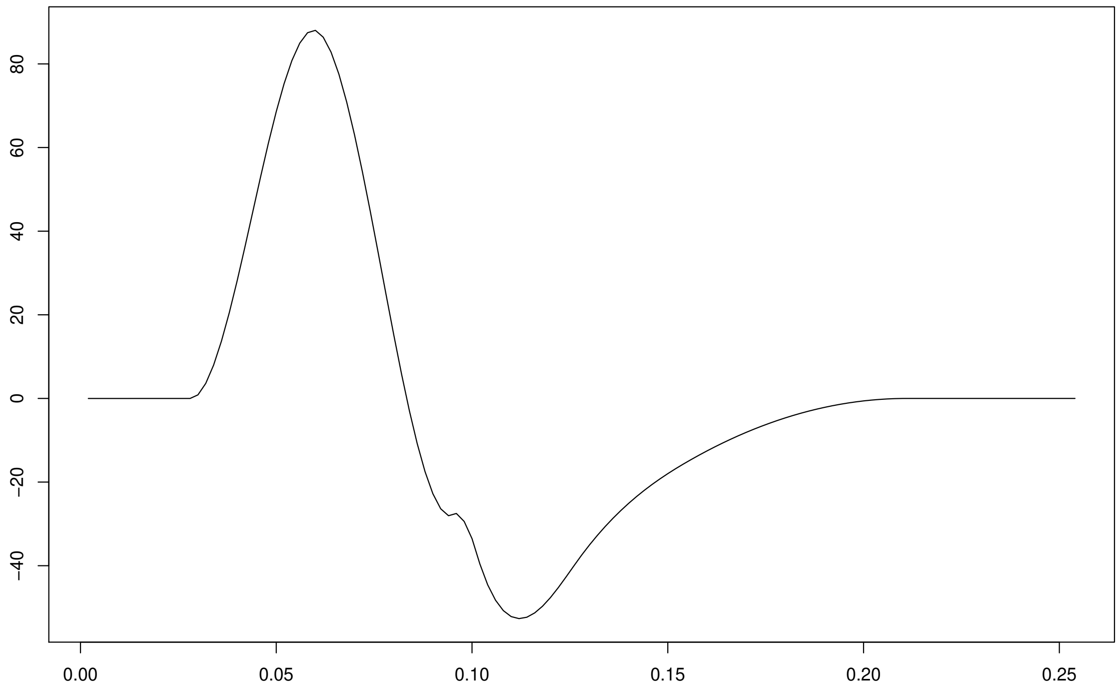


Abbildung 1.3: Die Ableitung der Schätzfunktion (vgl. Abbildung 1.2) für die Pinch-Daten. Die Methode liefert automatisch eine gute Approximation der Ableitung.

Das Problem der Glättung mit Zusatzinformationen über das Monotonieverhalten (*formerhaltende Bedingungen*) wird im dritten Kapitel betrachtet. Die vorher genannten Methoden können dann als eine Möglichkeit gesehen werden, diese Informationen zu ermitteln. Einige Ansätze werden vorgestellt, deren Existenz und Eindeutigkeit untersucht werden sollen.
Schließlich werden im vierten Kapitel die Diskretisierung und die Berechnung betrachtet. Die Schwierigkeit im Zusammenhang mit großen Datensätzen wird thematisiert. Ferner werden einige Verfahren der numerischen Optimierung vorgestellt, und einige Ansätze zur effizienteren Berechnung der Schätzfunktion werden besprochen.
Im fünften Kapitel wird die residuenbasierte Regression im Allgemeinen betrachtet. Hier wird ebenfalls eine effiziente Methode zur Berechnung implementiert und die Ergebnisse werden besprochen.
Abschließend beschreiben wir die Anwendung der Erkenntnisse auf das Problem der Dichteschätzung, die sich als kanonische Erweiterung der Taut-String-Idee bietet (vgl. [15]).
In dieser Arbeit werden keine wahrscheinlichkeitstheoretischen Grundkonzepte definiert. Insbesondere werden die Begriffe Zufallsvariable und Verteilung im üblichen Sinne benutzt und nicht weiter erklärt. Wir verweisen z.B. auf [26] oder [19].

In diesem Rahmen möchte ich allen danken, die mir während der Entstehung dieser Arbeit geholfen haben. Prof. Dr. Laurie Davies hat mein Interesse für die Fragestellungen der Statistik geweckt. Sein Enthusiasmus als Wissenschaftler wirkte auf mich stets herausfordernd. Neben Dr. Arne Kovac, der mich während der Studien- und Forschungszeit in Essen als Freund und Lehrer stets unterstützt hat, war Monika Meise mir eine großartige Arbeitskollegin und Freundin. Luisa Karmelita hat mir mit geistiger Frische und interessanten Anregungen, beim Verfassen der Arbeit geholfen. Meiner Familie, insbesondere meiner Mutter Behnaz Haftchenari und meinem Bruder Dr. Mohammad Majidi, der mich für das Studium der Mathematik begeistert hat, möchte ich ebenfalls herzlich danken.

20. Juni 2003 Ali Majidi

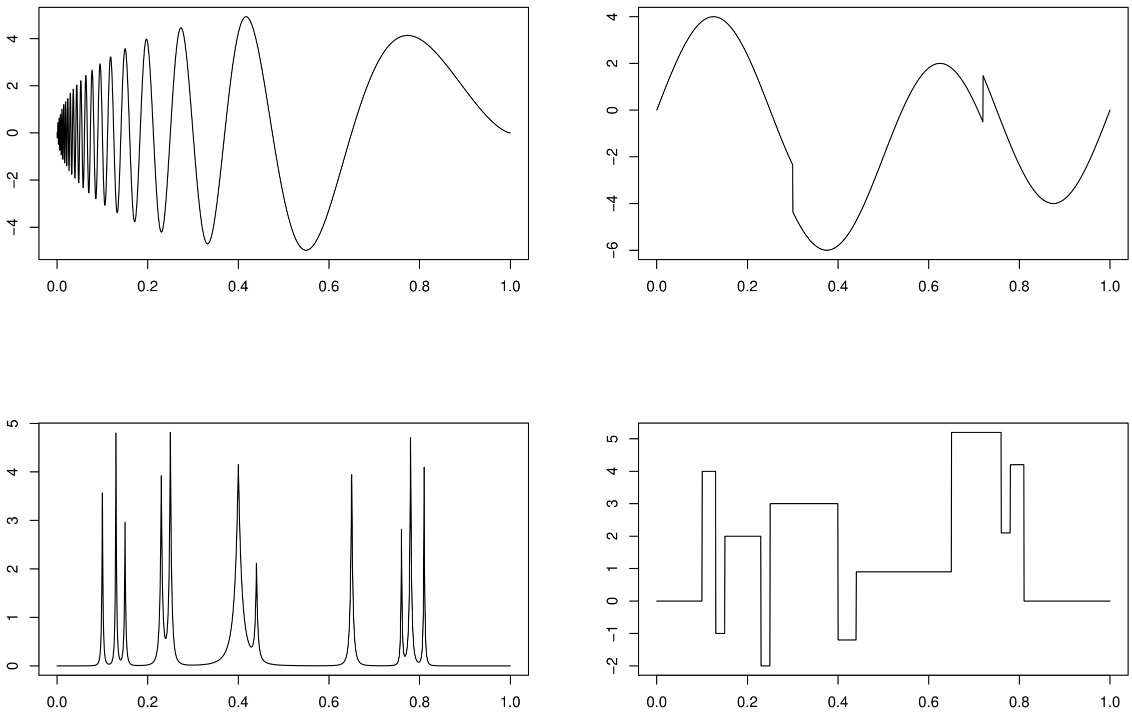


Abbildung 1.4: Die Funktionen Doppler, Heavisine, Bumps und Blocks werden im folgenden das Signal darstellen.

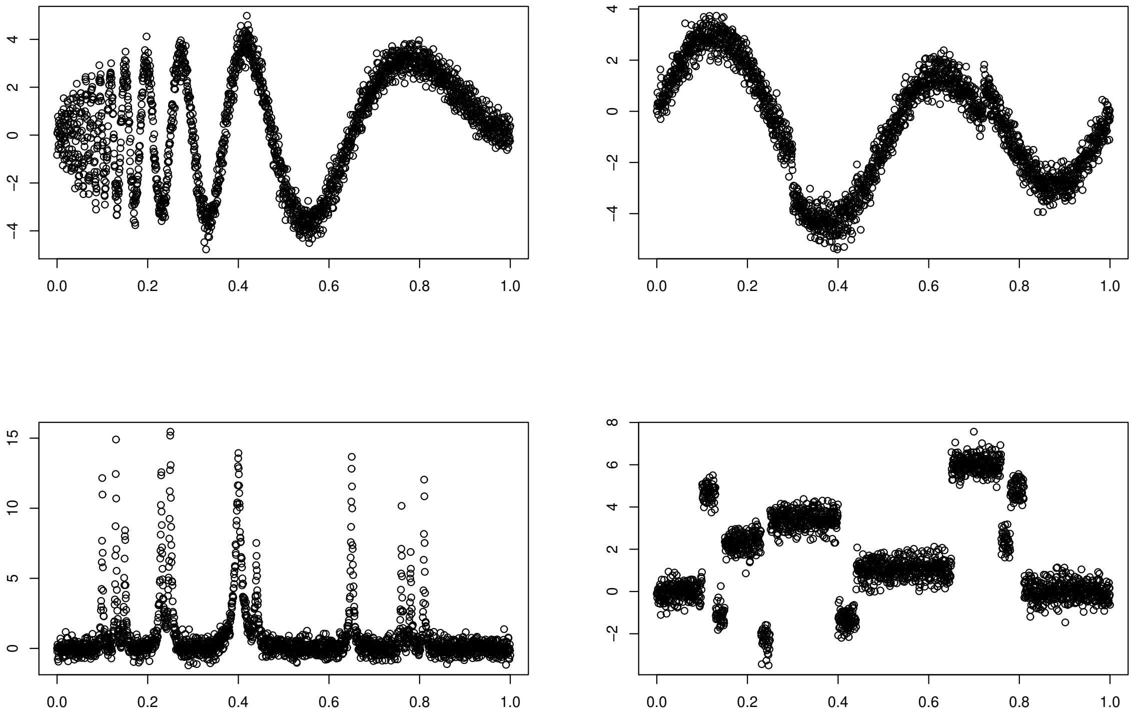


Abbildung 1.5: Die verrauschten DJ-Datensätze (eingeführt von [18]) werden im folgenden zur Veranschaulichung der Methoden verwendet.

Kapitel 2

Regression und Extremwertkontrolle

In diesem Kapitel wird zunächst eine kurze Einführung in die Problematik der Regression gegeben. Danach werden die gängigsten Methoden vorgestellt. Für eine genauere Beschreibung seien z.B. die Bücher [51], [25], [55] und [35] genannt. Schließlich werden die Vorteile einer Regressionsmethode mit Extremwertkontrolle[1] diskutiert.

2.1 Problemstellung

Gegeben sei ein Datensatz

$$\{(t_1, y_1), \ldots, (t_n, y_n)\}$$

mit

$$0 \leq t_1 < t_2 < \cdots < t_n \leq 1, \ y_i \in \mathbb{R}, \ (i = 1, \ldots, n).$$

Im folgenden sprechen wir von *dem Datensatz* und werden ihn mit y bezeichnen. Es sei angenommen, dass die Daten folgendem gängigen Modell entsprechen:

$$y_i = f(t_i) + \varepsilon_i \ \ (i = 1, \ldots, n), \tag{2.1}$$

[1] Die Bedeutung wird später in diesem Kapitel erklärt.

wobei $f : [0,1] \to \mathbb{R}$ eine Funktion sein soll, die bei Bedarf genauer spezifiziert wird. Außerdem sei das *Rauschen* $\{\varepsilon_i | 1 \leq i \leq n\}$ normal verteilt, das heißt:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma), \ (i = 1, \ldots, n, \sigma > 0).$$

Dies wird häufig als "Weißes Rauschen" bezeichnet. In der Regression[2] versucht man, unter der Annahme des obigen Modells, aus dem verrauschten Datensatz y die ursprüngliche Struktur der Funktion f zurück zu gewinnen. Hierzu bestimmt man eine Funktion $\hat{f}$, die die Struktur von f wiedergibt. Diese Funktion wird häufig als *Schätzer* für f bezeichnet. Es gibt unzählige Ansätze zur Lösung der Regressionsaufgabe in Form von (2.1), die je nach Art der Problemstellung zur Anwendung geeignet sind.

2.2 Parametrische Ansätze

Offensichtlich sind die Bedingungen, die in (2.1) gestellt werden, nicht präzise genug, um einen Schätzer für f zu bestimmen. Daher stellt man zusätzliche Bedingungen an f, die das Modell (2.1) weiter einschränken sollen. Eine Möglichkeit, f zu spezifizieren, ist die Annahme, dass f einer bestimmten (parametrischen) Familie angehört. Genauer formuliert, hat man eine Familie

$$f_\lambda, \lambda \in \Lambda, (\Lambda \in \{\mathbb{Z}, \mathbb{N}, \mathbb{R}, \cdots\}),$$

mit einem $\lambda^* \in \Lambda$, so dass gilt: $f = f_{\lambda^*}$. Nun versucht man $\hat{\lambda}$ so zu bestimmen, dass

$$\hat{\lambda} \approx \lambda^* \text{ und damit } \hat{f} := f_{\hat{\lambda}} \approx f$$

gilt. Dabei soll "$\approx$" bedeuten, dass $\|\hat{f} - f\|$ möglichst klein wird, für eine bestimmte Norm $\|.\|$. Eines der bekanntesten Methoden, den Abstand zu messen, ist die euklidische Norm. In diesem Fall liefert

$$\hat{\lambda} := \underset{\lambda \in \Lambda}{\operatorname{argmin}} \sum_{i=1}^{n} (f_\lambda(t_i) - y_i)^2,$$

eine Approximation für f mittels $\hat{f} := f_{\hat{\lambda}}$. Wählt man eine bestimmte Familie und ein geeignetes Abstandsmaß, so haben parametrische Ansätze viele Vorteile. Unter anderem

[2] Das Wort Regression bedeutet: Rückgang, Rückbewegung.

lässt sich der Parameter λ häufig einfach interpretieren. Die Wahl einer geeigneten Familie setzt jedoch Zusatzinformationen über den Datensatz voraus. Sind diese Informationen nicht ausreichend, um eine Wahl zu treffen, so können die Folgen für das Ergebnis der Methode gravierend sein (vgl. [29]). Diese Tatsache wird oft ignoriert, und parametrische Methoden werden ohne eine fundierte Begründung benutzt (vgl. [23]).
In Situationen, die neu und kompliziert sind, haben wir oft nur die Daten zur Verfügung. und kompliziert sind. Nehmen wir z.B. an, die Daten würden die Oberflächenmatrialien aus einem unbekannten Planeten beschreiben. In diesem Fall können wir kaum gerechtfertigte Annahmen über die Funktion f machen. Daher benötigt man flexiblere Methoden für das Regressionsproblem. In der nichtparametrischen Regression wird die notwendige Flexibilität durch eine Vergrößerung der Menge der zulässigen Funktionen:

$$\{f_\lambda, \lambda \in \Lambda\},$$

erreicht. Zum Beispiel kann man versuchen einen geeigneten Schätzer für f in der Menge aller differenzierbaren Funktionen zu finden. Diese Vorgehensweise erhöht den Einfluss des Datensatzes y auf die Schätzfunktion $\hat{f}$. Dadurch "[...] lässt man die Daten für sich selbst sprechen." (vgl. [23]).

2.3 Nichtparametrische Ansätze

Im folgenden werden die wichtigsten Methoden der nichtparametrischen Regression vorgestellt. Dabei sei angemerkt, dass diese Disziplin innerhalb der Statistik in den letzten Jahren stark gewachsen ist. Daher geben wir hier eine grobe Einführung und verweisen auf die Literatur z.B. [23] oder [68].

2.3.1 Kernschätzer

Ist die Funktion f in (2.1) eine Konstante, so ist der Mittelwert über y:

$$\frac{1}{n}\sum_{i=1}^{n} y_i,$$

ein guter Schätzwert für f. Dies ist der Grundgedanke für den *Kernschätzer*, einer der ältesten Methoden der nichtparametrischen Regression. Man approximiert dabei f an einem Punkt t mit Hilfe gewichteter, lokaler Mittelwerte. Bestimme also $w_i : [0,1] \to \mathbb{R}, (i = 1, \ldots, n)$ geeignet und setze

$$\hat{f}(t) := \sum_{i=1}^{n} w_i(t) y_i.$$

Die Regression der Funktion f reduziert sich demnach auf die Wahl der Funktionen w_i, die mit Hilfe einer *Kernfunktion* definiert werden.

Definition 2.1 (Kernfunktion) Zu $k \geq 1$, nennt man eine symmetrische Funktion $K : \mathbb{R} \to \mathbb{R}$ mit kompaktem Träger und

$$\int K(t)^2 dt < \infty, \int t^j K(t) dt = \begin{cases} 1, & (j = 0), \\ 0, & (j = 1, \ldots k-1), \\ \alpha \neq 0, & (j = k) \end{cases}$$

eine *Kernfunktion* k-ten Grades.

In der Literatur beschränkt man sich oft auf eine Kernfunktion zweiten Grades. Mit Hilfe der Kernfunktion K kann man die Gewichtsfunktionen $w_j(t)$ mittels

$$w_j(t) := K\left(\frac{t_j - t}{h}\right)$$

definieren. Dabei bestimmt der Parameter $h > 0$, welche Beobachtungen bei der Berechnung der Funktion $\hat{f}$ berücksichtigt werden. Die wesentliche Rolle von h wird an der Abbildung 2.1 deutlich.

2.3.2 Strafterme

Eine weitere Methode der nichtparametrischen Regression ist das Verwenden eines modifizierten Abstandmaßes. In den parametrischen Modellen bewirkt die Wahl einer Familie, dass wir eine einfache Funktion als Schätzer erhalten. Die Wiedergabe der Struktur von f soll dabei durch die "Nähe" zum Datensatz y garantiert werden. Lassen wir z.B. alle differenzierbaren Funktionen als Schätzer zu, so existieren Funktionen, die zulässig sind

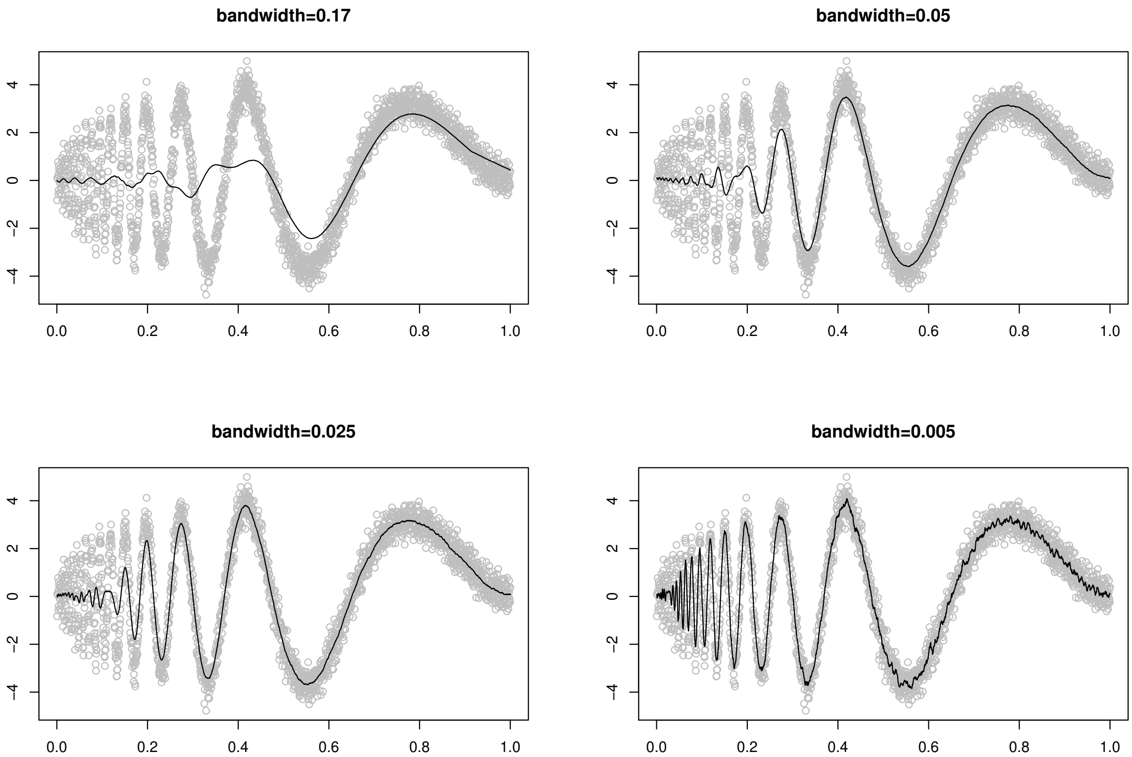


Abbildung 2.1: Das Bild zeigt die Anwendung von einem Kernschätzer auf dem Doppler-Datensatz mit einer festen Bandbreite. Die verschiedenen Bilder verdeutlichen die Problematik der richtigen Parameterwahl.

und y interpolieren. Somit haben sie den geringsten Abstand zum Datensatz y, bewirken aber keine Vereinfachung des Modells. Um gleichzeitig eine Vereinfachung des Modells zu bewirken, erweitert man den Abstandsbegriff, in dem man das Modell um einen Strafterm ergänzt:

$$\hat{f} := \arg\min \sum_{i=1}^{n} (g(t_i) - y_i) + \lambda S(g).$$

Häufig verwendet man:

$$S(g) = \int_0^1 g^{(m)}(t)^2 dt,$$

wobei $g^{(m)}$ die m-te Ableitung von g bezeichnet. In diesem Fall führt das Minimierungsproblem auf einen *Spline* (vgl. [67]). Im nächsten Kapitel werden wir diese Vorgehensweise näher beschreiben.

Die beiden genannten Methoden haben folgende Gemeinsamkeiten:

1) Es gibt (mindestens) einen Parameter in der Methode, der die Regressionsfunktion stark beeinflussen kann (vgl. Abbildung 2.1 und Abbildung 2.2). Man muss sich überlegen, wie dieser Parameter optimal zu wählen ist. Für komplexere Strukturen liefern alle Parameter unbefriedigende Ergebnisse.

2) Bei allen Methoden können zusätzliche Extremwerte durch die Regressionsmethode erzeugt werden.

Ein Weg, das erste Problem zu behandeln, ist die Einführung von lokalen Parametern. Die Verwendung von lokalen Parametern führt jedoch häufig dazu, dass Artefakte verstärkt auftreten. Außerdem sind für eine gute Wahl der lokalen Parameter zusätzliche Informationen erforderlich. Es gibt verschiedene Vorschläge zur Wahl der lokalen Bandbreiten. Die meisten Vorschläge benötigen jedoch Zusatzinformationen, wie z.B. eine Approximation der zweiten Ableitung (vgl. z.B. [28]), um die lokalen Parameter zu bestimmen. Diese Vorgehensweise erzeugt andere, zum Teil komplexere Probleme (vgl. z.B. [36]). In [49] wird die residuenbasierte Wahl der lokalen Parameter besprochen. Dadurch wird zwar die Anpassungseigenschaft der Kernschätzer verbessert, die grundsätzliche Problematik mit ungewollten Artefakten bleibt jedoch bestehen.

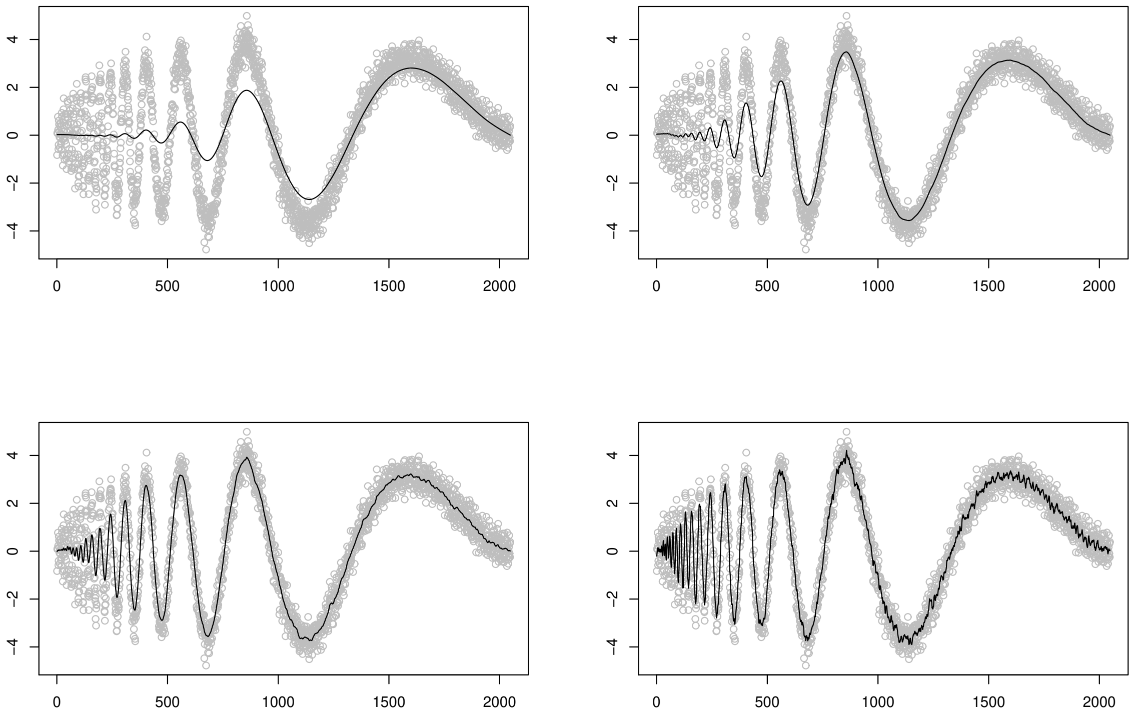


Abbildung 2.2: Das Bild zeigt die Anwendung von Smoothing-Splines mit verschiedenen λ. Die Auswirkung des Parameters λ auf den Schätzer ist ähnlich wie die Bandbreite beim Kernschätzer. Wählt man h groß, so werden die Extremwerte in der ersten Hälfte des Datensatzes stark abgeschnitten. Bei einem kleineren h wird die glatte Rechte Hälfte wieder verrauscht.

2.3.3 Basisfunktion-Methoden

Wie bereits erwähnt, ist die Menge der zulässigen Funktionen in der nichtparametrischen Regression sehr groß. Nehmen wir zum Beispiel an, dass die Funktion f[3] ein Element eines euklidischen $\mathbb{R}$-Vektorraumes mit einer abzählbaren Basis ist, so lässt sich f mittels

$$f = \sum_{i=1}^{L} \alpha_i^* \; \Phi_i, \; (\alpha_i^* \in \mathbb{R}),$$

darstellen, wobei $(\Phi_j)_{j=0}^{\infty}$ eine Basis des besagten Raumes sein soll. Setzt man $L \in \mathbb{N}$ als bekannt voraus, so lassen sich die Koeffinzienten α_i^* mit Hilfe des Abstandsbegriffes approximieren und liefern damit einen Schätzer für f mittels

$$\hat{f} := \sum_{i=1}^{L} \hat{\alpha}_i \; \Phi_i.$$

Die Festlegung einer geeigneten Basis kann dabei große Unterschiede im Resultat bewirken. Häufig benutzt man spezielle Eigenschaften der Basis, um den Datensatz zu entrauschen. Das beutet, dass zunächst zum Datensatz y eine Darstellung

$$\sum_{i=1}^{L_1} \hat{\beta}_i \; \Phi_i,$$

ermittelt wird. Danach verändert man die Koeffizienten, die "typische Eigenschaften" des Rauschens repräsentieren (das sog. *Thresholding*). Als spezielles Beispiel einer Basis seien hier die Fourier-Basis und die Wavelet-Basis (vgl. Abbildung 2.3) genannt. Wir verweisen an dieser Stelle auf die Literatur (vgl. z.B. [55], [12]).

2.3.4 Zusammenfassung

Die Regression von f mittels einer Basis kommt zwar gänzlich ohne die Wahl eines Parameters aus, das zweite Problem bleibt jedoch offensichtlich bestehen (vgl. Abbildung 2.3). Um das zweite Problem zu vermeiden gibt es ebenfalls Lösungsansätze (vgl. [24], [60], [59], [8]), wie man sinnvoller Weise Artefakte in der Schätzfunktion erkennt. Die meisten

[3] Die Frage, ob f wirklich existiert, soll hier nicht diskutiert werden.

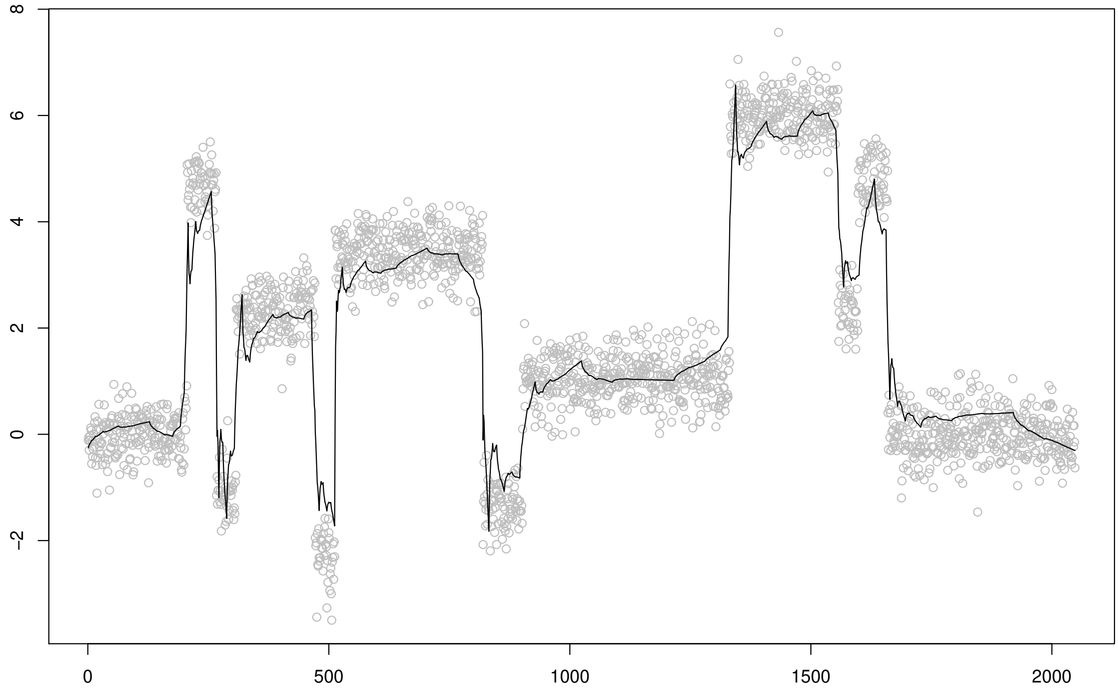


Abbildung 2.3: Das Bild zeigt die Anwendung von *Wavelet-Thresholding* mit Daubechies-Wavelets der Ordnung 3 auf den Blocks-Datensatz. Durch eine geeignete Wahl der Familie wird die Approximation besser, die zusätzlichen Extremwerte an den Sprüngen der Blocks bleiben jedoch stets bestehen

dieser Verbesserungsversuche sind jedoch rechenintensiv. Es wäre voteilhafter, wenn sich diese Artefakte von Anfang an vermeiden ließen. Im nächsten Abschnitt werden wir Methoden kennen lernen, in denen ungewollte Extremwerte dadurch vermieden werden, dass die "Verhinderung von Artefakten" explizit von der Regressionsmethode gefordert wird.

2.4 Extremwertkontrolle und Residuenbasierte Regression

Gehen wir zurück zum Anfang und betrachten das Problem (2.1) der Regression. Haben wir eine mögliche Schätzfunktion $\hat{f}$, so resultiert aus dem Problem (2.1) eine Forderung, die jeder Schätzer, unabhängig von den weiteren Annahmen über die Funktion f, erfüllen muss.
Die Struktur der Schätzfunktion $\hat{f}$ sollte

1) die *wesentlichen Merkmale* von f und

2) *keine weiteren*

beinhalten. Diese Forderung klingt zwar selbstverständlich, wird aber angesichts der vielfältigen Bereichen der Anwendung von Regressionsmethoden, häufig außer Acht gelassen oder als nebenläufiges "Wunschdenken" behandelt. Wir wollen hier nur unter der Annahme des Modells (2.1) dieses Prinzip konkretisieren. Zunächst stellen wir fest, dass die einzige wesentliche Forderung in (2.1), die Annahme einer Verteilung für das Rauschen ist. Daher formen wir die Gleichung um in

$$\varepsilon_i = y_i - f(t_i), (i = 1, \ldots, n).$$

Weiterhin werden mit

$$r_i := y_i - \hat{f}(t_i),\ (i = 1, \ldots, n), r = (r_1, \ldots, r_n)^T,$$

die *Residuen* zu $\hat{f}$ bezeichnet. Damit kann der erste Teil der obigen Forderung äquivalent formuliert werden als:

- r sollte die *wesentlichen Merkmale* von ε aufweisen.

Die Residuen der Schätzfunktion $\hat{f}$ müssen also wie Weißes Rauschen aussehen. Ist nur das Modell (2.1) vorausgesetzt, so ergibt sich, dass jede Funktion, deren Residuen wie Weißes Rauschen aussehen, die wesentlichen Merkmale der Funktion f beinhaltet. Nehmen wir also an, es gäbe einen Orakel, das uns die Frage, ob ein Datensatz Weißes Rauschen sei, beantworten würde, so könnten wir zu jeder Funktion $\hat{f}$ entscheiden, ob sie als Schätzer für f die erste Forderung erfüllen würde. Was ist aber mit der zweiten Forderung "*keine weiteren*". Machen wir uns klar, dass die zweite Forderung notwendig ist. Es sei eine gute Approximation $(r, \hat{f})$, d.h. $\hat{f} \approx f$, $r \approx$ "Weißes Rauschen" und r_1 sei ein Datensatz, der ebenfalls wie Weißes Rauschen aussieht. Setzen wir

$$r_2 := (r + r_1)/2,$$

so wird r_2 ebenfalls wie Weißes Rauschen mit dem selben $\sigma > 0$ wie r aussehen. Setzen wir dann

$$\hat{f}_2 := y - r_2,$$

so beinhaltet $(r_2, \hat{f}_2)$ zwar alle Informationen über f, ist jedoch alles andere als eine erwünschte Approximation denn

$$\hat{f}_2 = y - (r + r_1)/2 = f + (2\varepsilon - r - r_1)/2 = f + r_3,$$

wobei r_3 wie Weißes Rauschen mit Varianz 2σ aussieht. Es ist also klar, dass wir die zweite Forderung an der Schätzfunktion $\hat{f}$ präzisieren müssen. An dieser Stelle nehmen wir die Tatsache, dass f eine eindimensionale Funktion ist, zur Hilfe (vgl. Kapitel 1). Bei eindimensionalen Funktionen hat man eine Vorstellung davon wie eine Funktion aussieht, wenn die Extremwerte bekannt sind. Die Frage lautet also: Wie können wir garantieren, dass die Extremwerte von $\hat{f}$, die Extremwerte von f gut approximieren? Folgender Vorschlag scheint hier sinnvoll zu sein (vgl. [15]):

- Wähle unter allen Funktionen, deren Residuen wie Weißes Rauschen aussehen, eine mit minimaler Anzahl von Extremwerten.

Diese Forderung scheint, im Gegensatz zur ersten, realisierbar zu sein. Denn, obwohl kein Orakel zur Verfügung steht, können wir die Hypothese "könnte ein Datensatz Weißes Rauschen sein" für große Datensätze gut überprüfen. Trotzdem ist die Lösung des Regressionsproblems mit diesem Ansatz vermutlich zu rechenintensiv und für die Praxis nicht geeignet.
In [15] werden konkrete Tests für die *Rauschidentifizierung* vorgeschlagen. Ferner werden zwei Methoden, die Run-Methode und die Taut-String-Methode entwickelt, die mit einem vertretbaren Rechenaufwand eine Schätzfunktion liefern, die den Anforderungen 1) und 2) aus dem Regressionsproblem gerecht werden.
In der Run-Methode wird der *Run-Test* (vgl. [20],[13]) verwendet, um eine sinnvolle Lokation der Extremwerte für $\hat{f}$ zu ermitteln. Desweiteren werden Grenzen[4] ermittelt, die jeder Schätzer für f einhalten sollte (vgl. Abbildung 2.4). Aus diesen Informationen gewinnt man dann eine Schätzfunktion, die die wichtigsten Merkmale von f aufweist.
Die Taut-String-Methode ist ein Weg, eine Schätzfunktion mit wenig Rechenaufwand zu bestimmen, die unseren Ansprüchen genügt. Zunächt wird das Rauscherkennungskriterium präzisiert. Hierfür kann man beispielsweise das *Summenkriterium* verwenden:
Ein Datensatz r sieht aus wie Weißes Rauschen, wenn

$$|\sum_{i=a+1}^{b} r_i| \leq c_n(a,b), (a,b \in \{1,\ldots,n\}),$$

gilt, wobei $c_n(a,b)$ aus der Verteilung des Rauschens ermittelt wird (vgl. [21]). In [15] werden nur 2-er Potenzen als a und b zugelassen, um die Rechenzeit zu verringern. Daher wird dieses Kriterium als *Mutiresolutionskriterium* bezeichnet. Betrachten wir nun unseren Datensatz y, bezeichnen mit Y den aufintegierten Datensatz

$$Y_i := \sum_{j=1}^{i} y_j(t_j - t_{j-1})$$

und betrachten

$$u = Y + \lambda,\ l := Y - \lambda, \lambda = (\lambda_1, \ldots, \lambda_n)^T \in \mathbb{R}^n_+.$$

Wird nun eine "straffe Saite" (oder "Taut-String") durch diese Grenzen gezogen, so minimiert die Ableitung dieser Saite, die Anzahl der Extremwerte unter allen Funktionen, deren

[4] Diese Grenzen kann man auch als Konfidenzbänder für f verstehen.

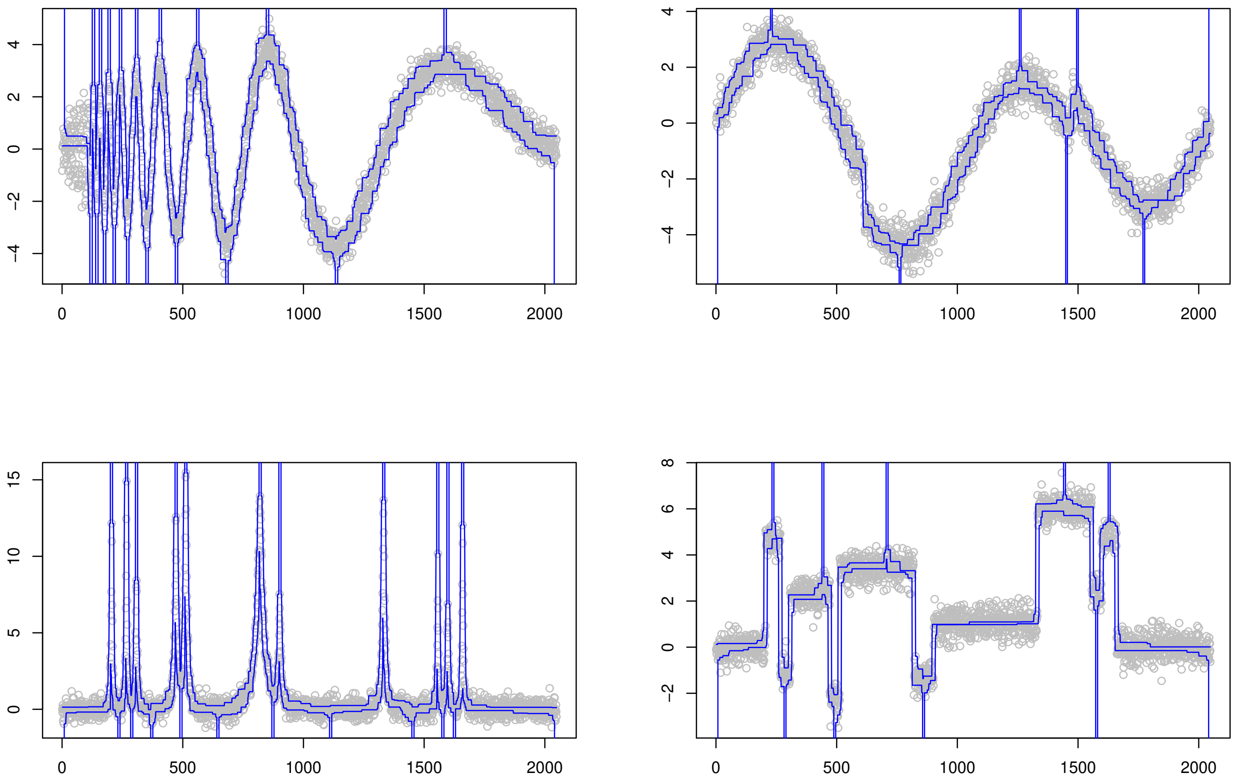


Abbildung 2.4: Das Ergebnis der Run-Methode angewendet auf die DJ-Datensätze. Die durchgezogenen Linien markieren die Grenzen, die eine “adäquate Approximation” von f einhalten muss.

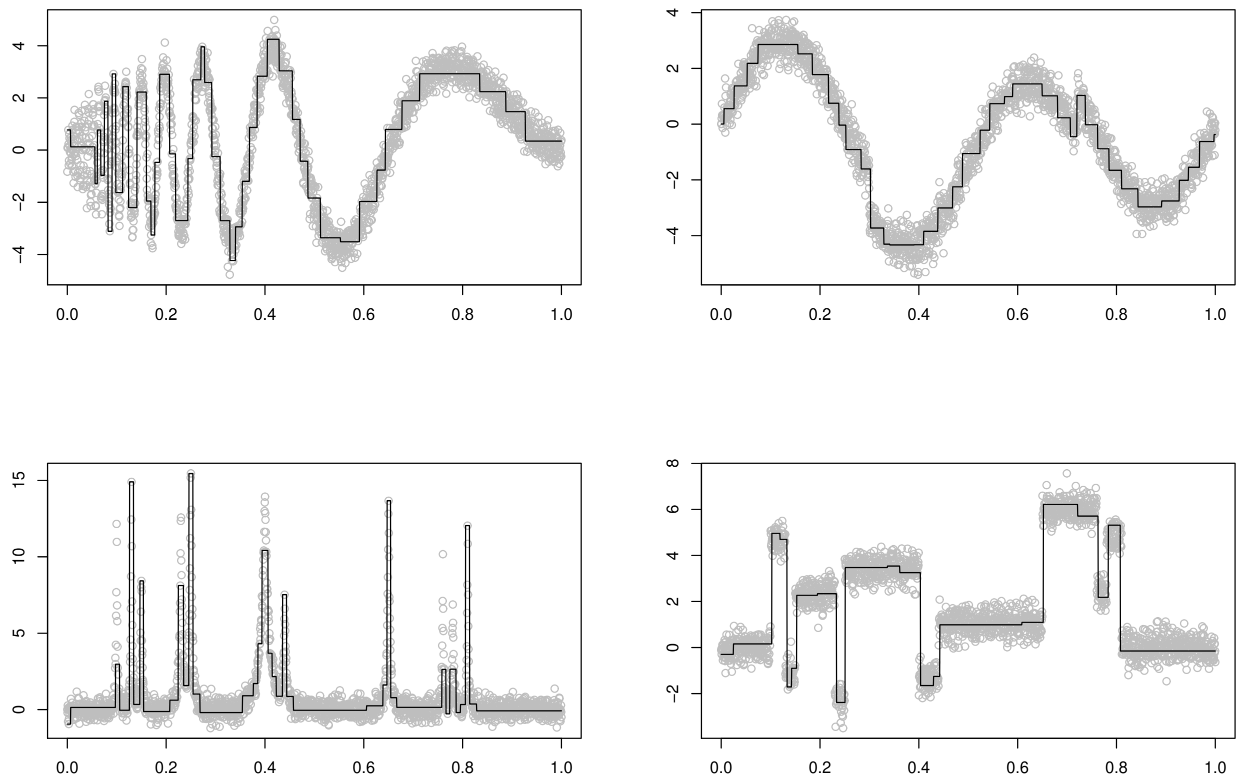


Abbildung 2.5: Die Schätzfunktionen, die sich aus der Run-Methode ergeben.

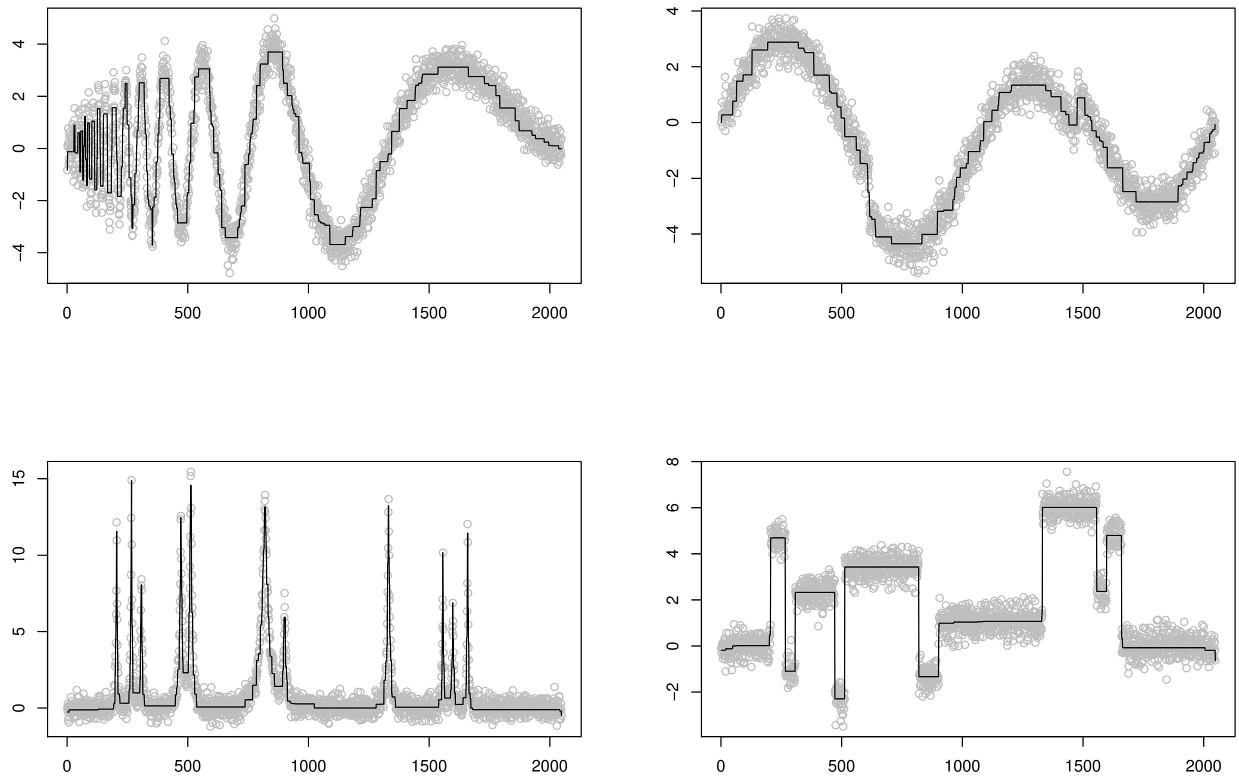


Abbildung 2.6: Das Bild zeigt das Resultat der Taut-String-Methode angewendet auf die DJ-Datensätze. Es gibt keine Artefakte, aber der Schätzer ist eine Treppenfunktion.

Stammfunktion zwischen l und u liegt. Wählt man nun λ geeignet, so liefert die daraus resultierende Saite, eine Schätzfunktion, die die “richtige” Anzahl von Extremwerten hat und die Multiresolutionsbedingungen erfüllt (vgl. [15]). Somit erfüllt der Taut-String beide Forderungen, die wir oben formuliert haben.

Es sei hier noch angemerkt, dass die Rauschidentifizierungsmethode für alle Datensätze einmal festgelegt wird und nicht jedesmal wieder neu gewählt werden muss, wie etwa der Parameter h bei den Kernschätzern. Somit braucht man für die Run-Methode und die Taut-String-Methode keinen Parameter zu wählen.

2.5 Zusammenfassung

Sowohl die Run-Methode, als auch die Taut-String-Methode bieten eine Lösung, die den kanonischen Anforderungen aus der Annahme des Modells (2.1) genügt (vgl. Abbildung 2.5 und Abbildung 2.6). Erweitert man nun die Annahmen im Modell (2.1), z.B. durch die Differenzierbarkeit von f, so stellt man fest, dass die resultierenden Schätzer beider Methoden, den neuen Ansprüchen nicht mehr genügen. Im nächsten Kapitel werden wir folgender Frage nachgehen: Gegeben das Resultat der Run-Methode oder der Taut-String-Methode, kann man eine glatte Funktion bestimmen, die die Vorzüge der ersten Schätzfunktion beibehält?

Kapitel 3

Das Glättungsproblem

Wir haben im letzten Kapitel verschiedene nichtparametrische Methoden kennen gelernt. Unter anderem haben wir die Run-Methode und die Taut-String-Methode vorgestellt, die zuverlässige Aussagen über die Anzahl und die Lokation der Extremwerte der Funktion f aus dem Problem (2.1) machen. Außerdem liefern beide Methoden Grenzen, die eine Schätzfunktion einhalten muss, um den Datensatz gut approximieren zu können. Ist f eine stetige oder sogar differenzierbare Funktion, so können wir mit dem stückweise konstanten Resultat der genannten Methoden nicht zufrieden sein.
Mit Hilfe der Erkenntnisse aus dem letzten Kapitel werden hier neue Erwartungen an einer Schätzfunktion für f formuliert. Wir suchen eine Funktion $\hat{f}$, die

1) den Datensatz y approximiert,

2) ein vorgegebenes Monotonieverhalten (bzw. Extremwertverhalten) besitzt,

3) möglichst glatt ist.

Im Folgenden beschreiben wir das Problem zunächst exakt. Danach werden Existenz und Eindeutigkeit der Lösung besprochen. Schließlich werden geeignete Diskretisierungen vorgestellt, um die Lösung zu berechnen.

3.1 Problemstellung

Gegeben sei der Datensatz y. Wir wollen eine möglichst einfache Funktion bestimmen, die den Datensatz y repräsentiert. Dazu wird das Standardmaß für die Glattheit definiert.

Definition 3.1 (Glättungsfunktional) Wir setzen

$$W_m := \{g : [0,1] \to \mathbb{R} | g \text{ m} \times \text{diff-bar und} \int_0^1 \left(g^{(m)}(t)\right)^2 dt < \infty\}$$

und

$$\|g^{(m)}\|_2 := \sqrt{\int_0^1 \left(g^{(m)}(t)\right)^2 dt}, (g \in W_m).$$

$\|.^{(2)}\|_2^2 : W_2 \to \mathbb{R}$ wird das *Glättungsfunktional* genannt.

Mit Hilfe dieses Funktionals können wir den Begriff "glatt" etwas genauer erklären. Gesucht ist also eine Funktion $\hat{f}$, die den Datensatz y approximiert, wobei $\|\hat{f}^{(m)}\|_2^2$ möglichst klein ist. Dabei interessiert uns hier hauptsächlich der Fall $m = 2$, da dieser Fall der intuitiven Vorstellung von dem Begriff "glatte Funktion" entspricht und daher in der Literatur am häufigsten verwendet wird. Wir müssen noch erklären was mit der "Approximation des Datensatzes" gemeint ist. Dazu wird in diesem Abschnitt angenommen, dass die Grenzen l, u und $t_{i_j}, (j = 1, \ldots, e)$ gegeben sind (z.B. von der Run-Methode oder der Taut-String-Methode), die eine *formerhaltende Approximation* der Daten garantieren. Das bedeutet, dass alle Funktionen die

1) $l_i \leq g(t_i) \leq u_i$ bzw. $l_i \leq G(t_i) \leq u_i$ (mit G Stammfunktion von g) erfüllen und

2) Extremwerte an den Stellen t_{i_j} haben,

die Struktur des Datensatzes y wiedergeben. Mit diesen Informationen können wir das Glättungsproblem exakt formulieren. Gesucht ist eine Funktion $\hat{f}$ mit

1) $l_i \leq \hat{f}(t_i) \leq u_i$ bzw. $l_i \leq \hat{F}(t_i) \leq u_i$,

2) $\hat{f}$ hat Extremwerte an den Stellen t_{i_j},

3) $\|\hat{f}^{(m)}\|_2^2$ ist minimal.

Im folgenden wird deutlich, dass die Lösung des Glättungsproblems (bei geeigneter Formulierung) eindeutig existiert. Zunächst geben wir einen Einblick in die Theorie und Anwendung von Splines. Insbesondere sind hier kubische Splines von Bedeutung. Wie wir später sehen werden, ist die Lösung des Glättungsproblems ein kubischer Spline.

3.2 Splines in der nichtparametrischen Regression

Splines gehören zu den fruchtbarsten mathematischen Hilfsmitteln im Zusammenhang mit dem Problem der Approximation.

Definition 3.2 (Spline) Zu einem Gitter $\xi = (t_1, \ldots, t_n)$ nennt man eine Funktion $s_\xi \in W_m$ einen *Spline k-ten Grades* $(k = m + 1, \ldots)$, falls s_ξ auf allen Intervallen $[t_i, t_{i+1}], (i = 1, \ldots, n-1)$ mit einem Polynom k-ten Grades übereinstimmt.

Einen Spline 3-ten Grades nennt man einen *kubischen Spline*. Der Ursprung der Splinetheorie liegt im Interpolationsproblem und im folgenden Satz.

Satz 3.3 (Holladay) Gegeben sei $\xi = (t_1, \ldots, t_n)$ mit

$$0 \leq t_1 < \cdots < t_n \leq 1 \text{ und } z_1, \ldots, z_n \in \mathbb{R}.$$

Es gibt genau einen kubischen Spline $s_\xi \in W_2$ mit

$$s_\xi(t_i) = z_i, (i = 1, \ldots, n) \text{ und } s_\xi^{(2)}(t_1) = s_\xi^{(2)}(t_n) = 0.$$

Die Optimerungsaufgabe

$$\|g^{(2)}\|_2^2 = \min$$

unter allen Funktionen $g \in W_2$ mit $g(t_i) = z_i$ $(i = 1, \ldots, n)$, wird durch s_ξ eindeutig gelöst. Wegen dieser Eigenschaft nennt man s_ξ einen *natürlichen* Spline.

Beweis:

Siehe z.B. [1].

■

Die Charakterisierung der Splines als Lösung von Minimierungsproblemen, hat mehrere Vorteile. In einigen Fällen kann man das Problem in einem allgemeineren Rahmen besser verstehen und die Lösung berechnen. Ferner ist die Verallgemeinerung auf kompliziertere Probleme oft einfacher, wenn man von einem Optimierungsproblem ausgeht (vgl. [11]). Hier werden nur natürliche kubische Splines betrachtet (d.h. Splines wie in Satz 3.3). Daher wird im folgenden auf die Präzisierung "natürlich" verzichtet. Wir bemerken nur, dass der Satz von Holladay sich auch auf höhere Ableitungen übertragen lässt. Weiterhin soll erwähnt werden, dass ähnliche Überlegungen wie in dieser Arbeit sich ebenfalls auf höhere Ableitungen übertragen lassen. Soll etwa die zweite Ableitung von f approximiert werden, so ist es sinnvoll höhere Ableitungen als Glättungsmaß zu betrachten (vgl. [55]). In der Statistik ist das Problem der Interpolation nicht von großem Interesse, da die Daten meistens verrauscht sind und eine Interpolation somit keine Vereinfachung des Modells bewirken würde. Mit Hilfe der Minimierungseigenschaft der Splines kann man jedoch allgemeinere Modelle angeben, die eine Approximation von f mit einer Glättung kombinieren. Wir geben hier einige Beispiele aus der nichtparametrischen Regression, weisen jedoch darauf hin, dass Splines in anderen Bereichen der Statistik, wie z.B. Zeitreihen, Dichteschätzung, etc. auch verwendet werden (vgl. [70], [65], [38], [16]).
Die Idee der *Smoothing-Splines* oder *Penalized-Splines* ist eines der ältesten Anwendungen der Splinetheorie in der Statistik. Dabei bestraft man den Least-Squares-Term $\sum_{i=0}^{n}(g(t_i)-y_i)^2$ mit einem Glättungsterm. Der Schätzer wir mittels

$$s_\lambda := \operatorname*{argmin}_{g\in W_2} \sum_{i=1}^{n}(g(t_i)-y_i)^2 + \lambda\|g^{(m)}\|_2^2, (\lambda > 0),$$

definiert. Dabei wählt man meistens $m \geq 2$. Falls man daran interessiert ist die Ableitungen der Funktion f zu schätzen, so wird $m > 2$ gewählt (vgl. z.B. [55]). Clark in [9] und Schönberg in [58] merken an, dass s_λ die Lösung des Minimierungsproblems

$$\|g^{(m)}\|_2^2 = \min \text{ s.t. } \sum_{i=1}^{n}(g(t_i)-y_i)^2 \leq U_0 \text{ [s.t. (= "subject to")]}$$

bzw.

$$\sum_{i=1}^{n}(g(t_i) - y_i)^2 = \min \ \text{s.t.} \ \|g^{(m)}\|_2 \leq g_0,$$

(mit U_0, g_0 geeignet) ist. Überlegen wir uns, dass die Interpolation, die Lösung des Problems

$$\|g^{(m)}\|_2^2 = \min \ \text{s.t.} \ \sum_{i=1}^{n}(g(t_i) - y_i)^2 \leq 0$$

ist, so wird deutlich, dass die Smoothing-Splines eine kanonische Anpassung des Interpolationsproblems an die Anforderungen der nichtparametrischen Regression darstellen. An dieser Art der Darstellung erkennt man besser, dass s_λ die glatteste Funktion ist, die sich in einer "Umgebung der Daten" befindet. Dabei ist $\sum_{i=1}^{n}(g(t_i) - y_i)^2$ offensichtlich ein grobes Maß für die "Nähe" zum Datensatz y. Eine Alternative bietet das Problem

$$\sum_{i=1}^{n}(g(t_i) - y_i))^2 + \lambda \mathrm{TV}(g^{(m)}) = \min,$$

wobei mit TV die Totale Variation bezeichnet wird (vgl. [47], [14]). Ist eine bessere Kontrolle der Funktion erwünscht, so betrachtet man das Problem:

$$\begin{aligned}
&\|g^{(m)}\|_2^2 = \min \ \text{s.t.} \\
&g \in W_m, \\
&l_i \leq g(t_i) \leq u_i, (i = 1, \ldots, n).
\end{aligned}$$

Für die statistische Interpretation nimmt man häufig an, dass die Fehler $\varepsilon_i, (i = 1, \ldots, n)$, einen kompakten Träger haben (vgl. [70]). Probleme dieser Art werden z.B. von [40] und [3] behandelt. Man kann sich leicht Situationen vorstellen, in denen es wünschenswert wäre, das Monotonieverhalten der Schätzfunktion ebenfalls zu kontrollieren. Dazu betrachtet man folgendes Problem:

$$\begin{aligned}
&\|g^{(m)}\|_2^2 = \min \ \text{s.t.} \\
&g \in W_m, \\
&l_i \leq g(t_i) \leq u_i, \ (i = 1, \ldots, n), \\
&g^{(1)}(t) \geq 0, \ (t \in [0, 1]),
\end{aligned}$$

was zu *monotonen* Splines führt (vgl. [4], [2], [69], [71]). Wright und Wegman betrachten in [69] etwas allgemeinere Probleme, die wir im folgenden ausführlicher erwähnen werden. Als Nächstes formulieren wir das Glättungsproblem exakt. Danach werden wir das Problem auf einem Gitter betrachten und Existenz und Eindeutigkeit überprüfen. Anschließend besprechen wir die Berechnung. Es sei zunächst ein *reguläres Gitter*

$$\xi = (t_1, \ldots, t_n) \text{ mit } 0 \leq t_1 < \cdots < t_n \leq 1,$$

gegeben. Außerdem sei $\mu = \mu_\xi : [0,1] \to \{1, -1\}$ eine Funktion, die das Monotonieverhalten vorschreibt, wobei 1 für *monoton wachsend* und -1 für *monoton fallend* steht. Wir setzen für $g \in W_m : (\mathcal{M}g) := g^{(1)}\mu$ und $(\mathcal{I}g)(x) := \int_0^x g(t)dt$ (der Integraloperator). Damit definieren wir das Glättungsproblem:

$$\begin{aligned} &\|g^{(m)}\|_2^2 = \min \text{ s.t.} \\ &g \in W_m, \\ &\mathcal{M}g \geq 0, \\ &l_i \leq (\mathcal{I}^k g)(t_i) \leq u_i (i = 1, \ldots, n), \end{aligned} \tag{3.1}$$

($k = 0$ oder 1). Dabei ist $\mathcal{M}g \geq 0$ komponentenweise zu verstehen. Wir interessieren uns hier hauptsächlich für den Fall $m = 2$ und setzen stets voraus, dass die Menge der zulässigen Funktionen für unsere Minimierungsaufgabe nicht leer ist. Dies stellt keine Einschränkung dar, da die Nebenbedingungen in einem ersten Schritt berechnet werden und daher keine leere Menge beschreiben. Hier werden maßtheoretische Probleme mit Nullmengen vernachlässigt und z.B. auf [5] verwiesen. Man sieht sofort, dass es Fälle gibt, in denen die Lösung nicht eindeutig ist. Hier nimmt man oft zusätzliche Bedingungen hinzu. Löwendick z.B. in [42], fixiert $g(0)$, $g(1)$, $g^{(p)}(0)$, $g^{(p)}(1)$, wobei $p = 1$ oder 2 sind. Diese Vorgehensweise kann, je nachdem wie die Werte $g(0), g(1)$ geschätzt werden, eine erhebliche Einschränkung bedeuten, wie man in der Praxis feststellt. Wir werden sehen, dass die Eindeutigkeit unter einfacheren Bedingungen gegeben ist, die bei uns immer erfüllt sind.

Definition 3.4 (Regularität) Wir betrachten ein Minimierungsproblem der Art (3.1) mit $m = 2$ und bezeichnen die Nebenbedingungen als *singulär*, genau dann, wenn es zwei verschiedene Geraden gibt, die zulässig sind (d.h. die Monotonie und Grenzbedingungen erfüllen). Das Problem heißt *regulär*, genau dann, wenn es nicht singulär ist.

Die *Regularität* ist zwar eine klar definierte Eigenschaft, jedoch nicht mit Hilfe der Nebenbedingungen beschrieben. Es folgt ein Beispiel für hinreichende Kriterien, die einfach nachgeprüft werden können.

Beispiel 3.5 Es sei das Problem (3.1) gegeben mit $\mu, l, u, k \in \{0, 1\}$. Die Nebenbedingungen sind regulär, falls gilt:

- μ ist nicht konstant und
  - k=0: es gibt i_1, i_2, so dass $u_{i_2} \leq l_{i_1}$
  - k=1: es gibt $i_1 < i_2 < i_3$, so dass

$$u_{i_2} \leq \min(l_{i_1}, l_{i_3}) \text{ oder } l_{i_2} \geq \max(u_{i_1}, u_{i_3}).$$

Beweis:
Ist μ nicht konstant, so sind die einzigen zulässigen Geraden, die konstanten Funktionen. Durch die zusätzlichen Bedingungen erreichen wir, dass höchstens eine konstante Funktion zulässig ist.

■

Satz 3.6 (Existenz und Eindeutigkeit)

1) Das Problem (3.1) besitzt eine Lösung $\hat{g} \in W_m$.
2) Im Fall $m = 2$ ist die Lösung eindeutig, genau dann, wenn die Nebenbedingungen regulär sind.

Beweis:

1) Es sei nochmals erwähnt, dass Nullmengen hier außer Acht gelassen werden! Betrachten wir die Menge $L_2 := \{g| \int_0^1 g(t)dt < \infty\}$, so ist $(L_2, <, >)$ mit

$$< h_1, h_2 >:= \int_0^1 h_1(t)h_2(t)dt,$$

ein Hilbert-Raum, mit der üblichen Äquivalenz auf Nullmengen (siehe z.B. [62]). Wir betrachten

$$C := \{h \in L_2 | \mathcal{I}^m h \text{ ist zulässig } \}.$$

Dabei sei $\mathcal{I}$ der Integraloperator. Die Menge C ist als Teilmenge von L_2 abgeschlossen und konvex. Daher nimmt die Normfunktion $\|.\|_2^2$ genau ein Minimum $\hat{h}$ auf C an. Wir setzen $\hat{g} := \mathcal{I}^m \hat{h}$, dann ist $\hat{g}$ ein Minimum des Minimierungsproblems. Diese Argumentationsfolge wird häufig benutzt (vgl. z.B. [64]).

2) Eindeutigkeit:

 – Singuläre Probleme haben offensichtlich unendlich viele Lösungen, nämlich

$$\alpha\Gamma_1 + (1-\alpha)\Gamma_2, (\alpha \in [0, 1]),$$

 wobei Γ_1, Γ_2 die zwei verschiedenen zulässigen Geraden sind.

 – Es seien $C \in \{C_0, C_1\}$, die Menge der zulässigen Funktionen für $k = 0$ bzw. $k = 1$ und $\hat{g}_1, \hat{g}_2$ seien zwei Lösungen des Problems

$$\min_{g \in C} \|g^{(2)}\|_2^2.$$

 Wir setzen $h := g_1 - g_2$. Aus der Konvexität von C folgt:

$$\|h^{(2)}\|_2 = 0.$$

 Hieraus ergibt sich $h^{(2)} \equiv 0$, also ist h eine Gerade.

Fall 1: Hat h zwei Nullstellen in $[0, 1]$, so muss $h \equiv 0$ sein und damit $g_1 = g_2$.

Fall 2: Hat h genau eine Nullstelle τ_κ auf $[0,1]$, so setzen wir

$$g_3 := 0.5g_1 + 0.5g_2,$$

dann ist g_3 ebenfalls eine Lösung des Problems (3.2).

Fall 2.1, k=0: g_3 berührt höchstens eine Grenzbedingung. Also ist g_3 die Lösung des Problems

$$\|g^{(2)}\|_2^2 = \min \text{ s.t. } \mathcal{M}g \geq 0, g(t_\kappa) = g_3(t_\kappa).$$

Damit folgt $\|g_3^{(2)}\|_2^2 = 0$, also ist g_3 eine Gerade. Wegen der Regularität gibt es höchstens eine zulässige Gerade, daher ist die Lösung hier ebenfalls eindeutig.

Fall 2.2, k=1: H hat folglich höchstens zwei Nullstellen $\tau_{\kappa_1}, \tau_{\kappa_2}$ in $[0,1]$. Also löst g_3 das entsprechende Problem:

$$\begin{aligned} &\|g^{(2)}\|_2^2 = \min \text{ s.t.} \\ &\mathcal{M}g \geq 0, \\ &G(\tau_{\kappa_1}) = c_1, G(\tau_{\kappa_2}) = c_2, \end{aligned}$$

mit $c_1 = G_3(\tau_{\kappa_1}), c_2 = G_3(\tau_{\kappa_2})$.

Fall 3: Hat h keine Nullstellen in $[0,1]$, so hat H höchstens eine Nullstelle τ_{κ_1} auf $[0,1]$. Damit löst g_3 das Problem

$$\begin{aligned} &\|g^{(2)}\|_2^2 = \min \text{ s.t.} \\ &\mathcal{M}g \geq 0 \end{aligned}$$

bzw.

$$\begin{aligned} &\|g^{(2)}\|_2^2 = \min \text{ s.t.} \\ &\mathcal{M}g \geq 0, \\ &G(t_{\kappa_1}) = G_3(\tau_{\kappa_1}). \end{aligned}$$

Dies liefert wiederum g_3 ist eine Gerade.

$\blacksquare$

Die Lösung des Glättungsproblems ist demnach nicht eindeutig, wenn wir Geraden als approximierende Funktionen zulassen. Diesen Fall kann man aber gesondert betrachten. Ist nämlich eine Gerade zulässig, so kann $\hat{f}$ durch einen parametrischen Ansatz, z.B. Least-Squares Gerade bestimmt werden. Die Bedingung $\mathcal{M}g \geq 0$ besteht aus überabzählbar vielen Ungleichungen. Die Frage ist nun, ob es eine geeignete, kleinere Teilmenge aus $[0,1]$ gibt, die ausreicht, um eine Formerhaltung der Lösung zu garantieren. Wegman und Wright beantworten diese Frage in einem Spezialfall (vgl. [70]). Wir notieren hier:

Satz 3.7 (Theorem 4.1 von Wright und Wegman [70]) Es sei das Problem (3.1) mit einem stetigen Funktional $M : W_2 \to W_1$ gegeben. Ferner existiere $g_1 \in W_2$ mit

$$Mg_1 > 0$$

und

$$l_i \leq g_1(t_i) \leq u_i, (i = 1, \ldots, n),$$

dann hat das Problem (3.1) eine Lösung, die ein kubischer Spline ist. Die Knotenmenge besteht aus $\{t_i | i = 1, \ldots, n\}$ und einer abzählbaren Teilmenge aus $[0,1]$.

Beweis:
vgl. [70]

■

Es gibt also eine kleinere Menge, die ausreicht, um in diesem Spezialfall die Monotonie zu garantieren. Da diese Knotenmenge trotzdem noch unendlich werden kann, können wir die formerhaltenden Nebenbedingungen nur in diskreter Form fordern. Machen wir diese Diskretisierung fein genug, so genügt das Resultat unseren Anforderungen.

3.3 Diskrete Nebenbedingungen

Wir wollen das Glättungsproblem (3.3) zunächst auf einem geeigneten Gitter betrachten. Es sei wieder ξ ein Gitter auf [0,1]. Um die Monotoniebedingungen zu diskretisieren, wollen wir die Forderung

"f ist monoton wachsend auf $I \subset [0,1]$"

ersetzen durch:

$$g(t_{i+1}) \geq g(t_i), (t_i \in I).$$

Dazu definieren wir folgendes Problem:

$$\begin{aligned}
&\|g^{(m)}\|_2^2 = \min \ \text{s.t.} \\
&g \in W_m, \\
&(g(t_{j+1}) - g(t_j))\, \mu_j \geq 0, (j = 1, \ldots, n-1), \\
&l_i \leq \mathcal{I}_n^k g(t_i) \leq u_i, (i = 1, \ldots, n).
\end{aligned} \tag{3.2}$$

Dabei sei $\mu_j = \mu(t_j), i = 1, \ldots, n$, die Funktion, die das Monotonieverhalten der Schätzfunktion vorschreibt, und $\mathcal{I}_n$ sei die diskrete Integration:

$$\mathcal{I}_n(t_i) := \sum_{j=1}^{i} g(t_j)(t_j - t_{j-1}), (i = 1, \ldots, n),$$

mit $t_0 := 0$. Wir wollen die Existenz und Eindeutigkeit der Lösung des Problems (3.2) beweisen. Hierfür beschreiben wir zunächst Eigenschaften der kubischen Splines, die eine einfache numerische Behandlung erlauben.

Lemma 3.8 (Kubische Splines) Zu $z_1, \dots, z_n \in \mathbb{R}$ gibt es genau einen natürlichen kubischen Interpolationsspline s_ξ (d.h. $s_\xi^{(2)}(t_1) = s_\xi^{(2)}(t_n) = 0$) auf einem regulären Gitter $\xi = (t_1, \dots, t_n)$. Wir setzen

$$\gamma_i := s_\xi^{(2)}(t_i), (i = 1, \dots, n), h_i := t_{i+1} - t_i, (i = 1, \dots, n-1),$$

dann ist s_ξ durch $\gamma = (\gamma_1, \dots, \gamma_n)^T$ eindeutig festgelegt. Es gilt:

- $R\gamma = D_2 z$

mit

$$R = \begin{pmatrix} \frac{1}{3}(h_1 + h_2) & \frac{1}{6}h_1 & & & \\ \frac{1}{6}h_1 & \frac{1}{3}(h_2 + h_3) & h_3 & & \\ \ddots & \ddots & \ddots & & \\ & \frac{1}{6}h_{n-3} & \frac{1}{3}(h_{n-3} + h_{n-2}) & \frac{1}{6}h_{n-2} & \\ & & \frac{1}{6}h_{n-2} & \frac{1}{3}(h_{n-2} + h_{n-1}) & \end{pmatrix}$$

und

$$D_2 = \begin{pmatrix} \frac{1}{h_1} & (\frac{-1}{h_1} + \frac{-1}{h_2}) & \frac{1}{h_2} & & & \\ & \frac{1}{h_2} & (\frac{-1}{h_2} + \frac{-1}{h_3}) & \frac{1}{h_3} & & \\ & \ddots & \ddots & \ddots & & \\ & & \frac{1}{h_{n-3}} & (\frac{-1}{h_{n-3}} + \frac{-1}{h_{n-2}}) & \frac{1}{h_{n-2}} & \\ & & & \frac{1}{h_{n-2}} & (\frac{-1}{h_{n-2}} + \frac{-1}{h_{n-1}}) & \frac{1}{h_{n-1}} \end{pmatrix}.$$

Für einen Spline u_ξ gilt:

$$\int_{t_1}^{t_n} (u_\xi^{(2)}(t))^2 dt = \gamma^T R \gamma,$$

wobei R eine symmetrische positiv-definite Matrix ist.

Beweis:

Wir berechnen nur die Matrizen:

$s_\xi^{(2)}$ ist eine stückweise lineare Funktion und es gilt:

$$s_\xi^{(2)}(t) = \gamma_i \frac{(t_{i+1} - t)}{h_i} + \gamma_{i+1} \frac{(t - t_i)}{h_i}, (t \in [t_i, t_{i+1}], i = 1, \dots, n-1).$$

Integrieren wir die Funktion zweimal, so folgt:

$$s_\xi(t) = \gamma_i \frac{(t_{i+1}-t)^3}{6h_i} + \gamma_{i+1}\frac{(t-t_i)^3}{6h_i} + A_i(t-t_i) + B_i, (t \in [t_i, t_{i+1}], i = 1, \ldots, n-1).$$

Setzen wir nun $z_i = s_\xi(t_i), z_{i+1} = s_\xi(t_{i+1})$ ein, so folgt:

$$z_i = \gamma_i \frac{h_i^2}{6} + B_i, z_{i+1} = \gamma_{i+1}\frac{h_i^2}{6} + A_i h_i + B_i, (i = 1, \ldots, n-1).$$

Schließlich bekommen wir:

$$B_i = z_i - \gamma_i \frac{h_i^2}{6}, A_i = \frac{z_{i+1}-z_i}{h_i} + h_i(\gamma_{i+1} - \gamma_i).$$

Investieren wir zusätzlich die Stetigkeit von $s^{(1)}$, so folgt:

$$\frac{h_i}{6}\gamma_i + \frac{h_i + h_{i+1}}{3}\gamma_{i+1} + \frac{h_{i+1}}{3}\gamma_{i+2} = \frac{z_{i+2}-z_{i+1}}{h_{i+1}} - \frac{z_{i+1}-z_i}{h_i}, (i = 1, \ldots, n-2).$$

Mit Hilfe des Satzes von Gerschgorin (vgl. z.B. [17]) sieht man, dass R positiv definit ist. Insbesondere ist R invertierbar.

$\blacksquare$

Satz 3.9 Es sei ξ ein reguläres Gitter, dann gilt:

1) Das Problem (3.2) hat eine Lösung s, die ein kubischer Spline auf dem Gitter ξ ist.

2) Sind die Nebenbedingungen regulär, so ist das Problem eindeutig lösbar.

3) Das Problem (3.2) lässt sich als ein quadratisches Optimierungsproblem auffassen.

Beweis:

1) Die Existenz einer Lösung sehen wir wie im Fall mit den stetigen Nebenbedingungen.

2) Es sei $s_{\xi,g}$ der natürliche Spline, der g auf dem Gitter ξ

$$(g(t_1), \ldots, g(t_n))$$

interpoliert. Für alle zulässigen Funktionen g gilt offensichtlich:

- Ist g zulässig, so ist auch $s_{\xi,g}$ zulässig.
- $\|s_{\xi,g}^{(2)}\|_2^2 \leq \|g^{(2)}\|_2^2$
- Insbesondere ist $\hat{g}$ eine Lösung des Problems, so gilt:

$$\|s_{\xi,\hat{g}}^{(2)}\|_2^2 \leq \|\hat{g}^{(2)}\|_2^2 \leq \|s_{\xi,\hat{g}}^{(2)}\|_2^2.$$

Wegen der Eindeutigkeit des natürlichen Splines folgt $\hat{g} = s_{\xi,\hat{g}}$.

Damit haben wir bewiesen, dass die Lösung ein Spline ist.

3) Sei C_ξ die Menge aller zulässigen Funktionen für das Problem (3.2). Folgende Probleme haben dieselbe Lösungsmenge:

- $\|g^{(2)}\|_2^2 = \min$ s.t. $g \in C_\xi$
- $\|s_{\xi,g}^{(2)}\|_2^2 = \min$ s.t. $g \in C_\xi$
- $\|s_{\xi}^{(2)}\|_2^2 = \min$ s.t. $s_\xi \in C_\xi$

Investieren wir noch Lemma 3.8, so lässt sich das Problem als ein diskretes QP auffassen.

■

Hier fragt man sich, ob und unter welchen Voraussetzungen, die diskreten Lösungen des Problems s_{ξ_n} gegen die Lösung des stetigen Problems (3.1) konvergieren, wobei ξ_n eine Folge von Gittern sei. Falls zum Beispiel die Gitter ξ_n äquidistant gewählt werden, so gilt $\cup\xi_n = \mathbb{Q}$, was eine dichte Teilmenge von $\mathbb{R}$ ist. Dieser Zusammenhang wird in [64] und [66] untersucht. Ferner könnte man mit Hilfe der exakten Lösungen eine bessere Charakterisierung des stetigen Problems erreichen. Wir wollen uns hier jedoch im weiteren Text der Berechnung der diskreten Lösung widmen.

3.4 Berechnung auf äquidistanten Gittern und QP

Wir haben im letzten Abschnitt das Problem (3.2) als die Lösung eines quadratischen Minimierungsproblems, unter linearen Nebenbedingungen aufgefasst. Solche Optimierungsprobleme sind gut untersucht worden. Trotzdem ist die Berechnung, bei größeren Datensätzen, sehr schwierig. Hier soll die Lösung des Minimierungsproblems berechnet werden. Dazu bemerken wir nochmal, dass die Matrix R aus dem Lemma (3.8), die das Verhältnis zwischen den Gewichten γ_i und die Werte $s(t_i)$ beschreibt, eine Tridiagonalmatrix ist. Bandmatrizen kann man bekanntlich auch in höheren Dimensionen effizient (in $O(n)$) behandeln. In der Nichtparametrischen Regression sind interessante Probleme oft mit großen Datensätzen verbunden. Daher soll die "Bandstruktur" des Problems hier erhalten werden. Wir Betrachten das Problem zunächst auf einem äquidistanten Gitter. Eine Möglichkeit, das Glättungsfunktional zu approximieren, ist der Vorschlag von [42] bzw. [50]: Approximiere die zweite Ableitung einer Funktion $g \in W_2$ durch die finiten Differenzen:

$$g^{(2)}(t_i) \approx n^2 \left(g(t_{i+1}) - 2g(t_i) + g(t_{i-1}) \right),$$

wobei $\xi = (1/n, 2/n, \ldots, (n-1)/n, 1)$. Mit der vorherigen Notation setzen wir also

$$g^n := (g(t_1), \ldots, g(t_n))^T \text{ und } g^{n,2} := nD_2 g^n,$$

wobei D_2 die Matrix aus Lemma 3.8 sein soll. Ferner wird angenommen, dass $\xi_1 = \xi_2 = \xi$ gilt. Damit kann $\gamma^T R \gamma$ in Lemma 3.8 durch

$$g^{n,2} R g^{n,2} = g_n Q_L g_n$$

mit

$$Q_L := (nD_2)^T R (nD_2) = \frac{n^3}{3} \begin{pmatrix} 4 & -7 & 2 & 1 & & & & & & \\ -7 & 16 & -10 & 0 & 1 & & & & & \\ 2 & -10 & 16 & -9 & 0 & 1 & & & & \\ 1 & 0 & -9 & 16 & -9 & 0 & 1 & & & \\ & \ddots & & & \ddots & & & \ddots & & \\ & 1 & & 0 & -9 & 16 & -9 & 0 & 1 & \\ & & 1 & 0 & -9 & 16 & -10 & 2 & & \\ & & & 1 & 0 & -10 & 16 & -7 & & \\ & & & & 1 & 2 & -7 & 4 & & \end{pmatrix}$$

approximiert werden. Eine zweite Möglichkeit ergibt sich aus der Matrixrelation $R\gamma = D_2 g_n$ in Lemma 3.8. Setzen wir das nämlich in $\gamma^T R\gamma$ ein, so folgt:

$$\gamma^T R\gamma = (R^{-1}D_2 g^n)^T R(R^{-1}D_2 g^n) = g^n D_2^T R^{-1} D_2 g^n.$$

Da R^{-1} keine Bandform hat, können wir nur die exakte Lösung berechnen und die Diagonalform beibehalten, indem wir ständig die Transformation $\gamma \leftrightarrow g^n$ durchführen. Dies würde den Algorithmus jedoch erheblich verlangsamen. Da viele Einträge von R^{-1} klein sind, bietet die folgende Vorgehensweise einen Kompromiss. Approximiere R^{-1} durch eine Tridiagonalmatrix, z.B.

$$R_1 := \frac{n}{5}\begin{pmatrix} 4 & -1 & & & \\ -1 & 4 & -1 & & \\ & & \ddots & & \\ & & -1 & 4 & -1 \\ & & & -1 & 4 \end{pmatrix}$$

und setze $\gamma^T R\gamma \approx g_n^T Q_M g_n$ mit

$$Q_M := D_2^T R_1 D_2 = \frac{n^3}{5}\begin{pmatrix} 4 & -9 & 6 & -1 & & & & & \\ -9 & 24 & -22 & 8 & -1 & & & & \\ 6 & -22 & 32 & -23 & 8 & -1 & & & \\ -1 & 8 & -23 & 32 & -23 & 8 & -1 & & \\ & \ddots & & \ddots & & \ddots & & & \\ & & -1 & 8 & -23 & 32 & -23 & 8 & -1 \\ & & & -1 & 8 & -23 & 32 & -22 & 6 \\ & & & & -1 & 8 & -22 & 24 & -9 \\ & & & & & -1 & 6 & -9 & 4 \end{pmatrix}.$$

Bemerkung 3.10 Die Matrizen Q_L und Q_M sind symmetrisch und positiv-semidefinit.

Die Approximation des Funktionals durch beide Matrizen lässt in der Praxis kaum Unterschiede feststellen. Auf eine genauere Analyse der Approximationseigenschaften wollen wir an dieser Stelle verzichten und werden im folgenden nur von $Q_2 \in \{Q_M, Q_L\}$ sprechen. Werden die Isotoniebedingungen und die Grenzbedingungen auch auf dem selben Gitter ξ gefordert, so kann das Problem mit diskretem Funktional wie folgt gestellt werden:

$$\begin{aligned}
&\|g^{(2)}\|_{n,2}^2 := \tfrac{1}{2} g^{nT} Q_2 g^n = \min \;\text{ s.t.} \\
&g^n \in \mathbb{R}^n, \\
&(g_{j+1}^n - g_j^n)\mu_j \geq 0, (j = 0, \ldots, n-1), \\
&l_i \leq \mathcal{I}^k g^n \leq u_i, (i = 1, \ldots, n).
\end{aligned} \tag{3.3}$$

Setzen wir $G_i^n = \sum_{j \leq i} g_j^n \frac{1}{n}$, so ist der Fall $k = 1$ äquivalent zu:

$$\begin{aligned}
&\tfrac{1}{2} G^{nT} Q_3 G^n = \min \;\text{ s.t.} \\
&\left(G_{j+1}^n - 2G_j^n + G_{j-1}^n\right) \mu_j \geq 0, (j = 0, \ldots, n-1), \\
&l \leq G^n \leq u,
\end{aligned} \tag{3.4}$$

mit $Q_3 := D_1^T Q_2 D_1$,

$$D_1 = \begin{pmatrix} 1 & -1 & & & & \\ & 1 & -1 & & & \\ & & \ddots & & & \\ & & & 1 & -1 & \\ & & & & 1 & -1 \end{pmatrix}.$$

Satz 3.11 Das Problem (3.3) ist lösbar. Die Lösung ist eindeutig, falls die Nebenbedingungen regulär sind.

Beweis:

Die Lösbarkeit ist klar. Die Eindeutigkeit folgt genau wie im stetigen Fall mit der folgenden Anmerkung. Aus

$$g^{nT} Q_2 g^n = 0,$$

folgt $D_2 g^n = 0$ und hieraus folgt g^n liegt auf einer Geraden.

■

3.5 Zusammenfassung

Wir haben hier das Glättungsproblem beschrieben und eine Diskretisierung vorgestellt. Dabei wurde sowohl das exakte Problem (3.2) auf einem diskreten Gitter, als auch das approximativ bestimmte Problem (3.3), bzw. (3.4) auf ein quadratisches Optimierungsproblem im $\mathbb{R}^n$, zurückgeführt. Im nächsten Kapitel werden Methoden vorgestellt, die solche Optimierungsprobleme lösen, und effiziente Methoden werden für unsere speziellen Optimierungsaufgaben entwickelt. Dabei steht die Möglichkeit, den Speicherverbrauch zu verringern, um große Probleme zu behandeln, im Vordergrund. Danach stellen wir die Ergebnisse vor und werden diese diskutieren.

Kapitel 4

Quadratisches Programmieren

In diesem Kapitel werden einige Methoden, zur Berechnung der Lösung eines quadratischen Optimierungsproblems[1] (QP) besprochen. Das Problem wird zunächst allgemein formuliert, um die Grundprinzipien der Methoden zu erläutern. Danach werden wir auf die Glättungsprobleme (3.3), bzw. (3.4) eingehen. Die Beschreibung der Methoden geschieht nicht detalliert, da hier nur ein grober Einblick in diesem Bereich ermöglicht werden soll. Zur näheren Beschreibung der Algorithmen verweisen wir zum Beispiel auf [54], [30], [31] bzw. auf das Buch von Nocedal und Wright [53].

4.1 Problemstellung

Wir betrachten hier einen Spezialfall eines Minimierungsproblems unter Nebenbedingungen. Das allgemeine Quadratische Programm wird formuliert als:

$$\begin{aligned}
&\min_x q(x) = \tfrac{1}{2}x^T Q x + x^T d \text{ s.t.} \\
&a_i^T x = b_i, (i = 1, \ldots, n_g), \\
&a_i^T x \geq b_i, (i = n_g + 1, \ldots, n_u),
\end{aligned} \tag{4.1}$$

wobei Q eine symmetrische $n \times n$ Matrix und $d, a_i \in \mathbb{R}^n, (i = 1, \ldots, n_u)$ sind. Hier sei stets $n_g = 0$ angenommen, was bedeutet, dass die Nebenbedingungen nur aus Ungleichungen

[1]Häufig wird hierfür der Ausdruck Quadratische Programmieren (quadratic programing) benutzt!

bestehen. Ferner nehmen wir an, dass Q positiv semidefinit ist. Machen wir uns klar, dass sowohl q, als auch die Menge der zulässigen Vektoren, konvex sind. In diesem Fall spricht man von einem *konvexem* QP.

Bemerkung 4.1

- Ein konvexes QP hat immer eine Lösung, falls die Menge der zulässigen Vektoren keine leere Menge ist.
- Die zulässige Menge definiert durch

$$\{x \in \mathbb{R}^n | Ax \geq b\}, A = [a_1, \ldots, a_u]^T, b = (b_1, \ldots, b_u)^T$$

  ist konvex (bzw. leer), wobei "$\geq$" komponentenweise zu verstehen ist.
- Wir definieren die Menge der aktiven Nebenbedingungen zu einem zulässigen Vektor x, mittels:

$$\mathcal{A}(x) = \{i | a_i^T x = b_i\}.$$

- Zu jeder Lösung x^* von (4.1) gibt es ein $\lambda^* = (\lambda_1^*, \ldots, \lambda_{n_u}^*)$, so dass folgende *Karush-Kuhn-Tucker* Bedingungen (KKT, vgl. [39]) erfüllt sind:

$$\begin{aligned} &Qx^* + d = \textstyle\sum_{i \in \mathcal{A}(x^*)} \lambda_i^* a_i, \\ &a_i^T x^* = b_i, i \in \mathcal{A}(x^*), \\ &\lambda_i^* \geq 0, (i = 1, \ldots, n_u). \end{aligned} \tag{4.2}$$

Beweis:
siehe [53].

■

Die Bedingungen (4.2) sind für die meisten QP Methoden von essentieller Bedeutung. Insbesondere sagen die *Lagrange Multiplikatoren* λ_i^* nicht nur etwas über das "aktiv sein" einer Bedingung aus, sondern auch über die Stärke dieser Bedingung (vgl. [53]). Wie bereits erwähnt, gibt es viele Vorschläge für die numerische Behandlung von (4.1). Ferner gibt es diverse Kombinationsmöglichkeiten dieser Methoden, die dazu führen, dass man für jedes Problem die geeignete Methode konstruieren könnte. Diese Methoden könnte man in verschiedenste Kategorien einteilen. Hier werden zwei repräsentative Kategorien vorgestellt.

4.2 Aktive-Menge-Methoden

Die Menge $\mathcal{A}(x^*)$, die in Bemerkung (4.1) definiert wurde, ist die Menge der aktiven Nebenbedingungen an der Lösung x^*. Wenn man diese Menge kennen würde, dann liesse sich das (QP) zu

$$\min_x q(x) \text{ s.t. } (a_i^T x = b_i, i \in \mathcal{A}(x^*))$$

vereinfachen. Bei den *Aktive-Menge-Methoden* versucht man die Indexmenge $\mathcal{A}(x^*)$ zu finden. Dazu definiert man eine Folge von Indexmengen $\mathcal{W}_k \subset \{1, \dots, n_u\}$ und stellt zu x^k das Problem

- Minimiere die Funktion q unter alle Vektoren x mit $(a_i^T x = b_i, i \in \mathcal{W}_k)$.

Dabei soll $q(x^k)$ eine fallende Folge sein. Es gibt diverse Möglichkeiten $\mathcal{W}_k$ zu bestimmen. Geeignete Beispiele für diese Methode werden in der nächsten Sektion mittels Anwendung auf unser spezielles Problem vorgestellt:

4.2.1 Implementierung einer Aktive-Menge-Methode

Hier wird eine Aktive-Menge-Methode genauer beschrieben, um sie unserer konkreten Situation anzupassen. Dabei suchen wir nach einem Weg, den Speicherbedarf der Methode, so weit wie möglich, niedrig zu halten, um möglichst große Datensätze bearbeiten zu können. Zunächst gehen wir vom Problem (4.1) aus und definieren zu x^k und beliebigem $\mathcal{W}_k \subset \mathcal{A}(x^k)$ das dazugehörige Subproblem:

$$\begin{aligned} &\min_p q(x^k + p) \text{ s.t.} \\ &a_i^T(x^k + p) = b_i, (i \in \mathcal{W}_k). \end{aligned} \tag{4.3}$$

Algorithmus 4.2 (Aktive-Menge-Algorithmus für konvexes QP)

- Sei $x^0 \in \mathbb{R}^n$ zulässig. Wir setzen $\mathcal{W}_0 := \mathcal{A}(x^0)$.
- **While** (Konvergenz nicht erreicht)
- $p^k := \underset{p}{\operatorname{argmin}}\ q(x^k + p)$ s.t. $a_i^T(x^k + p) = b_i, (i \in \mathcal{W}_k)$
  - **if** (p^k=0)
    - Bestimme Lagrange Multiplikatoren λ^k
    - Ist x^k die Lösung: **STOP**
    - Ist x^k nicht die Lösung,
      so wähle $i \in \mathcal{W}_k$ mit $\lambda_i^k < 0$ und setze
      $i := \operatorname{argmin}\{j|\lambda_j^k < 0\}$
      $\mathcal{W}_{k+1} = \mathcal{W}_k \setminus \{i\}$ (Das geht wegen Bemerkung 4.1)
    - $k \leftarrow k + 1$
  - **else**
    - setze $\alpha^k = \min(\min\{\alpha > 0|A(x^k + \alpha p^k) \geq b\}, 1)(> 0)$
    - $x^{k+1} = x^k + \alpha_k p^k$
    - $\mathcal{W}_{k+1} = \mathcal{A}(x^k)$
    - $k \leftarrow k + 1$
- end **While**

Die Funktion q wird in jedem Schritt in einem Unterraum von $\mathbb{R}^n$ minimiert, der durch $\mathcal{W}_k$ definiert ist. Stößt man dabei auf "Widerstand" durch Nebenbedingungen außerhalb von $\mathcal{W}_k$, so werden sie hinzu genommen. Ist das Minimum des Subproblems erreicht, so handelt es sich bei x^k entweder um die globale Lösung, oder wir können (mindestens) eine Nebenbedingung entfernen. Man kann zeigen, dass der Algorithmus 4.2 in jedem Schritt eine Verbesserung liefert und daher gegen die Lösung des Problems (4.1) konvergiert und sogar in endlich vielen Schritten mit der exakten Lösung abbricht (vgl. [53]). Diese Eigenschaft ist jedoch bei großen Problemen nicht von Bedeutung. Die Auswahl von α^k und die Selektion von $\mathcal{W}_k$ wird in der Literatur häufig diskutiert (vgl. z.B.[10] oder [32]). Wir wollen jedoch auf die Feinheiten der Methode verzichten und widmen uns mehr dem Speicherverbrauch. Die Lösungen p^k der Zwischenprobleme in Algorithmus 4.2 sind Lösungen linearer Gleichungssysteme, die jedoch abhängig von der Anzahl und Art von $a_i, i \in \mathcal{W}_k$ groß und singulär sein können. Im Folgenden betrachten wir das Problem (4.3) mit den einfachsten Nebenbedingungen nämlich Grenzen $l_i, u_i \in \mathbb{R}$ an den Komponenten x_i. Wir werden sehen, dass sich die Lösungen der Subprobleme sehr einfach berechnen lassen. Danach werden wir diese Vereinfachung durch formerhaltende Bedingungen erweitern und sehen, dass durch kleine Adaptionen der Methode, die Monotonie bzw. Konvexitätsnebenbedingungen ebenfalls berücksichtigt werden können. An dieser Stelle sei angemerkt, dass das spezielle Problem der Minimierung unter den Nebenbedingungen $l \leq x \leq u$ in der Literatur oft behandelt wurde. Eine effiziente Modifizierung des Aktive-Menge-Algorithmus für große Matrizen bietet z.B. [63]. Wir formulieren nun eine einfache Umschreibung des Subproblems in Algorithmus 4.2 (siehe z.B. [53]). Danach werden wir sehen, dass diese Formulierung uns erlaubt, den Speicherverbrauch für die Aktive-Menge-Methode, zu reduzieren. An dieser Stelle sei angemerkt, dass diese Speicherreduzierung in anderen Zusammenhängen ebenfalls angewendet werden könnte. Das Minimieren unter Grenzbedingungen ist, wie bereits erwähnt, von großem Interesse in der numerischen Optimierung (vgl. [44], [45], [43], [63]).

Satz 4.3 Betrachten wir das spezielle Problem

$$\min_{x\in\mathbb{R}^n} q(x) \text{ s.t. } l \leq x \leq u,$$

wobei $l, x, u \in \mathbb{R}^n$. Ferner sei zu x^k mit $l \leq x^k \leq u$ eine Arbeitsmenge $\mathcal{W}_k$ gegeben. Wir definieren

$$\mathcal{B}_k := \{i \in \{1 \dots, n\} : x_i^k = l_i \text{ oder } x_i^k = u_i\}, w := |\mathcal{B}_k|, s = n - w,$$

und damit

$$P_k : \mathbb{R}^n \to \mathbb{R}^s, P_k((x_1, \dots, x_n)^T) := (x_{i_1}, \dots, x_{i_s})^T,$$

wobei $\{i_1, \dots, i_s\} = \{1, \dots, n\} \setminus \mathcal{B}_k$. Dann lässt sich das Problem (4.3) wie folgt umschreiben:

i) $\min\limits_p q(x^k + p)^T$ s.t. $(p_i = 0, i \in \mathcal{B}_k)$.

ii) $\min\limits_p \frac{1}{2} p^T Q p + {g^k}^T p + c$ s.t. $p = P_k^T P_k p$, wobei $g^k = Qx^k + d$ und $c = \frac{1}{2} {x^k}^T Q x^k$.

iii) $\min\limits_r \frac{1}{2} r^T A_k r + {r^k}^T r + c$, wobei $A_k = P_k Q P_k^T, r^k = P_k g^k$.

Beweis:
i) ist exakt das Problem 4.3 in diesem speziellen Fall. Die Zielfunktion in ii) entsteht durch einfaches Ausmultiplizieren von i). Zur Umschreibung der Nebenbedingungen sei für $p \in \mathbb{R}^n$ angemerkt:

$$\{p | p_i = 0, i \in \mathcal{B}_k\} = \{p | p = P_k^T P_k p\}.$$

Schließlich sieht man iii) wie folgt. Wir substituieren $r := P_k p \in \mathbb{R}^s$ und setzen dies in ii) ein. Damit ergibt sich für die Nebenbedingungen: $p = P_k^T r$. Setzen wir dies wiederum in die Zielfunktion ein, so ergibt sich

$$\frac{1}{2}(P_k^T r)^T Q (P_k^T r) + {g^k}^T (P_k^T r) + c = \frac{1}{2} r^T A_k r + (P_k g^k)^T r + c.$$

■

Ist die Matrix $A_k \in \mathbb{R}^{s\times s}$ invertierbar, so lässt sich die Lösung p^k des Problems (4.3) mittels $p^k = -P_k^T A_k^{-1} r_k$ berechnen. An dieser Darstellung der Subprobleme sieht man, dass für den Aktive-Menge-Algorithmus in jedem Schritt nur die Lösung eines regulären LGS erforderlich ist. Man sieht weiterhin, dass A_k symmetrisch ist und die Dimension von A_k höchstens n sein kann. Wir brauchen also in jedem Schritt nur eine $s \times s$ Matrix, $(s \leq n)$ zu invertieren. Zunächst definieren wir eine *Bandmatrix.*

Definition 4.4 (Bandmatrix) Eine Matrix $B = (b_{ij}) \in \mathbb{R}^{n\times n}$ heisst eine *Bandmatrix mit Bandlänge* $m \geq 1$, falls gilt:

$$b_{ij} = 0, (|i-j| \geq m).$$

Eine Diagonalmatrix ist demnach eine Bandmatrix der Bandlänge 1. Damit notieren wir:

Bemerkung 4.5 Seien Q bzw. q und A_k wie oben gegeben.

i) Ist Q symmetrisch, so ist A_k ebenfalls symmetrisch für alle k.

ii) Ist Q eine Bandmatrix der Bandlänge m, so ist A_k ebenfalls eine Bandmatrix der Bandlänge m.

iii) Ist die Matrix Q eine Bandmatrix mit der Bandlänge m, so benötigt man für die Durchführung des Aktive-Menge-Algorithmus höchstens $2m+4$ Vektoren der Länge n.

Beweis:

i) Klar: $A_k^T = (P_k Q P_k^T)^T = P_k Q^T P_k^T = P_k Q P_k^T = A_k$.

ii) Es gilt offensichtlich:

$$A_k = (a_{\mu\nu})_{1\leq\mu,\nu\leq s} = (q_{i_\mu i_\nu})_{1\leq\mu,\nu\leq s},$$

Außerdem gilt mit (o.B.D.A) $\nu \geq \mu$:

$$\#\{i_\mu, i_\mu + 1, \ldots, i_\nu\} \geq \#\{\mu, \ldots, \nu\}.$$

Damit folgt:

$$m \leq |\mu - \nu| \Rightarrow m \leq |i_\mu - i_\nu| \Rightarrow a_{\mu\nu} = q_{i_\mu i_\nu} = 0.$$

iii) Wir können Q und d in $m+1$ Vektoren speichern (da Q symmetrisch ist!). Ferner müssen wir in jedem Schritt $\mathcal{B}_k, x^k, A_k, p_k$ speichern. (Die Lagrange-Multiplikatoren müssen hier nicht explizit berechnet werden!). Damit ergibt sich $m+1+1+1+m+1 = 2m+4$.

■

Dieser "Trick" versetzt uns in die Lage mit wenig Speicher, den Algorithmus zu programmieren. Dabei muss man bei der Invertierung der Matrizen A_k, die Bandform des Problems ausnutzen. Dies kann zum Beispiel realisiert werden, indem man die Cholesky Zerlegung der Matrix benutzt. Die oben beschriebene Methode wird in der Literatur häufig verwendet. Die Erhaltung der Bandstruktur der Matrix wird jedoch nicht erwähnt, da man in der Optimierungsliteratur Programme für allgemeinere Methoden haben möchte. Im weiteren Text wollen wir die Erhaltung der Bandstruktur auf unsere Probleme verallgemeinern. Dabei ist offensichtlich die Rolle der Projektion P_k sehr wichtig. Wir notieren zunächst eine einfache Feststellung.

Bemerkung 4.6 Es sei $\kappa : \mathbb{R}^n \to \mathbb{R}$ eine Abbildung und $\Omega \subset \mathbb{R}^n$ ein Unterraum von $\mathbb{R}^n$ mit $\dim(\Omega) = s, (s \leq n)$ gegeben. Ferner sei $P : \mathbb{R}^n \to \mathbb{R}^s$ eine lineare Abbildung mit vollem Rang s und die Zeilen von P seien alle Elemente aus Ω. Es gilt:

i) $P|_\Omega$ ist eine Bijektion und $P|_\Omega^{-1} = P^T$, falls die Zeilen von P eine Orthonormalbasis von Ω bilden

ii) $\Omega = \{x \in \mathbb{R}^n | x = P|_\Omega^{-1} P x\}$.

iii) Das Problem $\underset{x \in \mathrm{R}^n}{\operatorname{argmin}}\, q(x)$ s.t. $(x \in \Omega)$ wird gelöst durch

$$P|_{\Omega}^{-1} \underset{r \in \mathrm{R}^s}{\operatorname{argmin}}\, \kappa(P|_{\Omega}^{-1} r),$$

wobei die Lösung auch eine Menge sein kann.

Beweis:

i) Klar durch Nachrechnen.

ii) Klar nach i) und Lineare Algebra.

iii) Folgt aus der Substitution $r := Px$.

■

Im Allgemeinen ist die Bestimmung von P mit großem Rechenaufwand verbunden. In speziellen Fällen wie z.B. im Fall der Grenzbedingungen ist es jedoch möglich, P einfach anzugeben und die Bandstruktur der Matrix zu erhalten. Nun werden wir versuchen in unserer speziellen Situation die Monotonie bzw. Konvexitätsbedingungen einzubauen, ohne die Bandlänge der Matrix zu vergrößern. Zunächst betrachten wir das Problem (3.3) bzw. (3.4). Da hier der diskrete Fall betrachtet wird, wollen wir die Unbekannten zur Vereinfachung mit x bezeichnen.

$$\begin{aligned}
&\min_x \frac{1}{2} x^T Q x + d^T x \text{ s.t.} \\
&l \leq x \leq u, \\
&(x_{i+1} - x_i)\mu_i \geq 0, (i = 1, \dots n-1).
\end{aligned}$$

Sei x^k zulässig. Das dazugehörige Subproblem (4.3) lautet:

$$\begin{aligned}
&\min_p \tfrac{1}{2}(x^k + p)^T Q (x^k + p) + d^T (x^k + p) \text{ s.t.} \\
&p_i = 0, (i \in \mathcal{B}_k), \\
&p_{i+1} = p_i, \text{ für einige } i.
\end{aligned} \tag{4.4}$$

Wir wollen die aktiven Monotoniebedingungen, wie folgt notieren. Ist eine Folge von Monotoniebedingungen gegeben, so notieren wir dies als

$$x^k \text{ konstant auf } I_j := [\kappa_j : \rho_j] = \{\kappa_j, \kappa_i + 1, \ldots, \rho_j\}.$$

Wir wollen die Menge

$$\{I_j | x^k \text{ konstant auf } I_j\}$$

mit $\mathcal{S}_k$ bezeichnen. An dieser Stelle kann man sich überlegen, dass die Einfachheit von Grenzen als Nebenbedingungen, an folgender Tatsache liegt. An einer Komponente, kann nur eine Bedingung aktiv sein. Das Berühren beider Grenzen an einer Komponente bedeutet nämlich, dass die Grenzen an diesem Punkt übereinstimmen. Man kann also zur Lösung des Problems (4.3) eine Bedingung weglassen. Fügen wir nun die Isotoniebedingungen hinzu, so kann es sein, dass an einem Punkt sowohl eine Isotoniebedingung, als auch eine Grenzbedingung aktiv ist, also z.B. $x_i^k = l_i$ und $x_i^k = x_{i+1}^k$. Im folgenden Lemma werden wir sehen, dass man diese Nebenbedingungen in geeigneter Weise aufspalten kann.

Lemma 4.7 Wir betrachten das Problem (4.4) und setzen

$$\Omega := \{p \in \mathbb{R}^n | p_i = 0 \ (i \in \mathcal{B}_k), p \text{ konstant auf } I_j, I_j \in \mathcal{S}_k\},$$

$$\mathcal{B}_k^1 := \mathcal{B}_k \cup \{[i_s : j_s] | [i_s : j_s] \cap \mathcal{B}_k \neq \emptyset\} \text{ und } \mathcal{S}_k^1 := \mathcal{S}_k \setminus \mathcal{B}_k^1.$$

Damit gilt:

i) $\mathcal{B}_k^1 \cap \mathcal{S}_k^1 = \emptyset$

ii) $\Omega = \{p \in \mathbb{R}^n | p_i = 0 \ (i \in \mathcal{B}_k^1), p \text{ konstant auf } [i_s : j_s] \in \mathcal{S}_k^1\}$.

Beweis:

i) Ist gemäß der Definition klar.

ii) Gibt es in einem konstanten Intervall eine aktive Grenzbedingung, so müssen alle Punkte in diesem Intervall festgehalten werden.

■

Wir schließen aus dieser Eigenschaft:

Folgerung 4.1
Ist das Problem (4.4) gegeben, so kann man zu jedem Subproblem (4.4), die Aktive Menge $\mathcal{A}(x^k)$ in zwei disjunkte Indexmengen $\mathcal{S}_k$ und $\mathcal{B}_k$ zerlegen, die die aktiven Monotonie bzw. Grenzbedingungen beschreiben.

Damit können wir, wie im Fall der Grenzbedingungen, eine Beziehung zwischen den einzelnen Komponenten x_i und die möglichen aktiven Nebenbedingungen herstellen. Wir beschreiben den Aktive-Menge-Algorithmus für unser spezielles Problem im Fall der Monotoniebedingungen. Doch zuvor wird eine vereinfachende Notation eingeführt. Wir wollen auch einelementige Intervalle I_j erlauben. Die freien Komponenten werden also als einelementige Intervalle aus $\mathcal{S}_k$ aufgefasst, so dass gilt: $\{1, \ldots, n\} = \mathcal{B}_k \dot{\cup} (\dot{\bigcup\limits_{I_j \in \mathcal{S}_k}} I_j)$.

Satz 4.8 (AMA für Monotoniebedingungen) Wir betrachten das Problem (4.4), wobei die Mengen $\mathcal{B}_k$ und $\mathcal{S}_k$ disjunkt sein sollen. Wir setzen $s := |\mathcal{S}_k|$ und definieren damit $P_k := P(\mathcal{B}_k, \mathcal{S}_k) : \mathbb{R}^n \to \mathbb{R}^s$ mit $(P_k x)_j = \frac{1}{\sqrt{|I_j|}} \sum\limits_{i \in I_j} x_i$. Damit gilt:

i) Das Problem (4.4) ist äquivalent zu: $\min\limits_{p \in \mathbb{R}^n} q(x^k + p)$ s.t. $p = P_k^T P_k p$.

ii) Ist r^* die Lösung von
$$\min_{r \in \mathbb{R}^s} \frac{1}{2} r^T A_k r + r_k^T r + c,$$
mit
$$A_k = P_k Q P_k^T, r_k = P_k(Qx^k + d), c = \frac{1}{2} x^{k^T} Q x^k$$
und p^* die Lösung von i), so gilt: $p^* = P_k^T r^*$.

iii) Ist Q eine Bandmatrix der Bandlänge m, so ist A_k ebenfalls eine Bandmatrix der Bandlänge m, und der AMA kann mit einem $O(n)$ Speicherverbrauch berechnet werden. Ferner lassen sich die Lagrange-Multiplikatoren in jedem Schritt mit einem Rechenaufwand von $O(n)$ berechnen.

Beweis:

i) Wir müssen zeigen, dass $P_k|_\Omega$ eine Bijektion ist. Dazu sei $x \in \Omega$ und $i \in I_j$. Es gilt:

$$(P_k^T P_k x)_i = \frac{1}{\sqrt{|I_j|}} \left(\frac{1}{\sqrt{|I_j|}} \sum_{\kappa \in I_j} x_\kappa \right) \underset{x \text{ konstant auf } I_j}{=} \frac{1}{|I_j|}(|I_j| x_i) = x_i.$$

Damit ist $P_k|_\Omega$ insbesondere bijektiv!

ii) Folgt durch die Substitution $r := P_k p$.

iii) Es sei $Q = (Q_1, \ldots, Q_n) = (q_{ij})_{1 \le i,j \le n}$, $P_k Q P_k^T = (h_{\kappa\rho})_{1 \le \kappa,\rho \le s}$, dann ist:

$$h_{\kappa\rho} = \frac{1}{\sqrt{|I_\kappa|}} \sum_{i \in I_\kappa} \left(\frac{1}{\sqrt{|I_\rho|}} \sum_{j \in I_\rho} Q_j \right)_i = \frac{1}{\sqrt{|I_\kappa||I_\rho|}} \sum_{i \in I_\kappa, j \in I_\rho} q_{ij}.$$

Nehmen wir $|\rho - \kappa| \ge m$ an, so folgt für alle i, j mit $i \in I_\kappa, j \in I_\rho$:

$$|i - j| \ge m \text{ und damit } q_{i_j} = 0.$$

Betrachten wir die Zerlegung wie in Lemma 4.7, so ist die Bemerkung über die Lagrange Multiplikatoren klar.

■

Wir notieren die entsprechenden Sätze für konvexe Nebenbedingungen. Dabei müssen wir mehrere Fälle unterscheiden. An dieser Stelle sei angemerkt, dass die meisten dieser Überlegungen auf beliebige Nebenbedingungen, die jedoch eine Bandstruktur haben sollten, übertragen werden können. Das heisst Nebenbedingungen, die z.B. der Form $c_1 x_i - c_2 x_{i+1} \ge 0$ sind. Wir wollen jedoch hier darauf verzichten, da diese Verallgemeinerung in dieser Arbeit nur unnötige Notationen verursachen würde. Wir notieren das Subproblem für die diskreten Konvexitätbedingungen.

$$\begin{aligned} &\min_p (x^k + p)^T Q (x^k + p) + d^T (x^k + p) \text{ s.t.} \\ &p_i = 0, (i \in \mathcal{B}_k), \\ &p \text{ linear auf } ([i_\rho : j_\rho] \in \mathcal{K}_k). \end{aligned} \tag{4.5}$$

Wir wollen außerdem einen *Intervallzug*

$$[i_\rho : j_\rho], \dots, [i_{\rho+\kappa} : j_{\rho+\kappa}]$$

mit

$$[i_\rho : j_\rho] \cup \dots \cup [i_{\rho+\kappa} : j_{\rho+\kappa}] = [i_\rho : j_{\rho+\kappa}]$$

als

$$[i_\rho : j_\rho : \dots : i_{\rho+\kappa} : j_{\rho+\kappa}]$$

notieren.

Lemma 4.9 Es sei Subproblem 4.5 mit $\mathcal{B}_k$ und $\mathcal{K}_k$ gegeben. Ferner seien wieder zur Vereinfachung einelementige Intervalle erlaubt

$$\Omega = \{p \in \mathbb{R}^n | p_i = 0 \ (i \in \mathcal{B}_k), p \text{ linear auf } [i_\rho : j_\rho] \in \mathcal{K}_k, \rho = 1, \dots, s\}.$$

Wir trennen wieder die Nebenbedingungen in der Form

$$\mathcal{B}_k^1 := \mathcal{B}_k \cup \{[i : j] \in \mathcal{K}_k || [i : j] \cap \mathcal{B}_k| \geq 2\}$$

und

$$\mathcal{K}_k^1 := \mathcal{K}_k \setminus \{[i : j] \in \mathcal{K}_k || [i : j] \cap \mathcal{B}_k| \geq 2\}.$$

Damit gilt $\mathcal{K}_k^1$ und $\mathcal{B}_k^1$ beschreiben die Aktive Menge, d.h.

$$\Omega = \{p \in \mathbb{R}^n | p_i = 0 \ (i \in \mathcal{B}_k^1), \ p \text{ linear auf } [i_\rho : j_\rho] \in \mathcal{K}_k^1, \ \rho = 1, \dots, s\}.$$

Beweis:
Ist ein Linearitätsintervall gegeben, und sind mehr als zwei p_i gleich Null, so können wir das ganze Intervall gleich Null setzen.

■

Wir wollen nun mit Hilfe dieser Aussage, den Algorithmus für das Problem mit den Konvexitätsbedingungen so implementieren, dass die Bandstruktur erhalten bleibt. Dazu werden wir wieder zunächst zur Vereinfachung der Notation, die diskrete Menge $[\rho : \kappa] \setminus \{\rho, \kappa\}$ mit $(\rho : \kappa)$ bezeichnen.

Satz 4.10 (AMA für Konvexitätsnebenbedingungen) Es sei das Subproblem (4.5) mit $\mathcal{B}_k$ und $\mathcal{K}_k$ wie im Lemma 4.9 gegeben. Es sei

$$\Omega = \{p \in \mathbb{R}^n | p_i = 0 \ (i \in \mathcal{B}_k), p \text{ linear auf } I, I \in \mathcal{K}_k\},$$

Wir setzen $s_1 := |\mathcal{K}_k| - |\{I_j | I_j \cap I_{j+1} \neq \emptyset\}|$. Ferner sei

$$\{i_1 \ldots, i_{s_1}\} = \{1, \ldots, n\} \setminus \{i | i \in (\kappa : \rho), [\kappa : \rho] \in \mathcal{K}_k\}.$$

Es sei

$$\{j_1, \ldots, j_s\} = \{i_1, \ldots, i_{s_1}\} \setminus \{i | \exists j \text{ mit } [i : j) \cap \mathcal{B}_k \neq \emptyset\}$$

Wir setzen $P_k := P(\mathcal{B}_k, \mathcal{K}_k) : \mathbb{R}^n \to \mathbb{R}^s$ als die Projektion auf die Komponenten $\{j_1, \ldots, j_s\}$.

i) $P_k|_\Omega$ ist eine Bijektion und dim(Ω)= s.

ii) $P_k|_\Omega^{-1}$ ist definiert durch $P_k|_\Omega^{-1} x = z$ mit

- $z_{j_\kappa} = x_\kappa, \kappa = 1, \ldots, s$
- Für $\kappa = s_1, s_1 - 1, \ldots$ definieren wir rekusiv zu

$$i_\kappa \in \{i_1, \ldots, i_{s_1}\} \text{ mit } \exists j, \tau_\kappa := [i_\kappa : i_{\kappa+1}) \cap \mathcal{B}_k \neq \emptyset :$$

  $z_{i_\kappa} := \frac{(i_\kappa - \tau_\kappa)}{(i_{\kappa+1} - \tau_\kappa)} z_{i_{\kappa+1}}$.
- Definiere für $i \in (i_\kappa : i_{\kappa+1}), z_i$ durch lineare Interpolation.

ii) Die Lösung von (4.5) lässt sich bestimmen durch:

$$p^* = P_k|_\Omega^{-1} r^*, r^* := \text{ argmin } \frac{1}{2} r^T A_k r + r_k^T r + c,$$

wobei $A_k = (P_k|\Omega^{-1})^T Q P_k|_\Omega^{-1}, r_k = (P_k|\Omega^{-1})^T (Qx^k + d), c = x^{k^T} Q x^k$.

iii) Ist Q eine Bandmatrix, so ist A_k ebenfalls eine Bandmatrix.

Beweis:

i) Wir müssen bemerken, dass eine zulässige Funktion entweder keine Berührung mit den Grenzen im Inneren hat oder genau eine. Falls sie keine Berührung hat, so kann man sie durch die beiden Endpunkte repräsentieren. Hat sie genau einen Punkt, so wird sie nur von einem Parameter repräsentiert. Somit bleiben s Parameter übrig!

ii) Hier müssen wir uns klar machen, dass die Bedingung der Linearität in einem Intervall und die Einhaltung einer Grenze innerhalb desselben Intervalls $[z_{i_\kappa} : z_{i_{\kappa+1}}]$ sich zusammenfassen lassen zu:

$$z_{i_\kappa} = \frac{(i_\kappa - \tau_\kappa)}{(i_{\kappa+1} - \tau_\kappa)} z_{i_{\kappa+1}}.$$

iii) $P_k|_\Omega^{-1}$ lässt sich in zwei Abbildungen $\mathbb{R}^s \to \mathbb{R}^{s_1} \to \mathbb{R}^n$ aufspalten. Die Abbildung $\mathbb{R}^{s_1} \to \mathbb{R}^n$ beschreibt die Interpolation in den entsprechenden Intervallen. Die Abbildung $\mathbb{R}^s \to \mathbb{R}^{s_1}$ überträgt eine Bedingung im Inneren eines Intervalls zu den Rändern. Beide Abbildungen erhalten die Bandform, was die Aussage insgesamt beweist!

■

Diese Sätze zeigen, dass wir, zumindest beim Speicherverbrauch, die spezielle Struktur unseres Problems nutzen können. Wie bereits erwähnt, ist hierbei die Wahl der geeigneten Projektion P_k sehr wichtig. Um die Rolle der P_k besser zu verstehen, machen wir die folgende Bemerkung.

Bemerkung 4.11 Betrachten wir das Problem (4.4). Sei nun Ω eine Menge, die durch aktive Nebenbedingungen definiert ist, dann ist für $x \in \mathbb{R}^n$, der Vektor $x^* := P^T P x$ die Lösung der Minimierungsaufgabe

$$\|p - x\|_2^2 = \min, (p \in \Omega).$$

Beweis:

Die Zeilen von $P = (p_1, \dots, p_s)^T$ bilden eine Orthonormalbasis für Ω. Laut Definition ist

$$x^* = \sum_{i=1}^{s} < x, p_i > p_i$$

also die Projektion von x in Ω.

$\blacksquare$

Demnach ist unser Bestreben, so einfach und rechengünstig wie möglich, eine Basis für Ω zu finden. Wir notieren die Implementierung des Aktive-Menge-Algorithmus für die oben angemerkten Spezialfälle.

Algorithmus 4.12 (AMA Implementierung) Es sei $Q \in \{Q_2, Q_3\}$ aus dem letzten Kapitel, ohne Einschränkung seien beide $n \times n$ Matrizen. Ferner sei $x^0 \in \mathbb{R}^n, k = 0$ zulässig. Berechne

$$\mathcal{W}_k = \mathcal{A}(x^k) = \{i : \text{ Form - oder Grenzbedingung aktiv in } i\}.$$

- **While**(keine Konvergenz)
  - Zerlege $\mathcal{W}_k$ in Mengen wie in Lemma 4.7 bzw. 4.9.
  - Bestimme $A_k = P_k Q P_k^T$ (ist invertierbar).
  - $r^k = -Qx^k$ $p^k = P_k|_\Omega^{-1} A_k^{-1}(P_k r^k)$
    - **if** $(p^k! = 0)$
      - $\alpha_k = \min(\text{ argmax } \{\alpha > 0 : x^k + \alpha p_k \text{ zulässig}\}, 1)$
      - $x^{k+1} = x^k + \alpha_k p_k$
      - Berechne $\mathcal{W}_{k+1} = \mathcal{A}(x^{k+1})$
      - $k \leftarrow k + 1$
    - **else**
      - Berechne die Lagrange-Multiplikatoren: $\lambda_1, \dots, \lambda_{2n}$, wobei $\{\lambda_i, i \leq n\}$ für die Grenzen und $\{\lambda_i, i > n\}$ für die Form-Nebenbedingungen sind.

- $j =$ argmin $\{i : \lambda_i = \min\{\lambda_\kappa\}\}$
- **if**$(\lambda_j \geq 0)$ **STOP**
- **else** $\mathcal{W}_{k+1} = \mathcal{W}_k \setminus \{j\}$
- $k \leftarrow k + 1$

- end **While**

Bemerkung 4.13 Der Algorithmus 4.12 liefert in jedem Schritt einen zulässigen Vektor. Das Glättungsfunktional wird in jedem Schritt kleiner. Die Methode bricht nach endlich vielen Schritten mit der Lösung als Ergebnis ab.

Beweis:
Dies ist offensichtlich der Algorithmus 4.2 in diesem speziellen Fall!

■

4.2.2 Aktive-Menge Methoden für den stetigen Fall

Wir geben hier eine Beschreibung für die Aktive Menge Methode mit stetigem Funktional. Das bedeutet, dass hier die Diskretisierung und die Annahme der äquidistanten Gitter nicht notwendig sind. Die Methode ist nicht so effizient wie die vorherigen, daher werden wir es bei dieser Beschreibung belassen und widmen uns im weiteren Text der Berechnung. Zunächst wollen wir die Lösung folgenden Minimimierungsproblems genauer beschreiben. Gegeben sei ein Gitter. Wir wollen das Glättungsfunktional $\|.^{(2)}\|_2^2$ minimieren, wobei die Funktionswerte an einigen Gitterpunkten vorgeschrieben werden, und an den anderen Stellen des Gitters wird gefordert, dass die Funktionswerte aufeinander folgender Gitterpunkte übereinstimmen sollen. Dies ist nämlich genau das entscheidende Subproblem, was in unserer Minimierungsaufgabe vorkommt.

Satz 4.14 (Kubische Splines mit konstanten Nebenbedingungen) Es sei $\xi = (t_1, \dots, t_n)$ ein Gitter auf $[0,1]$, $\mathcal{K} \subset \{1, \dots, n\}$ eine Indexmenge und $z = (z_1, \dots, z_n)^T \in \mathbb{R}^n$. Das Problem

$$\begin{aligned}
&\min_{g \in W_2} \|g^{(2)}\|_2^2 \text{ s.t.} \\
&g(t_i) = g(t_{i+1}), (i \in \mathcal{K}), \\
&g(t_i) = z_i, (i \in \{1, \dots, n\} \setminus (\mathcal{K} \cup \{i+1 : i \in \mathcal{K}\}).
\end{aligned}$$

ist eindeutig lösbar. Die Lösung ist ein kubischer Spline s mit

$$s^{(2)}(t_1) = s^{(2)}(t_n) = 0, s^{(3)}(t_i^-) = s^{(3)}(t_{i+1}^+), (i \in \mathcal{K}).$$

Dabei ist mit "-" die linksseitige und mit "+" die rechtsseitige Ableitung gemeint.

Beweis:

Wir berechnen für den Spline s und eine beliebige Funktion $g \in W_2$ (vgl. Lemma 3.8):

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} g^{(2)}(x)s^{(2)}(x)dx &= \left[g^{(1)}(x)s^{(1)}(x)\right]_{t_i}^{t_{i+1}} - \int_{t_i}^{t_{i+1}} g^{(1)}(x)s^{(3)}(x)dx \\
&= \left[g^{(1)}(x)s^{(1)}(x) - g(x)s^{(3)}(x)\right]_{t_i}^{t_{i+1}}
\end{aligned}$$

Hieraus folgt

$$\begin{aligned}
\int_{t_1}^{t_n} g^{(2)}(x)s^{(2)}(x)dx &= \sum_{i=1}^{n-1} \int_{t_i}^{t_{i+1}} g^{(2)}(x)s^{(2)}(x)dx \\
&= \left(g^{(1)}(t_n)s^{(2)}(t_n) - g^{(1)}(t_1)s^{(2)}(t_1)\right) \\
&+ \sum_{i=1}^{n-1} \left(g(t_{i+1})s^{(3)}(t_{i+1}^-) - g(t_i)s^{(3)}(t_i^+)\right)
\end{aligned}$$

Setzen wir an dieser Stelle für g, die Differenz $g_1 - s$ ein. Der erste Term verschwindet nun wegen der Forderung $s^{(2)}(t_1) = s^{(2)}(t_n) = 0$. In der Summe müssen wir zwei Fälle unterscheiden. Ist nun eine Interpolation an der Stelle t_i vorgegeben, so folgt aus $g_1(t_i) - s(t_i) = 0$, die Annulierung des entsprechenden Terms. Haben wir einen Gleichheitsbedingung, so erhalten wir für t_i, t_{i+1}:

$$g(t_i)s^{(3)}(t_i^-) + g(t_{i+1})s^{(3)}(t_{i+1}^-) - g(t_i)s^{(3)}(t_i^+) - g(t_{i+1})s^{(3)}(t_{i+1}^+).$$

Aus der Spline-Eigenschaft von s folgt $s^{(3)}(t_{i+1}^-) = s^{(3)}(t_i^+)$, was mit $g(t_{i+1}) = g(t_i)$ die Annulierung der beiden mittleren Terme ergibt. Es bleibt also

$$g(t_i)s^{(3)}(t_i^-) - g(t_{i+1})s^{(3)}(t_{i+1}^+),$$

was durch unsere Zusatzforderung ebenfalls verschwindet. Hieraus ergibt sich also für eine beliebige Funktion $g_1 \in W_2$, die die Interpolation und Gleichheitsbedingung erfüllt, gilt:

$$\|g_1^{(2)}\|_2^2 - \|s\|_2^2 = \|g_1 - s\|_2^2 \geq 0,$$

was unsere Behauptung beweist.

■

Mit Hilfe dieses Satzes kann man das Problem der Minimierung von $\|g^{(2)}\|_2^2$ unter Interpolationsnebenbedingungen und vorgegebenen konstanten Intervallen lösen. Damit ist der AMA-Schritt definiert. Insbesondere würde dies eine sinnvolle Ergänzung zu der Hybrid-Methode von [42] liefern.

Die Berechnung solcher Splines kann man ebenfalls durch die Lösung einer Bandmatrix bestimmen (ähnlich wie in Lemma 3.8). Damit ist ein AMA-Schritt erklärt. Wir können also die Bandform unseres Problems erhalten, ohne den Diskretisierungsschritt zu vollziehen. Leider wird die Methode trotzdem sehr langsam, da wir in jedem Schritt, die Transformation $s \to s^{(2)} \to s$ durchführen müssen. Daher verzichten wir an dieser Stelle auf genauere Ausführungen.

Bei allen geschilderten Methoden zur Lösung eines quadratischen Optimierungsproblems, benötigt man einen Startvektor. Dabei kann durch ungeeignete Wahl des Startvektors x^0, die Methode viele Iterationen benötigen, bis eine hinreichende Approximation der Lösung bzw. die Lösung selbst erreicht wird. Dieses Problem tritt in nahe zu jedem Aktive-Menge basiertem Optimierungsalgorithmus auf (vgl. [53]). Ein bewährtes Mittel um dieses Nachteil zu reduzieren ist die Kombination mit anderen Methoden. Als Nächstes geben wir eine

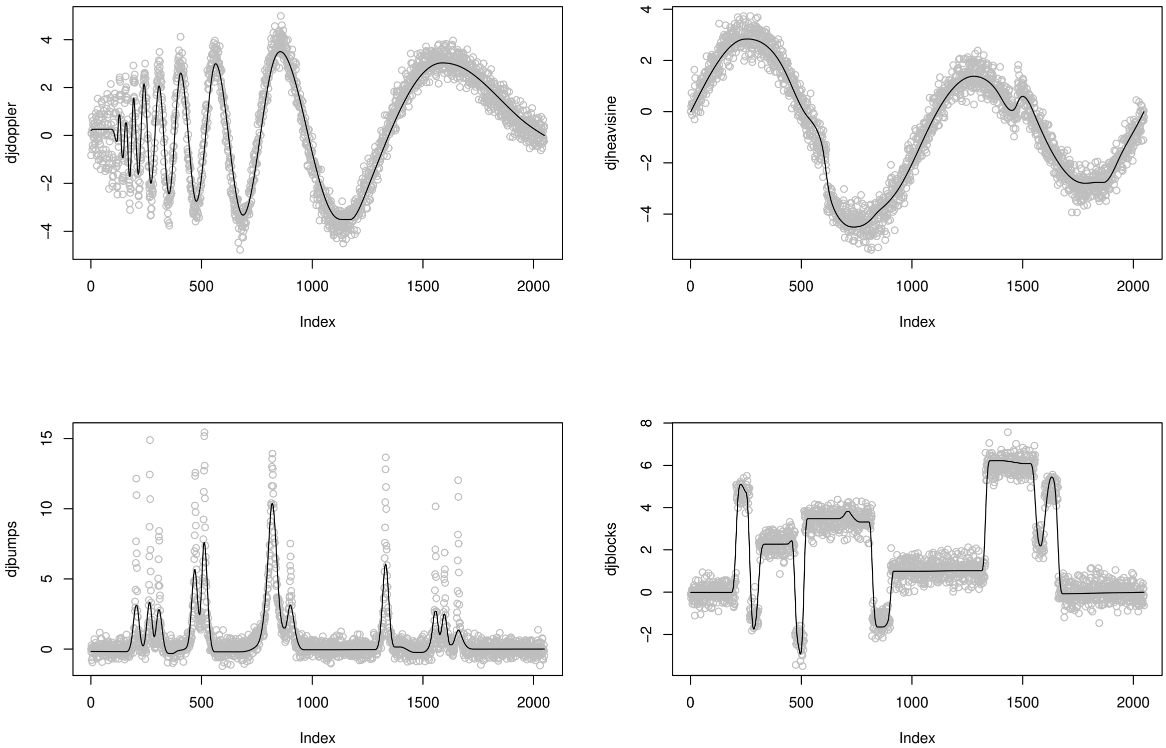


Abbildung 4.1: Das Resultat der **AMA** Methode, angewendet auf den DJ-Datensätzen unter Einhaltung der Run-Grenzen.

kurze Einführung in Gradientenbasierten Methoden. Als eine Art Spezialfall, stellen wir dann das QSOR vor und werden kurz auf diese Methoden eingehen. Danach werden wir zusammenfassend die praktischen Ergebnisse der Methoden vorstellen.

4.3 Gradient-Projektion-Methoden

Für bestimmte Probleme eignen sich die sogenannten Gradient-Projektion-Methoden. Diese Methoden sind meistens durch folgende Iteration erklärt:

- Bestimme den Gradienten p^k an der Stelle x^k.
- Finde α_k geeignet, so dass $x^k + \alpha_k p^k$ zulässig bleibt und setze

$$x^{k+1} := x^k + \alpha_k p^k.$$

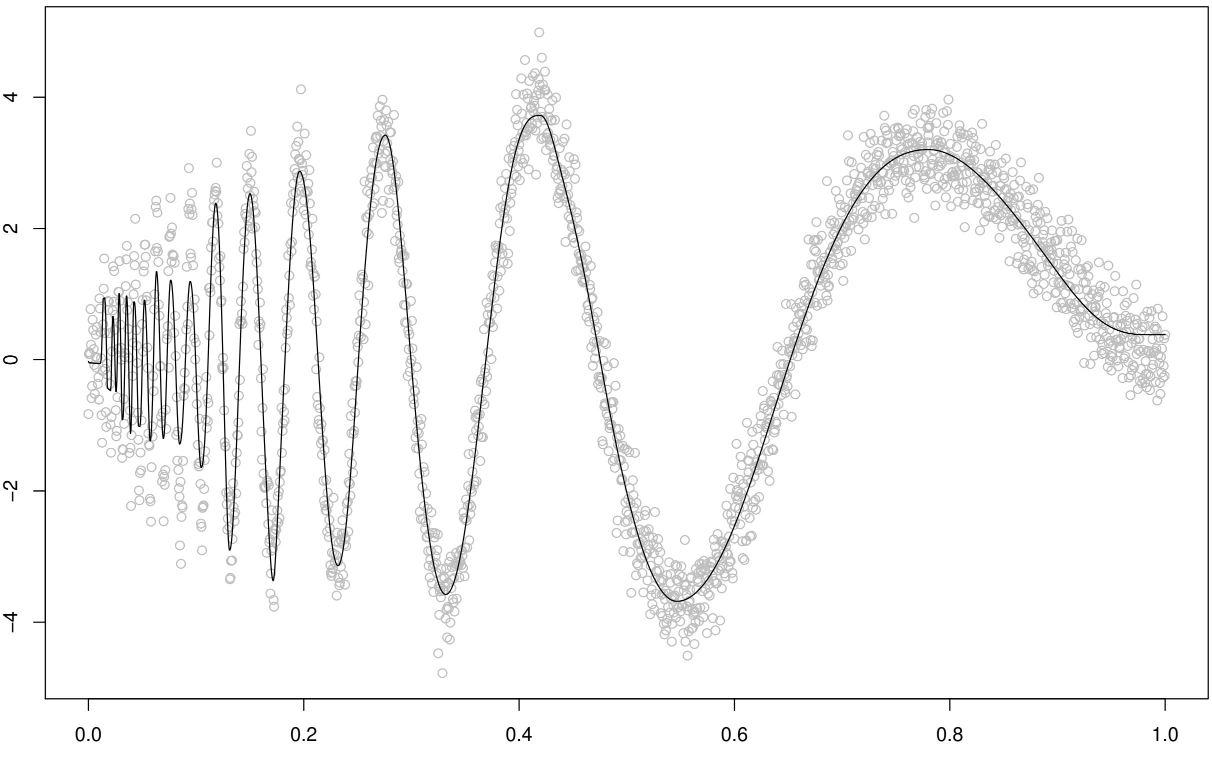


Abbildung 4.2: Das Resultat der **AMA** Methode, angewendet auf den Doppler-Datensatz unter Berücksichtigung der Taut-String-Bedingungen.

4.3.1 Implementierung einer Gradient-Projektion-Methode

Wir betrachten wieder das Problem (4.1) mit $n_g = 0$, d.h.

$$\min_{x \in \mathbb{R}^n} q(x) \text{ s.t. } Ax \geq 0,$$

Gradienten basierte Methoden sind iterative Methoden, die beginnend mit einem beliebigen Sartvektor x^0, die Lösung des Optimierungsproblems approximieren, indem sie in jedem Schritt durch Addition des Gradienten eine Minderung von q bewirken. Hierbei muss man wieder bedenken, dass in jedem Schritt die Zulässigkeit von dem aktuellem x_k garantiert werden muss. Genauer sind solche Methoden durch die Iteration

$$x^{k+1} := P(x^k - \alpha_k \nabla q(x^k))$$

gegeben wobei P die Projektion in den Zulässigkeitsraum sein soll, d.h.

$$P(x) := \text{argmin } \{\|z - x\| : Az \geq 0\}.$$

Man beachte, dass P wohldefiniert ist. Dabei ist $\|.\|$ eine Norm auf $\mathbb{R}^n$. Die Konvergenzkriterien für diese Methode wurden in der Literatur intensiv untersucht. In [7] wird die Konvergenz unter einfachen Bedingungen an q und $\|.\|$ bewiesen. Von der Wahl von α_k und $\|.\|$ hängt die Konvergenzgeschwindigkeit der Methode ab. [33] und [41] beweisen z.B.: ist ∇ Lipschitz stetig mit Konstante κ, so konvergiert die Methode, falls ein $\epsilon \in (0,1)$ existiert, mit:

$$\epsilon \leq \alpha_k \leq \frac{2}{\kappa}(1 - \epsilon).$$

Dabei ist es wichtig zu beachten, dass die Projektion P, häufig durchgeführt wird und daher nicht rechenintensiv sein sollte.

Beispiel 4.15 Es sei das Problem

$$\min_x q(x) \text{ s.t. } l \leq x \leq u$$

betrachtet. Ferner wählen wir für die Projektion $P_A = P_{[l,u]}$ die $\|.\|_2$ Norm in $\mathbb{R}^n$. Dann gilt:

$$P_A(x) = \text{median } \{l, x, u\},$$

wobei der Median hier auf die Komponenten angewendet wird.

In diesem Beispiel, kann man die Projektion in die Zulässigkeitsmenge sehr schnell ausrechnen. Nimmt man die Monotoniebedingungen hinzu, so wird die Projektion in die Menge der aktiven Nebenbedingungen zu teuer. Die Projektion P_k aus dem letzten Abschnitt ist ein Beispiel für eine "teure Projektion". Dabei muss man beachten, dass die Zerlegungen 4.7 bzw. 4.9 ebenfalls einen "Durchlauf" benötigen und damit die Berechnung der Projektion verteuern.

4.3.2 QSOR und Multigrid

Im folgenden werden wir das QSOR (Quasi SOR, vgl. [52]) kennen lernen. Diese Methode wurde von [50] entwickelt und besitzt die Eigenschaft, dass eine einzige Iteration sehr rechengünstig ist. Die Konvergenz jedoch ist sehr langsam. Diese Methode wird in [42] vollständig mit einem Konvergenzbeweis besprochen. Wir geben hier nur eine kurze Beschreibung.

Algorithmus 4.16 (QSOR)
Starte mit einem zulässigen Vektor $x^0 \in \mathbb{R}^n$ und wähle $\omega \in (0, 2)$

- **While**(Keine Konvergenz)
  - **for**(i in 1:n)
    - setze $\hat{x}_i^{k+1} = \text{median}\{l_i, u_i, \left(x_i^k - (Qx^k)_i\right)/\omega\}$
  - Korrigieren alle Konstanten (bzw. Linearen) Intervalle.
  - $k \leftarrow k + 1$
- end **While**

Für eine genauere Beschreibung siehe [42] bzw. [46]. In [42] wird die Konvergenz für das Problem (3.3) bewiesen. In [46] wird der Beweis auf das Problem (3.4) übertragen. Trotz der geringen Kosten pro Iteration ist das QSOR in vielen Fällen nicht effizient. Insbesondere, wenn die gesuchte Funktion f sehr glatt ist, braucht das QSOR sehr lange, um eine befriedigende Approximation der Lösung zu liefern. Diese Langsamkeit ist für die meisten Gradientenbasierten Methoden typisch. In [46] wird die QSOR Methode mit einem *Multigrid Ansatz* verbessert. Leider kann man die Verbesserung nur in numerischen Tests feststellen. Daher werden wir diese Idee kurz skizzieren, um uns danach der Berechnung zu zu wenden.

4.3.3 Multigrid-QSOR

Die *Multigrid-Technik* oder Mehrgitterverfahren sind allgemeine Ansätze aus der Numerik, die eine Verbesserung vieler Iterationsalgorithmen bewirken. Der Ansatz lautet:

- Führe die Iteration auf verschiedene Gitter aus.
- Führe ein *Gitterwechsel* durch, falls die Iteration zu langsam wird.

Dieser Ansatz wird bereits seit Jahren zum numerischen Lösen von Partiellen Differentialgleichungen benutzt. Die Effizienz wurde theoretisch untersucht (vgl. [34]). Für eine Einführung in Multigrid-Techniken verweisen wir auf [6]. Dieser Ansatz wurde für unseren Fall in [46] auf das QSOR übertragen, was eine Verbesserung der Methode bewirkte. Diese Effizienzsteigerung beruht jedoch auf numerischen Tests und konnte nicht theoretisch untermauert werden. Daher begnügen wir uns hier mit dieser Beschreibung und verweisen auf [46].

4.4 Zusammenfassung

Wir haben einen Aktive-Menge-Algorithmus und einen Gradientenbasierten Algorithmus eingeführt. Beide Methoden haben die Eigenschaft, dass sie mit wenig Speicher $O(n)$ auskommen. Man macht aber folgende "Beobachtung". Bei den Aktive-Menge-Methoden benötigt man viele Iterationen, um die gewünschte Aktive-Menge zu finden, falls der Startvektor nicht gut gewählt worden ist. Die Gradienten basierten Methoden kommen jedoch

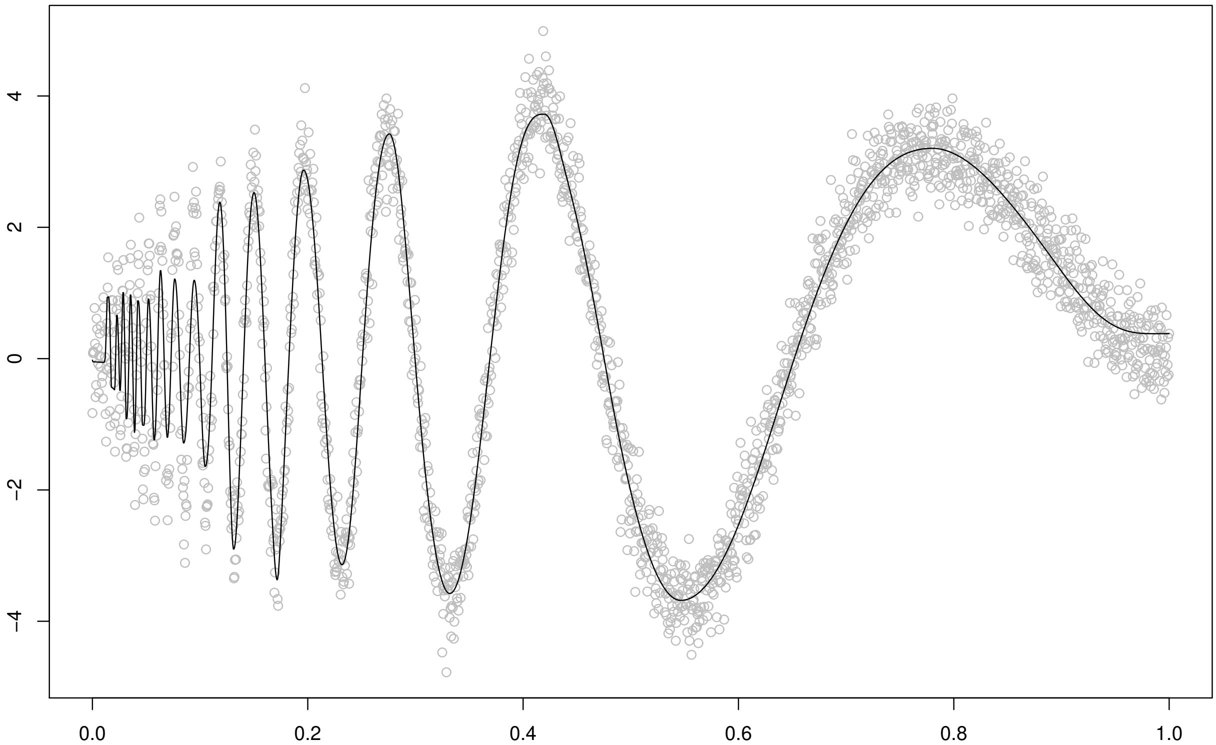


Abbildung 4.3: Das Resultat der **Multigrid-QSOR** Methode, angewendet auf den Doppler-Datensatz unter Einhaltung der Taut-String-Grenzen. Offensichtlich kann man mit dem Auge keinen Unterschied zwischen den Ergebnissen feststellen.

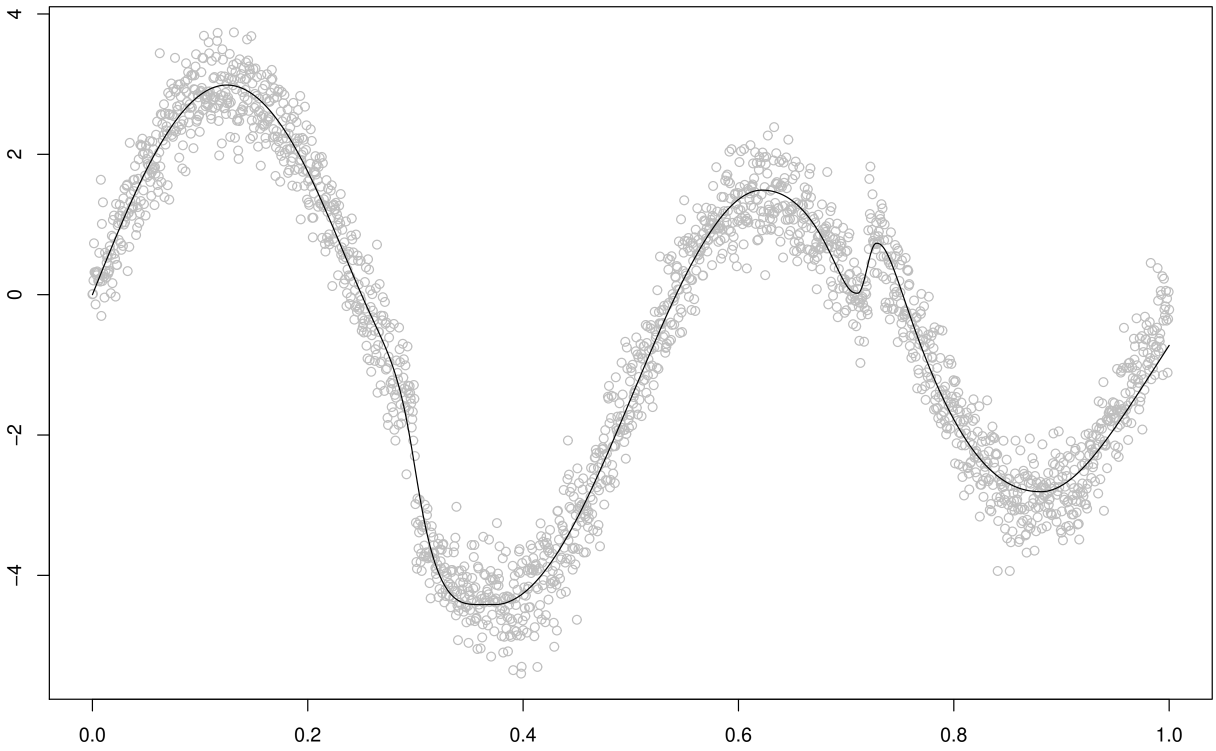


Abbildung 4.4: Das Resultat der **AMA** Methode, angewendet auf den Heavisine-Datensatz unter Einhaltung der Taut-String-Grenzen.

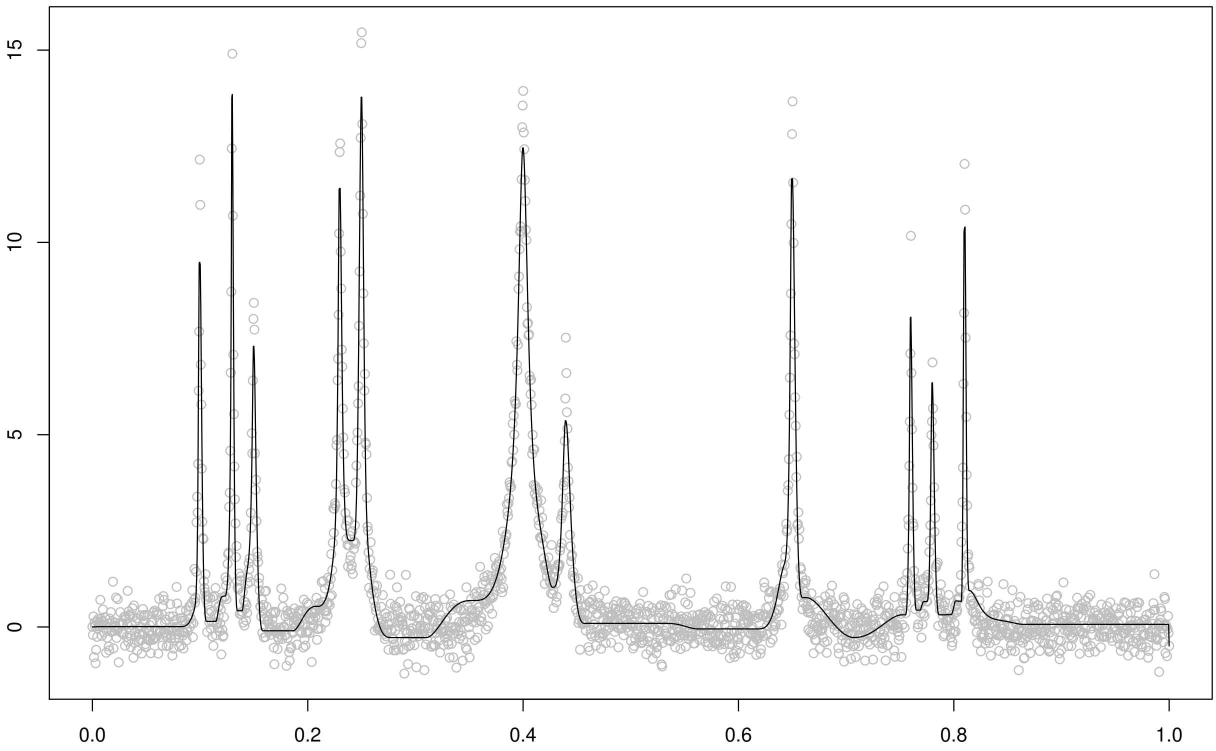


Abbildung 4.5: Das Resultat der **AMA** Methode, angewendet auf den Bumps-Datensatz unter Einhaltung der Taut-String-Grenzen.

sehr schnell in eine geeignete Umgebung der Lösung und werden danach extrem langsam (da in dieser Umgebung der Lösung, der Gradient sehr klein ist). In diesem Sinne ergänzen sich diese Methoden gegenseitig. Daher hat sich folgende Vorgehensweise in der Praxis bewährt.

i) benutze eine Gradienten Methode, bis sie langsam wird. Dies ist der Fall, wenn z.B. über viele Iterationen die Menge $\mathcal{A}(x^k)$ nicht ändert.

ii) benutze eine Aktive-Menge-Methode. Wechsle wieder zur Gradienten-Methode, falls $\|p^k\|$ und α_k zu klein werden, wobei p^k, α_k wie in (4.3) definiert sind.

Diese Vorgehensweise führt dazu, dass die Methode in vielen Fällen schneller als die einzelnen Methoden konvegiert. Folgende Tabelle zeigt die numerischen Ergebnisse für die vier DJ- Datensätze. Dabei steht in der Tabelle jeweils der Median der Rechenzeiten aus 101 Durchläufen[2] mit dem entsprechenden Datensatz. Ferner steht **MQSOR** für Multigrid-QSOR und **AMA** für den Aktive-Menge-Algorithmus. Außerdem muss man zu der Tabelle bemerken, dass das Resultat des **AMA**, wegen dem Abbruchkriterium, besser als das Resultat der **MQSOR** Methode war. Das Abbruchkriterium für die Kombination aus den beiden Mehoden ist jedoch derselbe, wie im **AMA**-Fall gewesen. Wir sehen, dass die Kombination in allen Fällen schnellere Ergebnisse liefert.

Tabelle der mittleren Rechenzeiten:

	MQSOR	AMM	Kombination
Doppler	51 sec.	68 sec.	50 sec.
Heavisine	65 sec.	66 sec.	59 sec.
Bumps	125 sec.	199 sec.	130 sec.
Blocks	260 sec.	390 sec.	320 sec.

[2]Die Durchläufe wurden alle auf einem Rechner mit einem AMD *Athlon* 1700XP Prozessor unter Einhaltung der Taut-String-Grenzen für Datensätze der Länge $n = 2048$ durchgeführt.

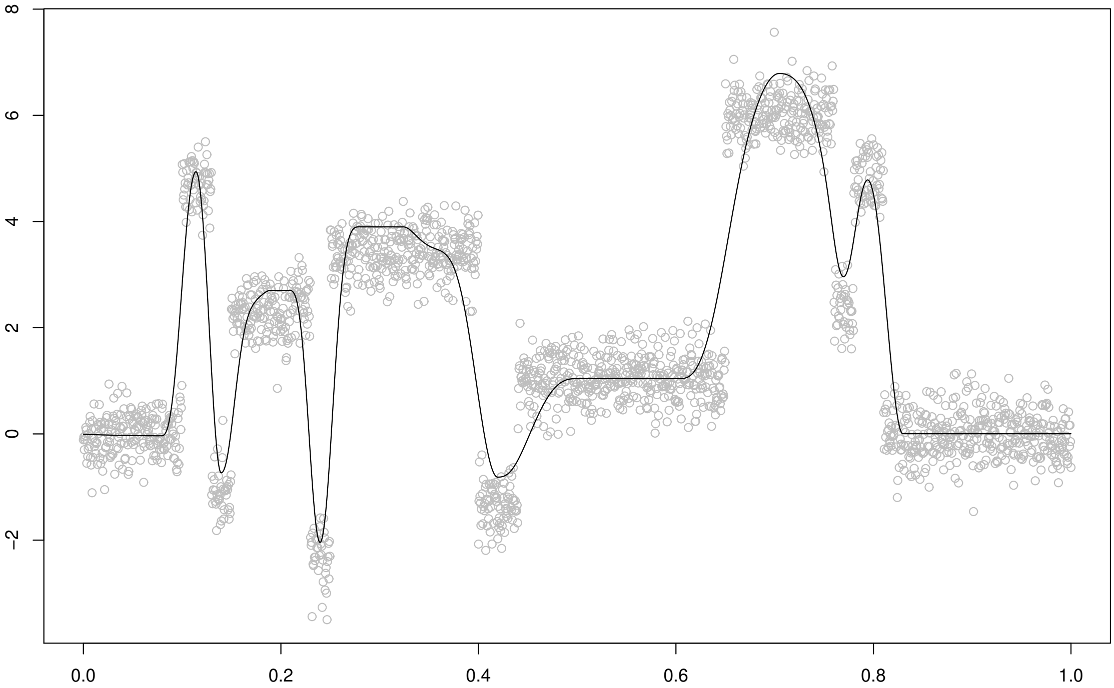


Abbildung 4.6: Das Resultat der **AMA** Methode angewendet auf den Blocks-Datensatz unter Einhaltung der Taut-String-Grenzen.

Kapitel 5

Rauscherkennungskriterium und Multiresolution

In den vorherigen Kapiteln haben wir Methoden zur Berechnung des folgenden Problems kennen gelernt:

- Bestimme die glatteste Funktion f, mit:
  - f hat das richtige Monotonieverhalten,
  - $l \leq \mathcal{I}^k f \leq u$.

Dabei wird das "richtige" Monotonieverhalten in einem ersten Schritt (z.B. durch die Run-Methode oder die Taut-String-Methode) berechnet. Im letzten Kapitel haben wir gesehen, dass die Bedingungen, die aus diesen Grenzen resultieren, manchmal (z.B. bei den Blocks-Daten) nicht genügen, um einen formerhaltenden Schätzer zu bekommen (vgl. Abbildung 5.1). Wir wollen zur Lösung dieses Problems im folgenden einen allgemeinen Ansatz zur *residuenbasierten nichtparametrischen Regression* vorstellen. Zunächst wird der Ansatz näher erläutert. Danach wird das asymptotische Verhalten eines konkreten Beispiels untersucht. Schließlich besprechen wir die Berechnung und stellen die Ergebnisse vor.

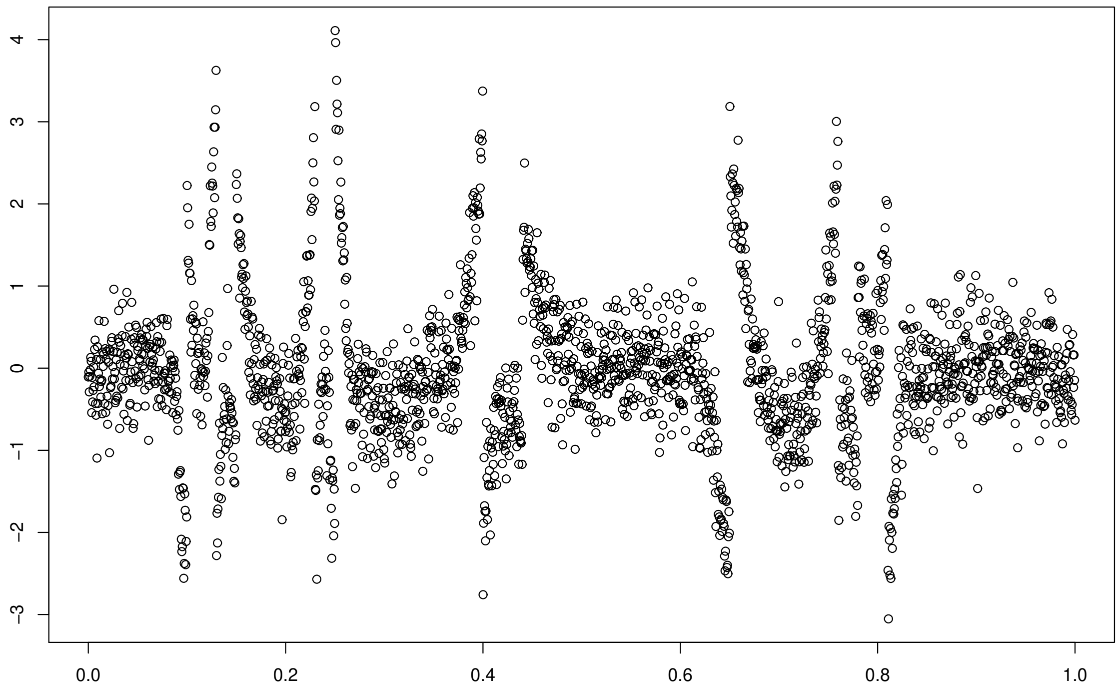


Abbildung 5.1: Die Residuen des glatten Schätzers für die Blocks-Daten. Offensichtlich sehen sie nicht wie Weißes Rauschen aus.

5.1 Nichtparametrische Regression, Residuen und Optimierung

Betrachten wir die Abbildung 4.4, so erscheint uns die Approximation der Daten nicht ausreichend. Die schlechte Qualität der Approximation kann man auf einem Test-bed, wo die wahre Funktion bekannt ist, auch mittels verschiedener Normen z.B. $\|.\|_\infty$ und $\|.\|_2$ feststellen. Das menschliche Auge scheint jedoch hier das geeignetste Maß für die Güte des Schätzers zu sein. Machen wir uns, klar warum einem Betrachter, die Approximation der Blocks-Funktion unbefriedigend erscheint. Wir wissen, dass unsere Daten aus einer einfachen Funktion und einem Rauschanteil bestehen. Daher sehen wir sofort, dass die Residuen $r_i = y_i - f_i$ nicht wie Rauschen aussehen können (vgl. 5.1).

Sehen die Residuen nicht wie Rauschen aus, so wird der Schätzer, egal welcher Art, keine gute Approximation der Funktion liefern können, da dieser dem Modell (2.1) nicht entsprechen kann. Daher können wir, abstrakt formuliert, die Forderung "die Residuen sehen aus wie Rauschen", mit höchster Priorität in die Schätzmethode aufnehmen. Natürlich versucht jede Methode diesem Anspruch gerecht zu werden. In den meisten Fällen spielt diese Forderung jedoch keine explizite Rolle. Die Taut-String-Methode (vgl. Kapitel 2) garantiert diese Eigenschaft durch die Multiresolutionbedingung und die Run-Methode durch die Runbedingung. Wir definieren allgemein:

Definition 5.1 (Rauscherkennungskriterium) Es seien $(\xi_n)_{n=1}^{\infty}$ äquidistante Gitter auf $[0,1]$. Ferner sei $\epsilon : [0,1]\times \rightarrow \mathbb{R}$ mit $,\epsilon(t) \sim \mathcal{N}(0,1)$. Eine Familie von Matrizen: $(T^k)_{k=0}^{\infty}$, $T^k = (T_1^k, \ldots, T_{n_k}^k)$ heißt (lineare) Rauscherkennungstransformation (ReT), falls gilt:

1) $\|T_i^k\|_2 = 1, (1 \leq i \leq n_k, k \in \mathbb{N})$

2) Für alle $\delta > 0$ gibt es ein $K_\delta \in \mathbb{N}$, so dass für $I \subset [0,1], |I| \geq \delta$ und $k \geq K_\delta$ ein $j \in \{1, \ldots, n_k\}$, und $c, C > 0$ existieren, mit:

   i) $(T_j^k)_\kappa = 0, (t_\kappa \notin I)$,

   ii) $c|I| \leq \frac{\#\{t_\kappa | (T_j^k)_\kappa \neq 0\}}{n_k} \leq C|I|$.

3) Es gilt: $\Pr\left(\max_{1\leq i\leq n_k} | < T_i^k, \epsilon(\xi_k) > | \leq \sqrt{\mathcal{T}\log(k)}\right) \rightarrow 1, (k \rightarrow \infty)$, wobei ϵ_{ξ_k} der Evaluationsvektor von ϵ in ξ_k sein soll.

Wir sagen Datensatz $r = \{r_1, \ldots, r_n\}$ erfüllt ein Rauscherkennungskriterium (ReK), falls es eine ReT $(T^k)_{k=0}^{\infty}$ gibt mit $\max_{1\leq i\leq n_k} | < T_i^k, r > | \leq \sqrt{\mathcal{T}\log(k)}$.

Beispiel 5.2 (ReK)

1) Die Identität $(E_n)_{n=1}^{\infty}$ ist keine ReT.

2) Das Summenkriterium für äquidistante Gitter, gegeben durch alle normierten Summen der Einheitsvektoren $e_i, (i = 1, \ldots, n)$ ist mit $\tau = 4\sigma^2$ ein ReK.

3) Die Multiresolutionskoeffizienten sind erklärt durch

$$w_{ij}(x) = 1/2^{i/2} \sum_{\kappa=2^j+2^i}^{2^{j+1}} x_\kappa, (\text{ für geeignete } i, j).$$

Die Mutiresolutionsbedingung für die Gitter

$$\xi_k := (1/2^k, \ldots, (2^k - 1)/2^k, 1), k \in \mathbb{N},$$

erklärt durch

$$\max |w_{ij}(x)| \leq \sqrt{2 \log(2^k)}\sigma,$$

ist ein ReK.

Beweis:

1) Die Eigenschaft 2) ist offensichtlich nicht erfüllt, da $\#\{t_\kappa | (T_j^k)_\kappa \neq 0\} = 1$, für beliebige k gilt.

2) Ist nach Definition klar.

3) Hier rechnen wir die Eigenschaft 2) nach. Zu $\delta > 0$ wählen wir K_δ so, dass $\delta/4 > 1/2^{K_\delta}$ gilt. Damit wählen wir zu $k \geq K_\delta, I \subset [0, 1], |I| \geq \delta$ ein i, so dass

$$2^k|I|/4 \leq 2^i \leq 2^k|I|/2.$$

Es gibt mindestens einen Intervall der Länge $2^i/2^k$, der vollständig in I liegt und damit die Voraussetzung erfüllt.

■

Hat man ein Rauscherkennungskriterium bestimmt, so kann man garantieren, dass für große n, die Residuen des Schätzers wie Rauschen aussehen.
Dabei ist zu beachten, dass die Wahl des Kriteriums, die Effizienz der Methode erheblich beeinflusst. Die Multiresolutionsbedingung ist für unsere Zwecke hier am besten geeignet und dazu leicht verständlich. Nun kann man Ansprüche an die Funktion f stellen. Sind diese Ansprüche konkret genug, so können wir das Problem als folgende Optimierungsaufgabe formulieren und lösen.

- Optimiere bzgl. eines Kriteriums unter der Bedingung, dass die Residuen wie Rauschen aussehen.

Beispiel 5.3 (Taut-String)

- Die Taut-String Methode-liefert einen effizienten Algorithmus für die Approximation der Lösung des Problems

  - Minimiere die Anzahl der lokalen Extremwerte von f unter der Nebenbedingung:
  $y - f$ genügt der Multiresolutionsbedingung.

Beispiel 5.4 (Wavelet-Thresholding) Sei $\mathcal{W} : \mathbb{R}^n \to \mathbb{R}^n$ eine orthonormale Transformation. Wir betrachten das Problem:

$$\min_{g \in \mathbb{R}^n} \sum_{i=1}^{n} g_i^2 = \min \text{ s.t. } |\mathcal{W}(g - y)| \leq c,$$

wobei $'' \leq''$ komponentenweise zu verstehen ist, und $c > 0$ sei eine Konstante. Setzen wir $w = \mathcal{W}f$, so genügt es, das Problem:

$$\min_{g \in \mathbb{R}^n} \|g\|_2^2 \text{ s.t. } \mathcal{W}y - c \leq w \leq \mathcal{W}y + c,$$

zu lösen, denn wegen der Orthonormalität gilt $\|g\|_2^2 = \|w\|_2^2$. Damit ist die Lösung des Problems $\hat{f} = \mathcal{W}^T \hat{w}$ mit

$$\begin{aligned} \hat{w}_i &= \{|(\mathcal{W}y)_i| > c\} \operatorname{sign}((\mathcal{W}y)_i) \ (|(\mathcal{W}y)_i| - c) \\ &= \begin{cases} (\mathcal{W}y)_i + c & ,((\mathcal{W}y)_i + c < 0) \\ (\mathcal{W}y)_i - c & ,((\mathcal{W}y)_i - c > 0) \\ 0 & ,\text{sonst.} \end{cases} \end{aligned}$$

Wählen wir für $\mathcal{W}$ eine diskrete Wavelet-Transformation, so ist diese Methode äquivalent zum Wavelet-Thresholding (vgl. [18]).

Im folgenden wird die Multiresolutionsbedingung betrachtet. Falls die Forderung an die Funktion f durch die Optimalität bezüglich eines Funktionals der Form $\|.\|_2^2$ gegeben ist, können wir den Schätzer mittels eines QP ausrechnen. In diesem Fall ist das Problem die Bestimmung der glattesten Funktion unter allen Funktionen, die der Multiresolutionsbedingung (mit Parameter $\sqrt{\tau \log(n)}\sigma_n$) genügen, wobei σ_n ein Schätzer für σ sein soll. Genauer:

$$\int_0^1 g^{(m)}(t)dt = \min \ \text{ s.t. } g - y \in \mathrm{S}(\xi_n, c_n), \tag{5.1}$$

wobei $c_n = \sqrt{\tau \log(n)}\sigma_n, \tau > 4$ und $\mathrm{S}(\xi_n, c_n)$ die Menge aller Funktionen, die auf dem Gitter ξ_n der Summenbedingung genügen (vgl. Kapitel 2). Im folgenden werden wir zunächst die beschriebene Methode für $m = 2$ näher untersuchen. An dieser Stelle sei noch erwähnt, dass die richtige Anzahl der Extremwerte für uns ebenfalls von Bedeutung ist. Hier betrachten wir aber zunächst die Glattheit im obigen Sinne als Anforderung an $\hat{f}$. Das richtige Monotonieverhalten wird später bei der Berechnung hinzu genommen. Zuerst besprechen wir die asymptotischen Eigenschaften dieser Methode. Wegen der einfacheren Notation benutzen wir hier das Summenkriterium. Die Beweise lassen sich auf allgemeinere ReT übertragen. Dazu benötigt man jedoch Zusatzvoraussetzungen an T_i^k[1], die wir hier jedoch nicht genauer erläutern wollen. An dieser Stelle sei wieder erwähnt, dass die Aussagen in dieser Arbeit sich in erster Linie auf den Fall $m = 2$ beschränken.

[1] Zum Beispiel kann man verlangen, dass $\|T_i^k\|_1$ von 0 beschränkt ist.

5.2 Asymptotische Konvergenz

In diesem Kapitel wollen wir einige asymptotische Eigenschaften notieren. Wir bezeichnen hier wie im letzten Kapitel:

$$W_2 = W_2[0,1] := \{g : [0,1] \to \mathbb{R} | g\ 2\times \text{ diff-bar und } \int_0^1 g^{(2)}(t)^2 dt < \infty\},$$

für

$$g \in W_2, \|g^{(2)}\|_2 := \sqrt{\int_0^1 g^{(2)}(t)^2 dt}$$

und $\|g\|_\infty := \sup_{x\in[0,1]} |g(x)|$.

Definition 5.5 Gegeben sei $g : [0,1] \to \mathbb{R}$ und

$$\xi = \xi_n = (0, t_1, \ldots, t_{n-1}, 1),$$

ein Gitter auf $[0,1]$. Wir sagen g *genügt der Summenbedingung mit einer Konstanten* $c > 0$, falls gilt:

$$|\sum_{i\in\xi\cap I} g(t_i)| \leq \sqrt{c\log(n+1)}\sqrt{l_n},$$

mit

$$l_n = l_n(I) := \#\{t_i | 0 \leq i \leq n\} \cap I$$

für beliebige Intervalle $I \subset [0,1]$.
Wir notieren dies mittels

$$g \in \mathrm{S}(\xi, c).$$

Zuerst wird folgendes Lemma über die Summen der Ableitungen bewiesen. Da wir die zweite Ableitung der Funktion minimieren möchten, setzten wir $\hat{f} \in \mathcal{W}_2$ voraus. Damit können wir im folgenden, die Existenz der ersten beiden Ableitungen voraussetzen, ohne zusätzliche Eischränkungen des Modells vornehmen zu müssen.

Lemma 5.6 Es sei

$$\xi = (t_1, \ldots, t_n)^T$$

ein reguläres Gitter und mit $c > 0$, sei

$$g \in W_2[0,1] \cap \mathrm{S}(\xi, c),$$

gegeben. Weiterhin seien

$$j, l \in \{1, \ldots, n\}, k \in \mathbb{Z}$$

mit $1 \leq j + 1 + k, j + k + l \leq n$.

i) Es gilt:

$$|\sum_{i=j+1}^{j+l} g^{(1)}(t_i)(t_{i+k} - t_i)| \leq 2\sqrt{c \log(n+1)}\sqrt{l} + \frac{\|g^{(2)}\|_2}{\sqrt{3}} \sum_{i=j+1}^{j+l} |t_{i+k} - t_i|^{\frac{3}{2}}.$$

ii) Insbesondere ist ξ äquidistant, so gilt:

$$|\sum_{i=j+1}^{j+l} g^{(1)}(t_i)||\frac{k}{n}| \leq 2\sqrt{cl \log(n+1)} + \frac{\|g^{(2)}\|_2}{\sqrt{3}} l |\frac{k}{n}|^{\frac{3}{2}}.$$

Mit Hilfe dieses Lemmas können wir die Bedingungen an die Funktionswerte $g(t_i)$ auf die Ableitung $g^{(1)}$ übertragen.

Beweis:

Wir benutzen die Taylorsche Formel mit Integraldarstelung des Restglieds, um die Funktion g an der Stelle t_{i+k} dar zu stellen:

$$\begin{aligned} g(t_{i+k}) &= g(t_i) + g^{(1)}(t_i)(t_{i+k} - t_i) \\ &\quad + \int_{t_i}^{t_{i+k}} g^{(2)}(x)(t_{i+k} - x)dx, (i = j+1, \ldots, j+l). \end{aligned}$$

Summieren wir über $i = j+1, \ldots, j+l$, so ergibt sich:

$$\begin{aligned}
& |\sum_{i=j+1}^{j+l} g^{(1)}(t_i)(t_{i+k}-t_i)| \\
= \;& |\sum_{i=j+1}^{j+l} g(t_{i+k}) - g(t_i) - \sum_{i=j+1}^{j+l} \int_{t_i}^{t_{i+k}} g^{(2)}(x)(t_{i+k}-x)dx| \\
\leq \;& |\sum_{i=j+1}^{j+l} g(t_{i+k})| + |\sum_{i=j+1}^{j+l} g(t_i)| + |\sum_{i=j+1}^{j+l} \int_{t_i}^{t_{i+k}} g^{(2)}(x)(t_{i+k}-x)dx| \\
\leq \;& 2\sqrt{cl\log(n+1)} + \sum_{i=j+1}^{j+l} |\int_{t_i}^{t_{i+k}} g^{(2)}(x)(t_{i+k}-x)dx|, (g \in \mathrm{S}(\xi, c)) \\
\leq \;& 2\sqrt{cl\log(n+1)} + \sum_{i=j+1}^{j+l} \sqrt{|\int_{t_i}^{t_{i+k}} g^{(2)}(x)^2 dx|}\sqrt{|\int_{t_i}^{t_{i+k}} (t_{i+k}-x)^2 dx|} \\
\leq \;& 2\sqrt{cl\log(n+1)} + \|g^{(2)}\|_2 \sum_{i=j+1}^{j+l} \frac{1}{\sqrt{3}} |t_{i+k}-t_i|^{\frac{3}{2}}.
\end{aligned}$$

Die vorletzte Ungleichung folgt aus der Cauchy-Schwarz'schen Ungleichung. Setzen wir nun $t_{i+k} - t_i = \frac{k}{n}$ ein, so folgt die zweite Behauptung des Lemmas.

■

Satz 5.7 Es sei $f \in W_2$ fest, $y(t) = f(t) + \varepsilon(t)$, wobei $\varepsilon(t)$ eine $\mathcal{N}(0,1)$ verteilte Zufallsvariable für beliebige $t \in [0,1]$ sein soll. Ferner sei $(\xi_n)_{n=1}^{\infty}$ eine Folge von Gittern und $f_n \in W_2$ die Lösung des Minimierungsproblems

$$\|g^{(2)}\|_2^2 = \min \text{ s.t. } g - y \in \mathrm{S}(\xi_n, \tau_1), (\tau_1 > 4).$$

i) Für $\delta > 0$ gilt:

$$\Pr(n^{1/8}\log(n+1)^{-\alpha}\|f^{(1)} - f_n^{(1)}\|_\infty \leq \delta) \to 1, (n \to \infty).$$

ii) Für $\delta > 0$ gilt:

$$\Pr(n^{3/8}\log(n+1)^{-\beta}\|f - f_n\|_\infty \leq \delta) \to 1, (n \to \infty).$$

Dabei sind α, β beliebige Konstanten > 0.5.

Beweis:

Beweis von i):

Wir nehmen zunächst an, dass $f - y \in \mathrm{S}(\xi_n, \tau_1)$ gilt. Damit folgt:

$$\|f^{(2)} - f_n^{(2)}\|_2 \leq \|f^{(2)}\|_2 + \|f_n^{(2)}\|_2 \leq 2\|f\|_2 =: K.$$

Wegen $f_n - y \in \mathrm{S}(\xi_n, \tau_1)$ gilt:

$$h_n := f - f_n = f - y - (f_n - y) \in \mathrm{S}(\xi_n, 2\tau_1).$$

Wir setzen

$$\tau := 2\tau_1, \mu_n := n^{-1/4}, I_n(t) := [t - \mu_n, t + \mu_n] \cap [0,1], j_n(t), l_n(t) \in \mathbb{N},$$

mit

$$\{t_i | 1 \leq i \leq n\} \cap I_n(t) = \{(j_n(t)+1)/n, \ldots, (j_n(t) + l_n(t))/n\}.$$

Im folgenden werden wir von I_n, j_n, l_n sprechen. Wählen wir nun n so groß, dass $[t+\mu_n, t+5\mu_n] \subset [0,1]$ oder $[t - 5\mu_n, t - \mu_n] \subset [0,1]$ gilt und setzen $k_n := l_n$, falls $(j_n + 2l_n \leq n)$ und $k_n := -l_n$ sonst, gilt: $1 \leq j_n + 1 + k_n, j_n + l_n + k_n$. Betrachten wir:

$$h_n^{(1)}(t) = h_n^{(1)}(t_i) + \int_{t_i}^{t} h_n^{(2)}(x)dx, (i = j_n + 1, \ldots, j_n + l_n).$$

Nach Summation über $i = j_n + 1 \ldots j_n + l_n$ ergibt sich:

$$\begin{aligned}
l_n|h_n^{(1)}(t)| &= |\sum_{i=j_n+1}^{j_n+l_n} h_n^{(1)}(t_i) + \sum_{i=j_n+1}^{j_n+l_n} \int_{t_i}^{t} h_n^{(2)}(x)dx| \\
&\leq |\sum_{i=j_n+1}^{j_n+l_n} h_n^{(1)}(t_i)| + \sum_{i=j_n+1}^{j_n+l_n} |\int_{t_i}^{t} h_n^{(2)}(x)dx| \\
&\leq |\sum_{i=j_n+1}^{j_n+l_n} h_n^{(1)}(t_i)| + l_n \sqrt{|\int_{t_i}^{t} dx|} \sqrt{\int_{t_i}^{t} |h_n^{(2)}(x)^2 dx|} \\
&\leq 2|\frac{n}{k_n}| \sqrt{l_n \log(n+1)\tau} + \frac{K}{\sqrt{3}} l_n \sqrt{|\frac{k_n}{n}|} + l_n K \sqrt{|t - t_i|}, \text{(Lemma 5.6)} \\
&\leq 2|\frac{n}{k_n}| \sqrt{l_n \log(n+1)\tau} + \frac{K}{\sqrt{3}} l_n \sqrt{|\frac{k_n}{n}|} + l_n K \sqrt{|I_n|}.
\end{aligned}$$

Setzen wir $|k_n| = l_n$ und dividieren durch l_n, so folgt:

$$|h_n^{(1)}(t)| \leq 2\sqrt{\frac{\tau_1 n^2 \log(n+1)}{l_n^3}} + \frac{K}{\sqrt{3}}\sqrt{\frac{l_n}{n}} + K\sqrt{|I_n|}. \tag{5.2}$$

Nach der Definition von μ_n gilt:

$$n\mu_n \leq l_n \leq 2\mu_n n, \mu_n \leq |I_n| \leq 2\mu_n.$$

mit $n\mu_n = n^{3/4}$ folgt:

$$|h_n^{(1)}(t)| \leq 2\sqrt{\frac{\tau_1 n^2 \log(n+1)}{n^{9/4}}} + \frac{K}{\sqrt{3}}\sqrt{\frac{2n^{3/4}}{n}} + K\sqrt{2n^{-1/4}}.$$

Da die rechte Seite unabhängig von t ist können wir folgern:

$$\begin{aligned}
& n^{1/8}\log(n+1)^{-\alpha}\|h_n^{(1)}\|_\infty \\
\leq\ & n^{1/8}\log(n+1)^{-\alpha}\left(\sqrt{\frac{4\tau_1\log(n+1)n^2}{n^{\frac{9}{4}}}} + K\sqrt{2}\left(\frac{1}{\sqrt{3}}\sqrt{n^{-1/4}} + \sqrt{n^{-1/4}}\right)\right) \\
=\ & \sqrt{4\tau_1}\log(n+1)^{-(\alpha-1/2)}n^{(1+1/8-9/8)} + K\sqrt{2}\frac{1+\sqrt{3}}{\sqrt{3}}\log(n+1)^{-\alpha}n^{(1/8-1/8)} \\
=\ & \sqrt{4\tau_1}\log(n+1)^{-(\alpha-1/2)} + \left(K\sqrt{2}\frac{1+\sqrt{3}}{\sqrt{3}}\right)\log(n+1)^{-\alpha} \to 0, (n\to\infty).
\end{aligned}$$

Damit können wir für $\delta > 0$ fix, $N_\delta \in \mathbb{N}$ so wählen, dass die rechte Seite kleiner als δ wird. Dann folgt für alle $n \geq N_\delta$:

$$\Pr(\{f \in \mathrm{S}(\xi_n, \tau_1)\}) \leq \Pr(n^{1/8}\log(n+1)^{-\alpha}\|f - f_n\|_\infty \leq \delta)$$

und damit

$$\Pr(n^{1/8}\log(n+1)^{-1}\|f - f_n\|_\infty \leq \delta) \to 1, (n \to \infty).$$

Beweis von $ii)$:

Wir wählen zunächst zu $t \in [0,1]$ beliebig, I_n, j_n, l_n genauso wie oben, dann folgt:

$$h_n(t) = h_n(t_i) + h_n^{(1)}(t_i)(t - t_i) + \int_{t_i}^{t} h_n^{(2)}(x)(t-x)dx, (i = j_n + 1, \ldots, j_n + l_n),$$

und Summieren über i ergibt:

$$\begin{aligned} l_n|h(t)| &\leq |\sum_{i=j_n+1}^{j_n+l_l} h(t_i)| + \|h_n^{(1)}\|_\infty \sum_{i=j_n+1}^{j_n+l_n} |t-t_i| + \frac{K}{\sqrt{3}} l_n |I_n|^{3/2} \\ &\leq \sqrt{\tau_1 \log(n+1) l_n} + \|h_n^{(1)}\|_\infty l_n |I_n| + \frac{K}{\sqrt{3}} l_n |I_n|^{3/2}. \end{aligned}$$

Es sei $\beta > 0.5$ beliebig gegeben. Wählen wir $0.5 < \alpha_1 < \beta$ und f_n, so dass $\|h_n^{(1)}\|_\infty \leq n^{-1/8} \log(n+1)^{\alpha_1}$ gilt. Investieren wir

$$n\mu_n \leq l_n \leq 2\mu_n n, \mu_n \leq |I_n| \leq 2\mu_n,$$

so ergibt sich nach der Division durch l_n:

$$\begin{aligned} & |h_n(t)| \\ \leq\ & 2\sqrt{\frac{\tau_1 \log(n+1)}{l_n}} + \log(n+1)^\alpha n^{-1/8} |I_n| + \frac{K}{\sqrt{3}} |I_n|^{3/2} \\ \leq\ & 2\sqrt{\frac{\tau_1 \log(n+1)}{n^{3/4}}} + 2\log(n+1)^\alpha n^{(-1/8-1/4)} + K\sqrt{\frac{8}{3}} n^{-3/8} \end{aligned} \tag{5.3}$$

Damit können wir, wegen $\beta > \alpha_1$, folgern:

$$\begin{aligned} & n^{3/8} \log(n+1)^{-\beta} \|h_n\|_\infty \\ \leq\ & n^{3/8} \log(n+1)^{-\beta} \left(2\sqrt{\frac{\tau \log(n+1)}{n^{3/4}}} + 2\log(n+1)^{\alpha_1} n^{(-3/8)} + K\sqrt{\frac{8}{3}} n^{-3/8} \right) \\ =\ & 2\sqrt{\tau_1} \log(n+1)^{-(\beta-1/2)} + 2\log(n+1)^{-(\beta-\alpha_1)} + K\sqrt{\frac{8}{3}} \log(n+1)^{-\beta} \\ \to\ & 0, (n \to \infty). \end{aligned}$$

Wählen wir zu $\delta > 0$ N_δ, so dass die rechte Seite $\leq \delta$ wird. Dann folgt für alle $n \geq N_\delta$:

$$\begin{aligned} & \Pr\left(\{f \in \mathrm{S}(\xi_n, \tau_1)\} \cap \{\|f^{(1)} - f_n^{(1)}\|_\infty \leq n^{-1/8} \log(n+1)^{\alpha_1}\}\right) \\ \leq\ & \Pr(n^{3/8} \log(n+1)^{-\beta} \|f - f_n\|_\infty \leq \delta). \end{aligned}$$

■

In den oben aufgeführten Beweisen kann man erkennen, dass für den Konvergenzbeweis, die Begingungen nicht auf allen Intervallen verlangt werden müssen. Es reicht z.B. alle duadischen Intervalle der Länge $\leq 2/n^{-1/4}$ zu betrachten. Wir wollen jedoch auf diese Feinheiten nicht eigehen, sondern versuchen die Konvergenzgeschwindigkeit zu verbessern. Mit einer weiteren Voraussetzung an f kann man eine schnellere Konvergenz erreichen.

Lemma 5.8 Es sei $f : [0,1] \to \mathbb{R}$ $3\times$ mal differenzierbar und $|\int_0^1 f^{(3)}(t)dt| < \infty$. Ferner seien $f_n \in W_2$ wie in Satz 5.7 gewählt. Für $\delta > 0, h_n := f - f_n$ beliebig gelten dann folgende Aussagen:

i) Gilt mit $\alpha > 0.5$, $\Pr(n^\nu \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta) \to 1, (n \to \infty)$, so folgt

$$\Pr(n^{\nu/2} \log(n+1)^{-\beta} \|h_n^{(2)}\|_2 \leq \delta) \to 1, (n \to \infty),$$

für $\beta > \alpha/2$.

ii) Gilt mit $\eta \in (0,1), \beta > 0.5 : \Pr(n^{\eta/2} \log(n+1)^{-\beta} \|f_n^{(2)}\|_2 \leq \delta) \to 1, (n \to \infty)$, so folgt: $\Pr(n^{(1+3\eta)/8} \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta) \to 1, (n \to \infty)$, für $\alpha > \beta$.

Beweis:

Beweis i):

Setzen wir voraus, dass $f \in \mathrm{S}(\xi_n, \tau_1)$ ist, so folgt:

$$\begin{aligned}
\|f_n^{(2)} - f^{(2)}\|_2^2 &\leq \|f_n^{(2)} - f^{(2)}\|_2^2 + \|f^{(2)}\|_2^2 - \|f_n^{(2)}\|_2^2 \\
&= \|f^{(2)}\|_2^2 - 2\int_0^1 f^{(2)}(t) f_n^{(2)}(t)dt + \|f_n\|_2^2 + \|f^{(2)}\|_2^2 - \|f_n^{(2)}\|_2^2 \\
&= 2\int_0^1 (f^{(2)}(t) - f_n^{(2)}(t)) f^{(2)}(t)dt \\
&= 2[(f^{(1)}(x) - f_n^{(1)}(x)) f^{(2)}(x)]_0^1 - 2\int_0^1 h_n^{(1)}(t) f^{(3)}(t)dt \\
&\leq \|h_n^{(1)}\|_\infty |f^{(2)}(1) - f^{(2)}(0)| + \|h_n^{(1)}\|_\infty |\int_0^1 f^{(3)}(t)dt|
\end{aligned}$$

Nehmen wir für festes $\delta > 0$ zusätzlich an, dass

$$n^\nu \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta$$

gilt, so folgt mit $c := |f^{(2)}(1) - f^{(2)}(0)| + |\int_0^1 f^{(3)}(t)dt| < \infty$:

$$\begin{aligned} n^{\nu/2} \log(n+1)^{-\beta} \|h_n^{(2)}\|_2 &\leq n^{\nu/2} \log(n+1)^{-\beta} \sqrt{\|h_n^{(2)}\|_\infty c\delta} \\ &\leq n^{\nu/2} \log(n+1)^{-\beta} \sqrt{n^{-\nu/2} \log(n+1)^\alpha c\delta} \\ &= \log(n+1)^{-(\beta-\alpha/2)} \sqrt{c\delta} \to 0 (n \to \infty). \end{aligned}$$

Nun wählen wir $N_\delta \in \mathbb{N}$, so dass die rechte Seite kleiner als δ wird. Damit gilt für $n \geq N_\delta$:

$$\begin{aligned} & \Pr\left(\{f - y \in \mathrm{S}(\xi_n, \tau_1)\} \cap \{n^\nu \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta\right) \\ \leq \; & \Pr\left(n^{\nu/2} \log(n+1)^{-\beta} \|h_n^{(2)}\|_2 \leq \delta\right). \end{aligned}$$

Mit $n \to \infty$ ist die Behauptung bewiesen.

Beweis von ii)

Wähle $\mu_n := n^{(\eta-1)/4}$ und betrachte damit die Ungleichung (5.2):

$$|h_n^{(1)}(t)| \leq 2\sqrt{\frac{\tau n^2 \log(n+1)}{l_n^3}} + K_n \left(\sqrt{\frac{l_n}{3n}} + \sqrt{|I_n|} \right)$$

mit $K_n = 2\|h_n^{(2)}\|_2$. Wählen wir $K_n \leq \delta n^{-\eta/2} \log(n+1)^\beta$, so gilt mit

$$n\mu_n \leq l_n \leq n\mu_n, \mu_n \leq |I_n| \leq 2\mu_n$$

und $n\mu_n = n^{(3+\eta)/4}$:

$$\begin{aligned} \|h_n^{(1)}\|_\infty &\leq 2\sqrt{\frac{\tau_1 n^2 \log(n+1)}{n^{(9+3\eta)/4}}} + \delta n^{-\eta/2} \log(n+1)^\beta \left(\sqrt{\frac{2n\mu_n}{3n}} + \sqrt{2\mu_n} \right) \\ &= 2\sqrt{\tau} n^{-(1+3\eta)/8} \log(n+1)^{1/2} + \delta \sqrt{\frac{8}{3}} \log(n+1)^\beta n^{(\eta-1)/8-\eta/2} \end{aligned}$$

Demnach gilt für $\alpha > \beta$:

$$\begin{aligned}
& n^{(1+3\eta)} \log(n+1)^{-\alpha} \|h_n^{(2)}\|_\infty \\
\leq \; & n^{(1+3\eta)} \log(n+1)^{-\alpha} \\
& \left(2\sqrt{\tau_1} n^{-(1+3\eta)/8} \log(n+1)^{1/2} + \delta \sqrt{\frac{8}{3}} \log(n+1)^{\beta} n^{-(1+3\eta)/4} \right) \\
= \; & 2\sqrt{\tau} \log(n+1)^{1/2-\alpha} + \delta \sqrt{\frac{8}{3}} \log(n+1)^{\beta-\alpha} n^{-(1+3\eta)/4} \\
& \to 0, (n \to \infty), (\alpha > \beta > 1/2).
\end{aligned}$$

Damit folgt die Behauptung wie oben.

■

Satz 5.9 Es sei wieder f $3\times$ diff-bar und $|\int_0^1 f^{(3)}(t)| < \infty|$. Ferner sei h_n wie im Lemma 5.8 gegeben. Wir setzen

$$b_0 = \frac{1}{8}, \; b_{N+1} := \frac{1+3b_N}{8}$$

Es gelten folgende Aussagen:

1) Für alle $N \in \mathbb{N}, \alpha > 0.5, \delta > 0$ gilt:

$$\Pr\left(n^{b_N} \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta\right) \to 1, (n \to \infty).$$

2) Für alle $N \in \mathbb{N}$ gilt:

$$b_N \leq b_{N+1} \leq 1/5 \text{ und } \lim_{n\to\infty} b_N = \frac{1}{5}.$$

3) Für beliebige $0 < \nu \leq 1/5, \alpha > 0.5, \delta > 0$ gilt:

i) $\Pr(n^{\nu/2} \log(n+1)^{-\alpha} \|h_n^{(2)}\|_2 \leq \delta) \to 1, (n \to \infty)$

ii) $\Pr(n^{\nu} \log(n+1)^{-\alpha} \|h_n^{(1)}\|_\infty \leq \delta) \to 1, (n \to \infty)$

iii) $\Pr(n^{(1+3\nu)/4} \log(n+1)^{-\alpha} \|h_n\|_\infty \leq \delta) \to 1, (n \to \infty)$

Beweis:

1) Für $N = 0$ ist die Aussage im Satz 5.7 bereits bewiesen worden. Für den Fall $N > 1$ brauchen wir nur das Lemma 5.8 mit Induktion anzuwenden.

2) Wir müssen bemerken, dass

$$\frac{1+3b_N}{8} \leq \frac{1}{5} \Leftrightarrow b_N \leq \frac{1}{5},$$
$$b_N \leq \frac{1+3b_N}{8} \Leftrightarrow b_N \leq \frac{1}{5}.$$

3) Für $\nu < \frac{1}{5}$ folgt 3) aus 2). Für $\nu = 1/5$ müssen wir die Folge

$$\Pr\left(n^{b_n} \log(n+1)^{\alpha} \|h_n^{(1)}\|_{\infty} \leq \delta\right)$$

betrachten.

■

Folgerung 5.1
Insbesondere folgt aus dem letzten Satz in diesem Fall:

$$\Pr\left(n^{2/5} \log(n+1)^{\alpha} \|f - f_n\|_{\infty} \leq \delta\right) \to 1, (n \to \infty).$$

5.3 Berechnung

Hier soll $\hat{f}$, eine Approximation von f berechnet werden, mit

1) $y - \hat{f}$ sieht aus wie Rauschen,

2) $\hat{f}$ hat keine Artefakte in Form von zusätzlichen Extremwerten,

3) $\hat{f} \in W_2$ und möglichst glatt.

Betrachten wir das Problem mit einer Diskretisierung auf einem Gitter wie im zweiten Kapitel, so können wir das Problem formulieren als

$$\min \frac{1}{2} g^T Q_2 g \text{ s.t. } y - g \in \mathrm{MR}(\xi_n, c_n), \mathcal{M}g \geq 0. \tag{5.4}$$

Dabei bezeichnet MR das Multiresolutionskriterium, was wir zur Effizienz der Methode anstelle des Summenkriteriums benutzen. Führen wir uns wieder die Argumente in Kapitel 4 vor Augen, so wird deutlich, dass ein allgemeines exaktes QP hierfür sehr teuer sein würde. Die Methoden, mit denen wir die Speicherkapazität reduziert haben verlieren hier ihre Wirkung, da die Multiresolutionsbedingungen keine “lokalen” Nebenbedingungen wie die Grenzbedingungen sind. Im folgenden wird das Prinzip der Straffunktionen vorgestellt. Für eine genauere Beschreibung wird auf die Literatur verwiesen (vgl. [61] oder [53]). Danach betrachten wir die Ergebnisse.

5.3.1 Strafterm-Algorithmen und LANCELOT

Betrachten wir das Optimierungsproblem, so wird deutlich, dass hier im Vergleich zum Quadratschen Programm im letzten Kapitel, eine große Anzahl von Nebenbedingungen hinzu kommen. Außerdem geht hier die “lokale” Struktur der Nebenbedingungen verloren. Es gibt Multiresolutionskoeffizienten die sich auf große Intervalle beziehen. Hier wollen wir auf die sogenannten Strafterm-Algorithmen zurückgreifen, die sich in ähnlichen Zusammenhängen als sehr effizient erweisen. Die darauf basierende Methode, LANCELOT (vgl. [53]), wird in vielen kommerziellen Software-Paketen benutzt. Wir geben hier lediglich eine grobe Beschreibung und verweisen auf die Literatur. Gegeben sei das Problem:

$$\begin{aligned} &q(x) = \min \ \text{s.t.} \\ &Ax \geq b \\ &l \leq x \leq u. \end{aligned}$$

Wir schreiben das Problem zunächst um in:

$$
\begin{aligned}
&q(x) = \min \ \text{s.t.}\\
&Ax - s = b,\\
&s \geq 0,\\
&l \leq x \leq u,
\end{aligned}
$$

wobei $s = (s_1, \ldots, s_{n_u})^T$. Man behandelt hier die Nebenbedingungen $l \leq x \leq u$ gesondert, weil es effizientere Algorithmen gibt mit denen man solche Optimierungsprobleme lösen kann. Wir haben im letzten Kapitel Methoden gelernt, die uns erlauben die Monotoniebedingungen ebenfalls effizient zu behandeln. Damit können wir diese Nebenbedingungen ebenfalls gesondert betrachten.
Wir notieren also:

$$
\begin{aligned}
&q(x) = \min \ \text{s.t.}\\
&Ax - s = b,\\
&s \geq 0,\\
&l \leq \mathcal{I}^k x \leq u,\\
&\mathcal{M}\mathcal{D}^{k+1} x \geq 0,
\end{aligned}
$$

wobei $k \in \{0, 1\}$, und $\mathcal{I}, \mathcal{D}$, jeweils das diskrete Integrieren bzw. Differenzieren und $\mathcal{M}$ das Monotonieverhalten beschreiben sollte (vgl. Kapitel 2). Die Zusatzparameter s_i haben einen stabilisierenden Einfluß auf den Algorithmus (siehe [53] für eine genauere Diskussion). In dem Algorithmus werden eine Folge von Subproblemen gelöst, die mit Hilfe dieser Funktion definiert werden.
Wir betrachten die Funktion

$$\mathcal{L}_\rho(\lambda, x, s_k) := q(x) - \sum_{i=1}^{n_u} \lambda_i (a_i^T x - b_i - s_i) + \frac{1}{2\rho} \sum_{i=1}^{n_u} (a_i^T x - b_i - s_i)^2,$$

mit $(\rho > 0)$. Der LANCELOT-Algorithmus definiert eine Folge $\rho_k \to 0$, so dass die Lösung des Problems

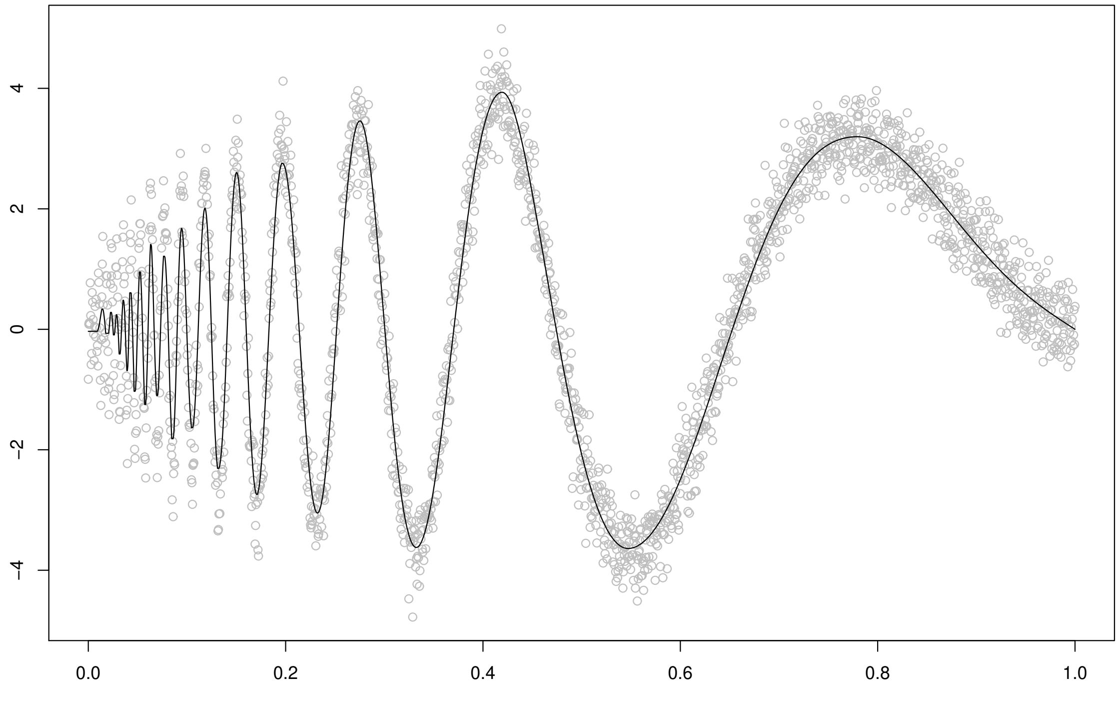


Abbildung 5.2: Das Ergebnis von LANCELOT, angewendet auf den Doppler-Datensatz, unter Einhaltung der Monotonie- und Multiresolutionsbedingungen.

$$\begin{aligned} &\mathcal{L}_{\rho_k}(\lambda^k, x, s) = \min \text{ s.t.} \\ &l \leq \mathcal{I}^k x \leq u, \\ &\mathcal{MD}^{k+1} x \geq 0. \end{aligned} \tag{5.5}$$

gegen die globale Lösung des ursprünlichen Problems und die λ_i^k gegen die entsprechenden Lagrange-Multiplikatoren, konvergieren.

Algorithmus 5.10 (LANCELOT)

- Starte mit $x^0, s^0, \lambda^0, \rho_0$ fest!
  - **While**($\sum_i \lambda_i^- >$ tol)
    - Setze (x^{k+1}, s^{k+1}) als Minimum des Problems (5.5).
    - $\lambda^{k+1} := \max(\lambda^k - (Ax^k - b)/\rho_k, 0)$.
  - $\rho_{k+1} \leftarrow (1+\epsilon)\rho_k$
  - $k+1 \leftarrow k$
  - end **While**

In jedem Schritt wird ρ_k vergrößert und damit die verletzten Bedingungen mit einer größeren Konstante bestraft. Die Berechnung der Lösung des Minimierungsproblems (5.5) können wir, mit Hilfe der Algorithmen im letzten Kapitel, effizient durchführen. Wir wollen hier auf Details verzichten, verweisen auf [53] und widmen uns der Anwendung dieser Methode auf unser Problem.

Satz 5.11 Es sei das Q wie im letzten Kapitel gegeben und damit das Problem:

$$\begin{aligned} &\tfrac{1}{2} x^T Q x = \min \text{ s.t.} \\ &l \leq \mathcal{I}^k x \leq u, \\ &\mathcal{MD}^{k+1} x \geq 0 \end{aligned} \tag{5.6}$$

und x erfüllt die Multiresolutionsbedingungen bis zur Ordnung m, wobei m ein natürliche Zahl $\leq n$ ist. Dann lässt sich LANCELOT mit $O(n)$ Speicherverbrauch programmieren.

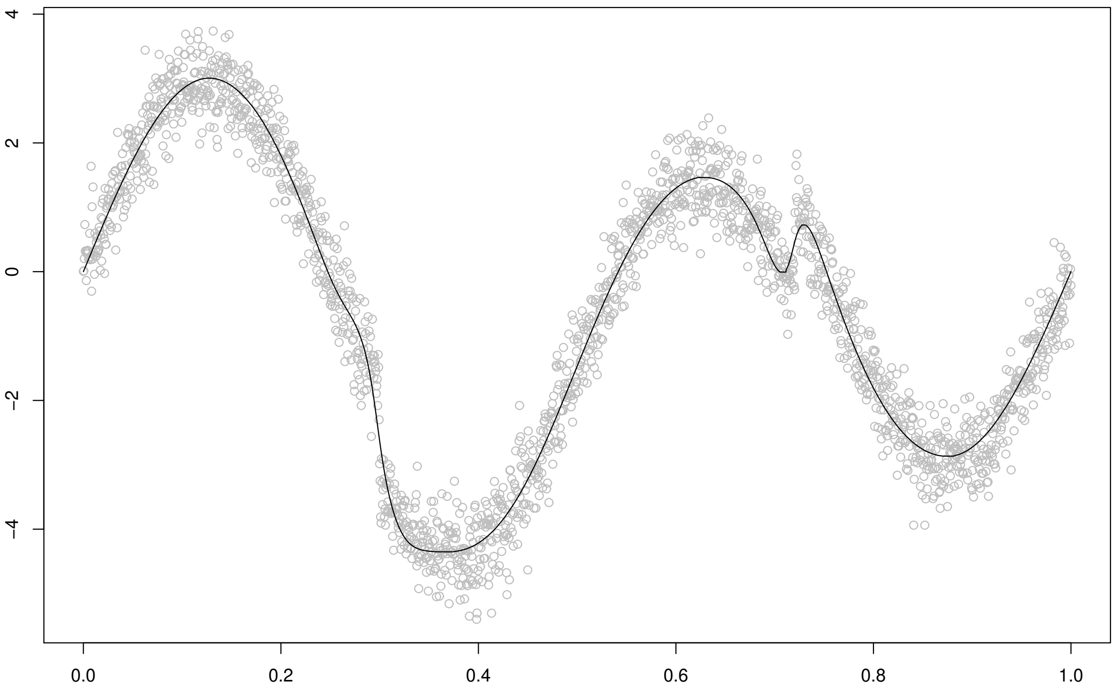
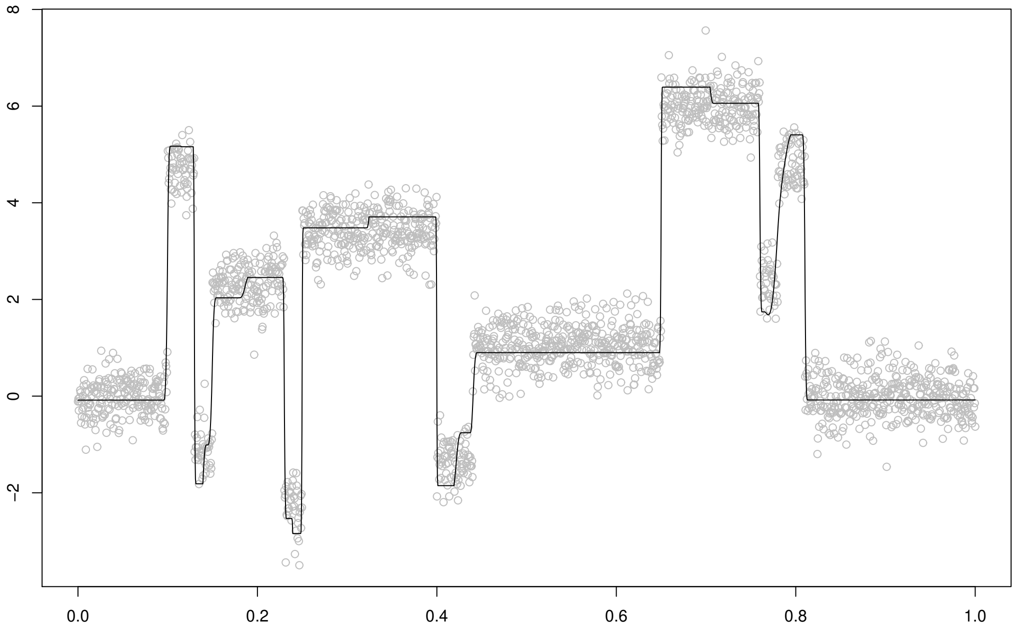


Abbildung 5.3: Das Ergebnis von LANCELOT, angewendet auf das Problem (5.6). Das Ergebnis von LANCELOT angewendet auf den Heavisine-Datensatz, unter Einhaltung der Monotonie- und Multiresolutionsbedingungen.



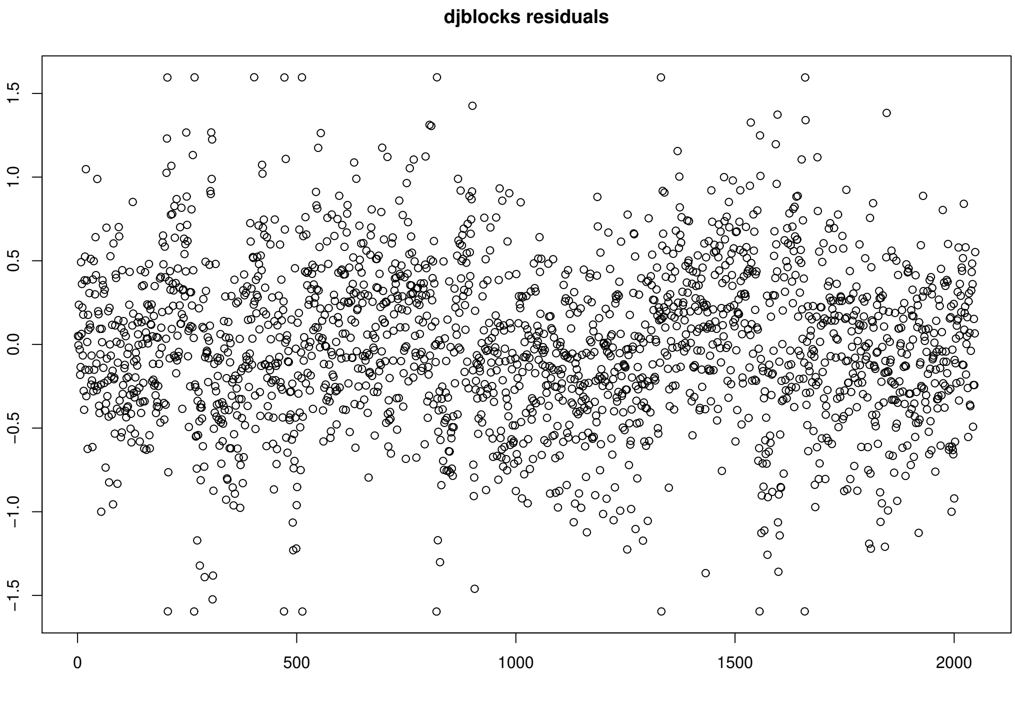


Abbildung 5.4: Das Ergebnis von LANCELOT, angewendet auf den Blocks-Datensatz, unter Einhaltung der Monotonie- und Multiresolutionsbedingungen. Im unteren Bild sind die Residuen geplottet, die nun der Multiresolutionsbedingung genügen.

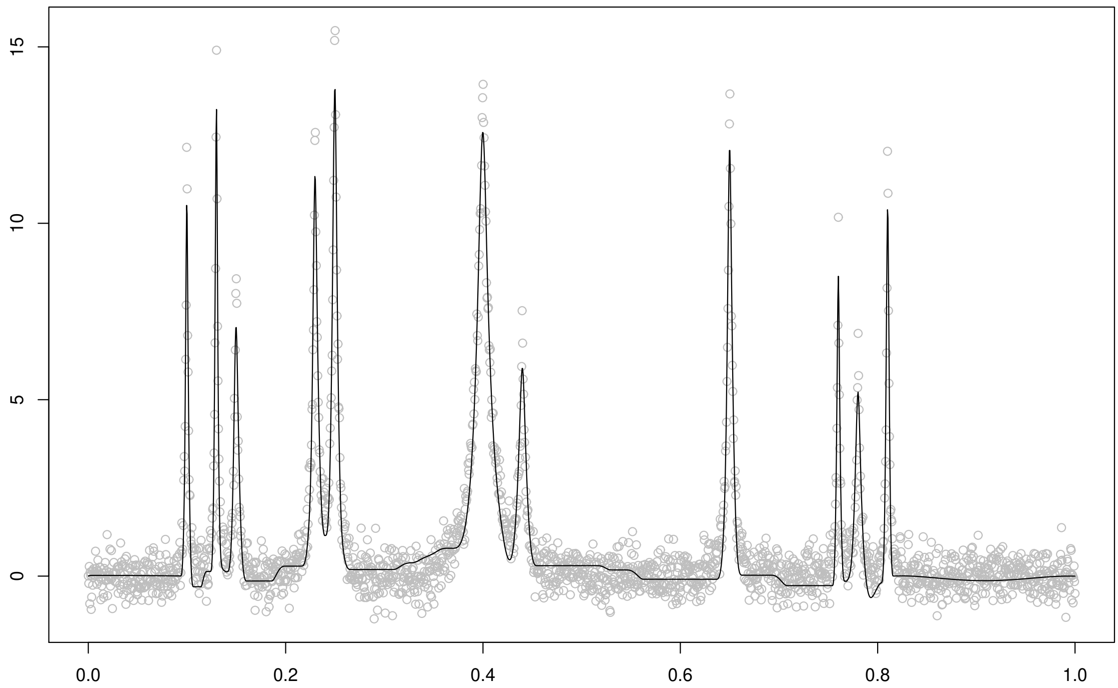


Abbildung 5.5: Das Ergebnis von LANCELOT, angewendet auf den Bumps-Datensatz, unter Einhaltung der Monotonie und Multiresolutionsbedingungen.

Beweis:

Wir müssen lediglich bemerken, dass im Fall $k = 1$, $a_i^T x = x_{j_i+2^\kappa} - x_{j_i}, (2^\kappa \leq m)$ und im Fall $k = 0$,

$$a_i^T x = \sum_{j_i+1}^{j_i+2^\kappa} x_j, (2^\kappa \leq m)$$

für manche j_i bedeuten. Damit führt das Minimieren der Funktion $\mathcal{L}_\rho$ auf eine Diagonalmatrix.

■

In der unteren Tabelle sehen wie die Mediane der Rechenzeiten in Sec. bei 101 Durchläufen[2] der LANCELOT-Methode, angewendet auf die DJ-Datensätze:

[2]Die Durchläufe wurden alle auf einem Rechner mit einem *AMD Athlon 1700XP* Prozessor unter Einhaltung der Multiresolutionsbedingungen auf allen duadischen Intervallen $\leq 1/16$ durchgeführt.

LANCELOT	Sec.
Doppler	4.5
Heavisine	6.3
Bumps	11.1
Blocks	14.2

5.4 Schätzen von Ableitungen

An dieser Stelle soll ein anderes Problem in der Nichtparametrischen Regression vorgestellt werden. Gegeben sei ein Datensatz $y = \{y_1, \ldots, y_n\}$, welches das Modell (2.1) erfüllt. Ist einer Approximation der Ableitungsfunktion f' von f erwünscht, so wird automatisch die Anforderung an der "Qualität" der Regressionsfunktion erhöht. Eine Möglichkeit, diese hohe "Qualität" zu erzielen ist die weitere Einschränkung der zulässigen Funktionen für die Regressionsfunktion. Sind wir z.B. daran interessiert die erste Ableitung von f mittels einer Straffunktion-Methode zu approximieren, so können wir z.B. als Strafterm

$$S(g) := \int_0^1 (g^{(4)}(t))^2 dt$$

wählen (vgl. [55]). Diese Vorgehensweise erhöht die Rechenzeit, die man für die Berechnung der Schätzfunktion benötigt. Außerdem bleiben die Probleme mit ungewollten Artefakten in Form von Extremwerten (vgl.Kapitel 2). Selbst kleine Artefakte, die an einem Plot der Regressionsfunktion kaum sichtbar sind, können sich auf die Ableitung viel stärker auswirken (vgl. Abbildung 5.6).

Wir haben in dem vorherigen Abschnitt gesehen, dass mit hoher Wahrscheinlichkeit, die erste und sogar die zweite Ableitung unserer Regressionsfunktion, die erste bzw. die zweite Ableitung der Funktion f sehr gut approximieren, ohne dass man zusätzliche Bedingungen stellen muss. Die Abbildung 5.7 demonstriert diese Eigenschaft. Die formerhaltenden Bedingungen erzwingen eine gute Anpassung an die Daten. Die Güte dieser Anpassung wird erst bei der Betrachtung der Ableitungen deutlich.

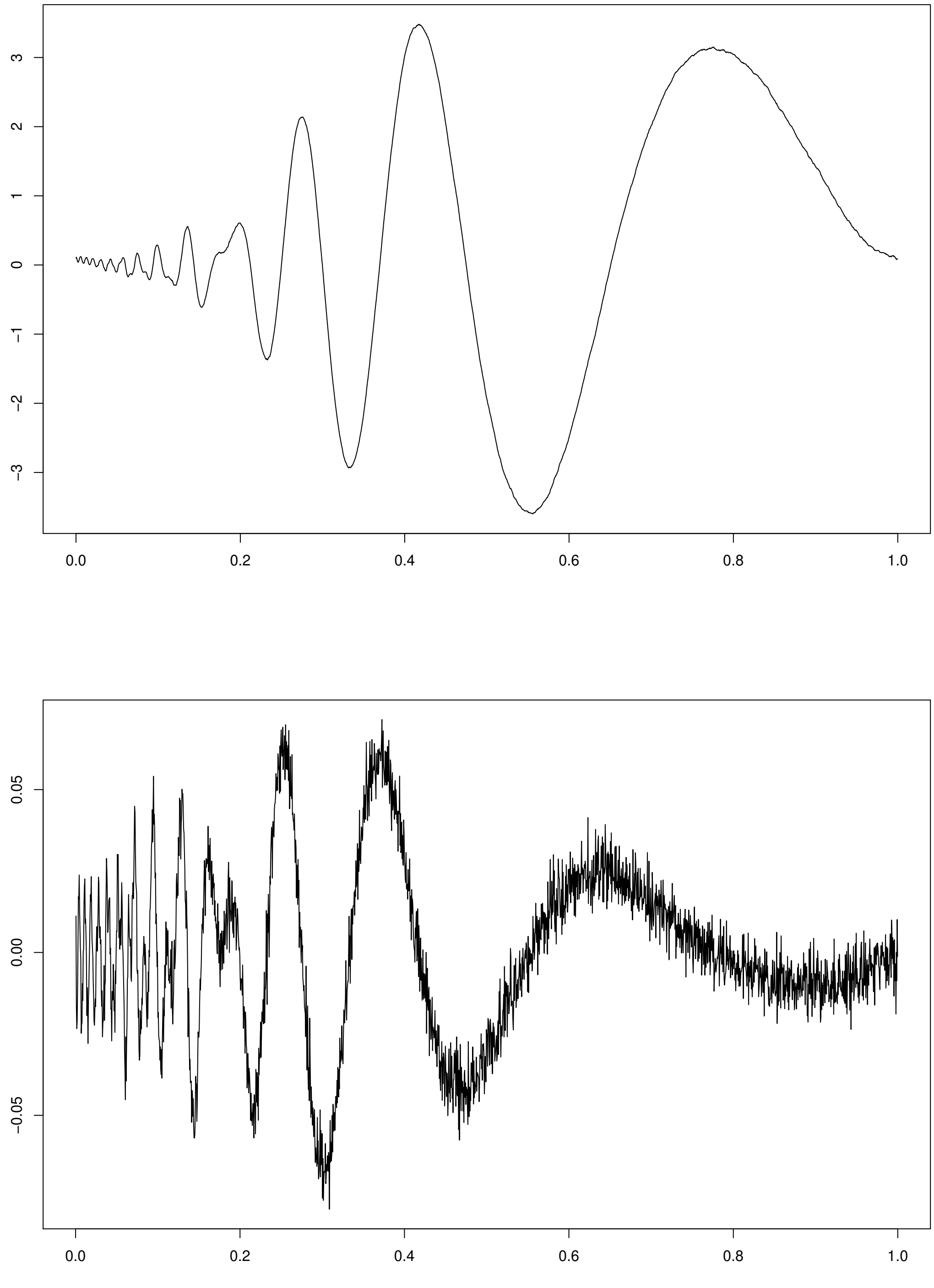


Abbildung 5.6: Das Ergebnis eines Kernschätzers im oberen Bild, scheint glatt zu sein. Die Approximation der Ableitung mittels der ersten Differenzen ist jedoch offensichtlich sehr schlecht.

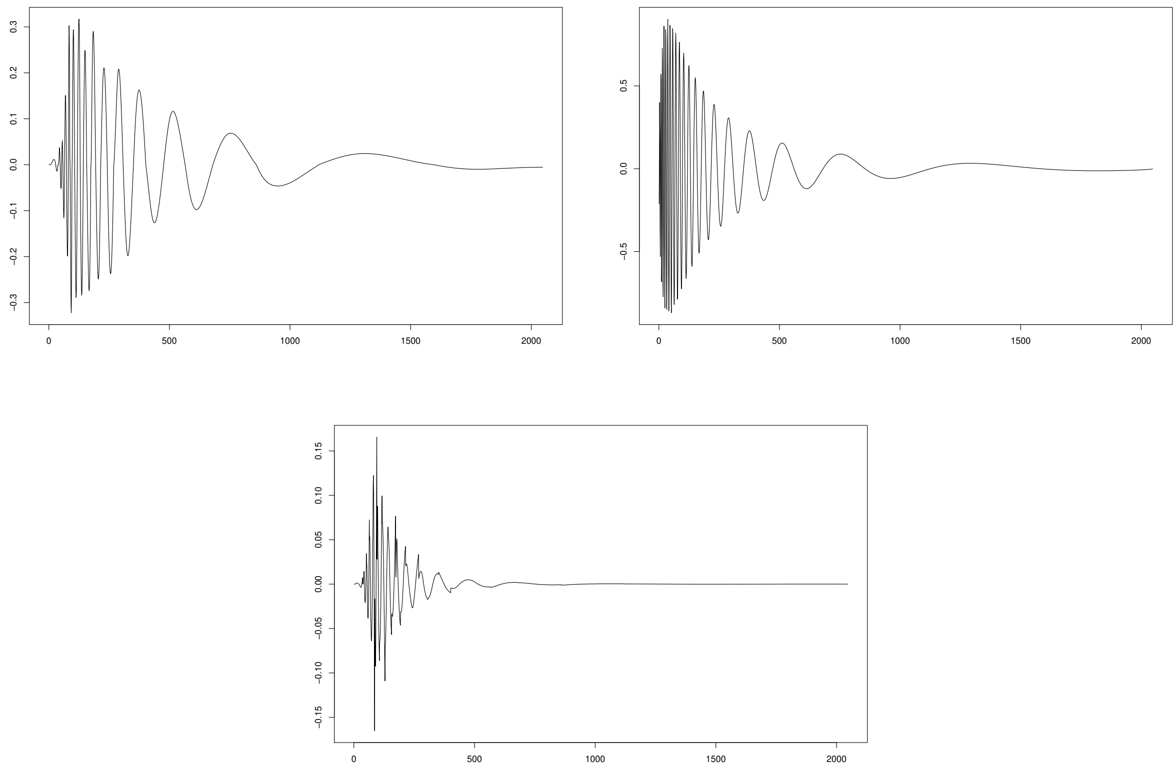


Abbildung 5.7: Die Differenzen ersten Grades des Ergebnisses im linken Bild, liefern ohne zusätzlichen Aufwand, eine gute Approximation der Ableitung (vgl. die Ableitung der Doppler-Funktion in der rechten Abbildung). Die Differenzen zweiter Ordnung im unteren Bild, liefern ebenfalls eine glatte Approximation für die zweite Ableitung.

Kapitel 6

Anwendung auf die Dichteschätzung

In diesem Kapitel wird zunächst das Problem der Dichteschätzung vorgestellt, um anschließend die Resultate aus den letzten Kapiteln hierauf anzuwenden. Für eine ausführliche Einführung in Dirchteschätzung verweisen wir z.B. auf [60].

6.1 Problemstellung

Gegeben sei ein Datensatz

$$\{x_1, \ldots, x_n\}.$$

Wir nehmen an, dass die Daten einen stochastischen Hintergrund haben. D.h. es gibt $X_1, \ldots, X_n$ unabhängig und identisch verteilte Zufallsvariablen mit "$x_i = X_i(\omega)$ ist eine Realisation von X_i". Um den zu Grunde liegenden Vorgang besser zu verstehen, ist es wichtig, eine Approximation der Verteilungsfunktion der Zufallsvariablen X_i

$$F_X(x) = \Pr(X \leq x),$$

bzw. der Dichtefunktion $f_X = F_X^{'}$ zu bestimmen. Hier stellen wir Methoden vor, die einen Schätzer für die Dichtefunktion liefern.
Für die Demonstration der Ideen in diesem Kapitel, wollen wir die Claw-Dichte verwenden.

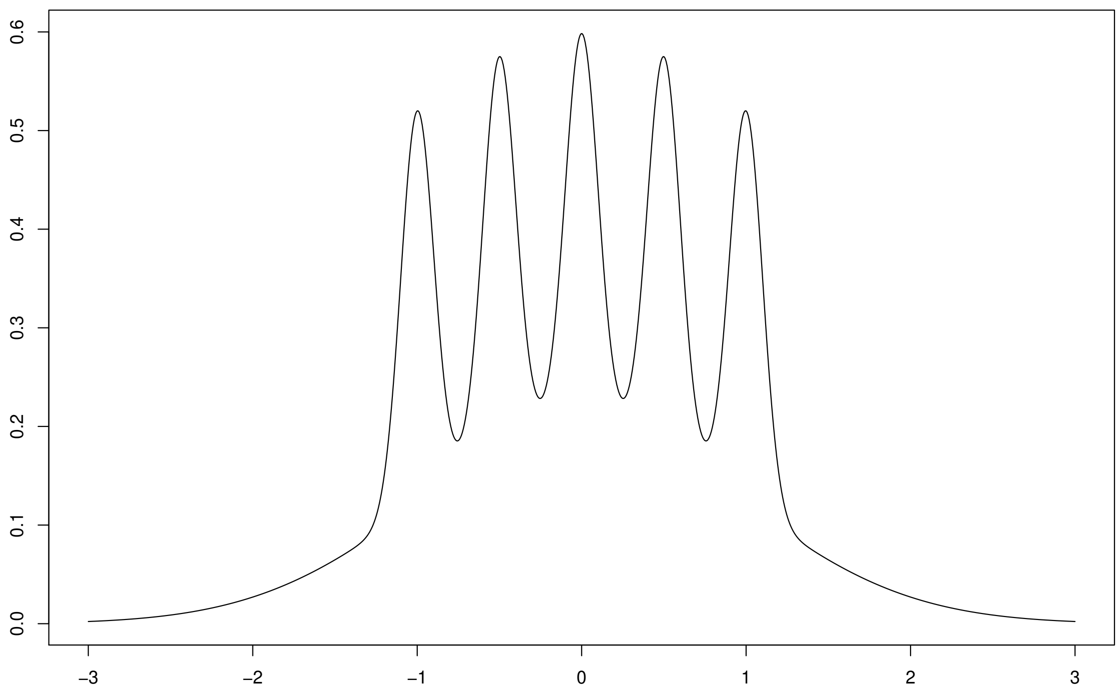


Abbildung 6.1: Die Claw-Funktion zwischen [-3,3] (vgl. [48]).

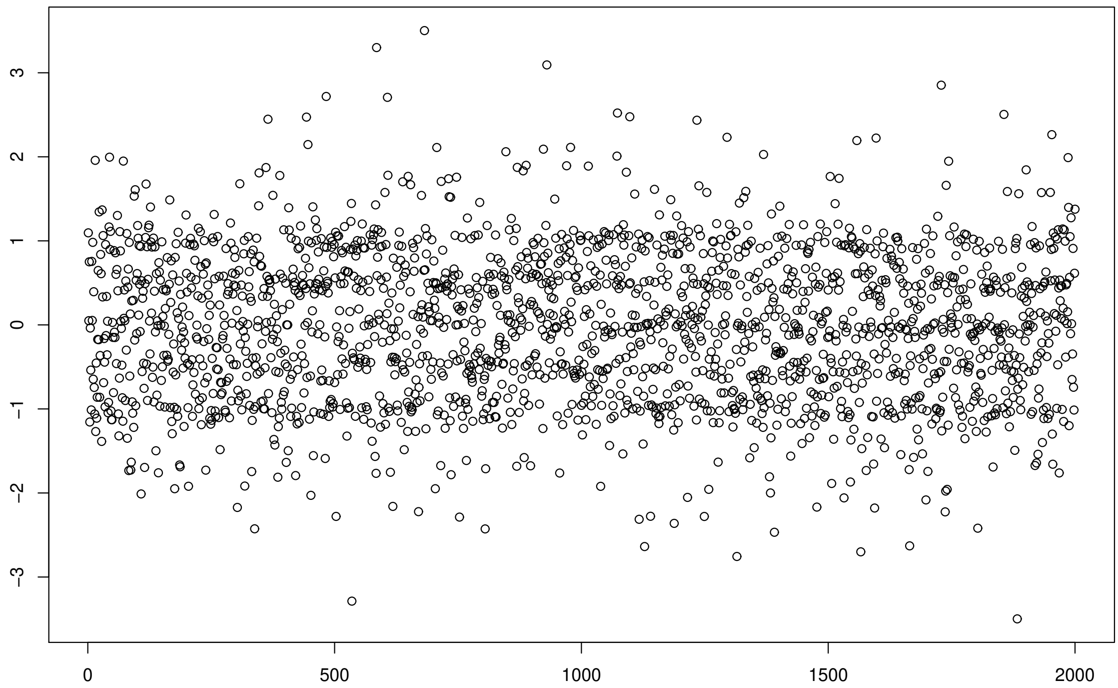


Abbildung 6.2: Ein Datensatz der Länge 2048 simuliert mit Hilfe der Claw-Dichte. Hier verwenden wir einen "großen" Datensatz, um die Effizienz der Optimierungsmethode zu demonstrieren.

Die Claw-Dichte wurde zuerst von [48] durch den Term

$$0.5 * \mathcal{N}(0,1) + 0.1 * \sum_{i=0}^{5} \mathcal{N}(i/2-1, 0.1) \tag{6.1}$$

definiert. Diese Dichte wird häufig in der Literatur benutzt (vgl. Abbildung 6.1). Ähnlich wie bei der nichtparametrischen Regression, wollen wir hier Daten (vgl. Abbildung 6.1 und Abbildung 6.2) mit der Claw-Verteilung simulieren und versuchen, die zu Grunde liegende Dichte zu approximieren.
Die einfachste Methode, eine Approximation der Dichtefunktion zu bestimmen ist der sog. *Histogramm* (vgl. Abbildung 6.3). Ist $t_1, \ldots, t_{n_H}$ eine Zerlegung des Intervalls

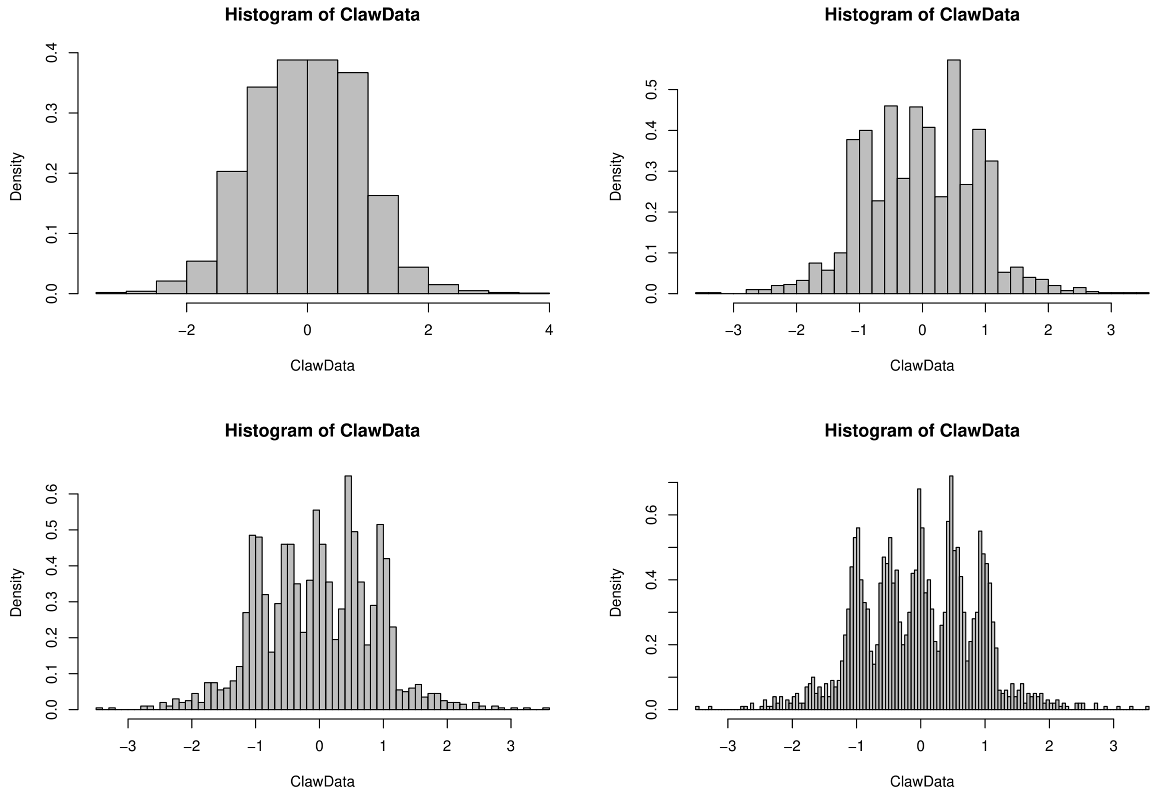


Abbildung 6.3: Verschiedene Histogramme für den Claw-Datensatz. Offensichtlich treten hier bei der Parameterwahl ähnliche Probleme auf, wie in der nichtparametrischen Regression.

$[\min\{x_i\}, \max\{x_i\}]$, so definiert man ein Histogramm mittels:

$$H(x) = \{x \in [t_i, t_{i+1}]\} \frac{|\{x_i \in [t_i, t_{i+1}]\}}{n}.$$

Meistens wird die Zerlegung des Intervalls äquidistant gemacht. Histogramme sind einfach zu berechnen und für viele Situationen sehr nützlich. Trotzdem gibt es häufig Fälle, in denen man einen besseren Schätzer benötigt. Insbesondere braucht man einen glatten Schätzer, wenn die Dichte oder sogar die Ableitung der Dichte in einer weitgehenderen Analyse der Daten verwendet wird (siehe z.B. [60], [35]). Hierfür gibt es wieder *parametrische* und *nichtparametrische* Ansätze.

Wir wollen hier als Beispiel für den nichtparametrischen Ansatz, den Kernschätzer erwähnen und verzichten auf weitere Erläuterungen. Eine Kernfunktion für Dichteschätzer hat die

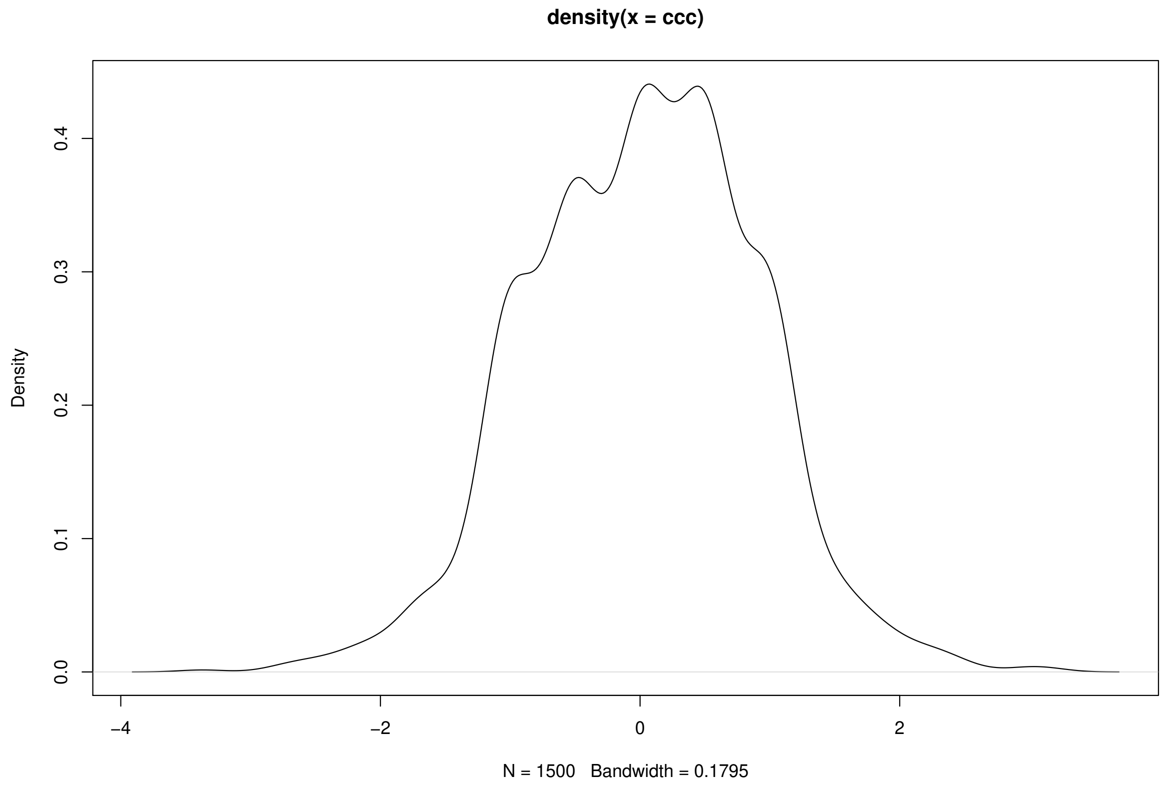


Abbildung 6.4: Ein Kernschätzer für den Claw-Datensatz. Offensichtlich sind die 5 Extremwerte überglättet worden. .

Eigenschaft:

$$\int_{-\infty}^{\infty} K(x)dx = 1.$$

Damit definiert man zu einem Datensatz $\{x_1 \ldots, x_n\}$, den Kernschätzer

$$\hat{f}(x) := \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right),$$

mit $h > 0$. Die Bandbereite beeinflusst den Schätzer ähnlich wie im Fall der Kernschätzer für die nichtparametrische Regression (vgl. Abbildung 6.4 und Abbildung 6.5).
Wir wollen hier jedoch auf genauere Ausführungen verzichten und verweisen auf [15]. Wie bereits bemerkt, wird die Dichteschätzung sehr oft mit ähnlichen Ansätzen wie bei der nichtparametrischen Regression behandelt. Daher bleibt das Problem der unbefriedigenden Approximation der Extremwerte bei der Dichteschätzung. Die zuverlässige Bestim-

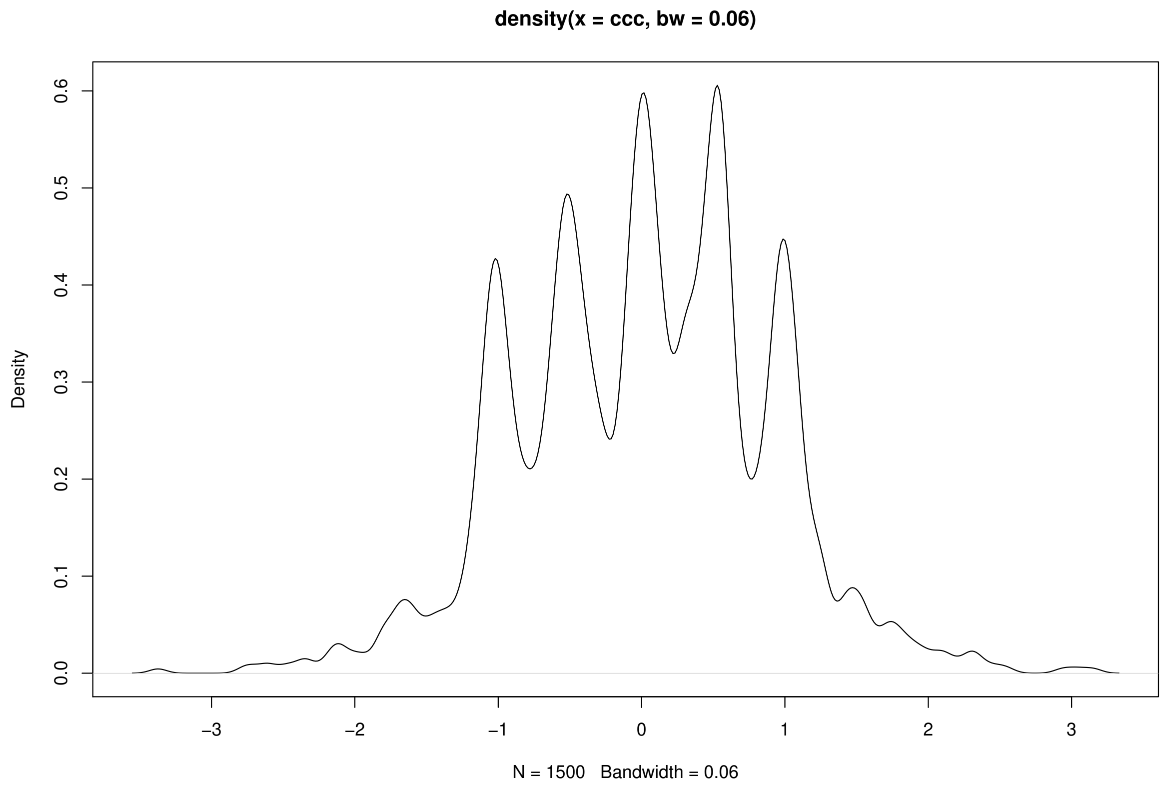


Abbildung 6.5: Ein Kernschätzer für den Claw-Datensatz mit einer kleineren Bandbereite. Die Extremstellen sind alle vorhanden. Am Rande treten jedoch Artefakte auf, die eine genauere Analyse erschweren. Sind wir an einer Ableitung der Dichtefunktion interessiert, so verstärkt sich die Wirkung dieser Artefakte.

mung der Extremwerte bei dieser Fragestellung ist mindestens genauso wichtig wie im Fall der nichtparametrischen Regression (vgl. [59]). Im Kapitel 2 haben wir die Taut-String-Methode als eine effiziente Möglichkeit, die Extremwerte zu schätzen und gleichzeitig eine Regressionsfunktion zu bestimmen, kennengelernt. Die Idee der Taut-String-Methode wurde in [15] auf das Problem der Dichteschätzung übertragen. Wir betrachten hier nur das Ergebnis und verweisen für weitere Details auf den Artikel [15]. Das Resultat der Taut-String-Methode in diesem Zusammenhang ist ebenfalls eine stückweise konstante Dichte, die mittels Grenzen l, u an der aufintegierten Funktion definiert wird. Wir wollen nun eine glatte Funktion finden, die unseren Datensatz ebenfalls approximiert. Daher definieren wir folgendes Problem

$$\begin{aligned}
&\min_{z\in\mathrm{R}^n} \tfrac{1}{2} z^T Q_2 z \text{ s.t.} \\
&l \leq Z \leq u, \\
&(z_{i+1} - z_i)\mu_i \geq 0, \quad (i = 1, \ldots, n-1), \\
&z_i \geq 0, \qquad\qquad\quad\; (i = 1, \ldots, n),
\end{aligned}$$

wobei Z die diskrete Integration von z und Q_2 wie im Kapitel 3 definiert ist. Wir haben hier lediglich die Forderung nach nichtnegativen Vektoren hinzugenommen, da das Resultat als Schätzer für eine Dichte, nichtnegativ sein muss.

Bemerkung 6.1 Der LANCELOT-Algorithmus kann für das Problem (6.2) mit einem Speicherverbrauch von $O(n)$ implementiert werden.

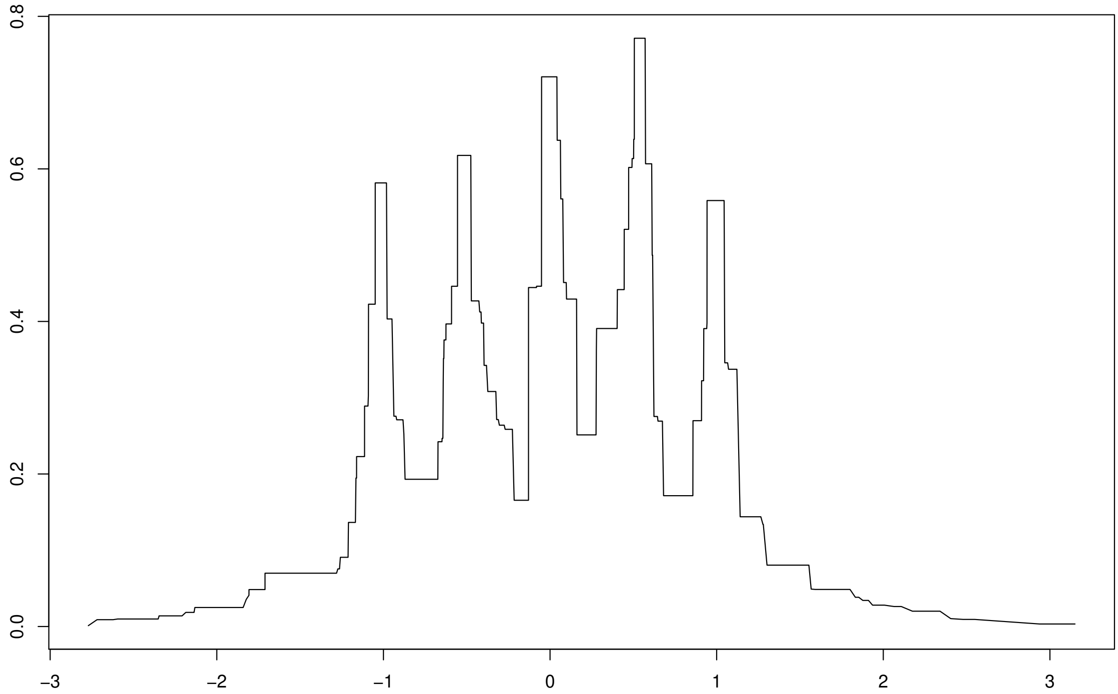


Abbildung 6.6: Das Resultat der Taut-String-Methode, angewendet auf den Claw-Datensatz. Alle Extremstellen sind ausgeprägt und es gibt keine Artefakte.

Beweis:

Betrachten wir die Funktion $\mathcal{L}_\rho$ aus dem letzten Kapitel, so führt die Optimierungsaufgabe in jedem Schritt zu einer Bandmatrix.

■

6.2 Zusammenfassung

Wir haben hier eine Methode vorgestellt, die in einem ersten Schritt, die Modalität einer Dichtefunktion anhand der zu grunde liegenden Daten ermittelt. In einem zeiten Schritt wird dann ein glatter Schätzer $\hat{f}$ für die Dichte bestimmt. Die Glättung können wir mittels der Überlegungen aus den letzten Kapiteln, effizient und schnell Durchführen. Die untere Tabelle zeigt das Ergebnis der Methode angewendet auf eine Datensatz der Länge $n = 2048$. Dabei wurden 101 Durchgänge[1] durchgeführt. In der linken Spalte stehen die Mediane der jeweiligen Rechenzeiten. In der rechten Spalte sehen wir die Anzahl der Versuche bei denen die Modalität richtig ermittelt wurde.

	Sec.	richtige Modalität
Normal	3.2	85 %
Claw	4.5	69 %

[1] Der verwendete Prozessor war hier der AMD *Athlon* 1700 XP.

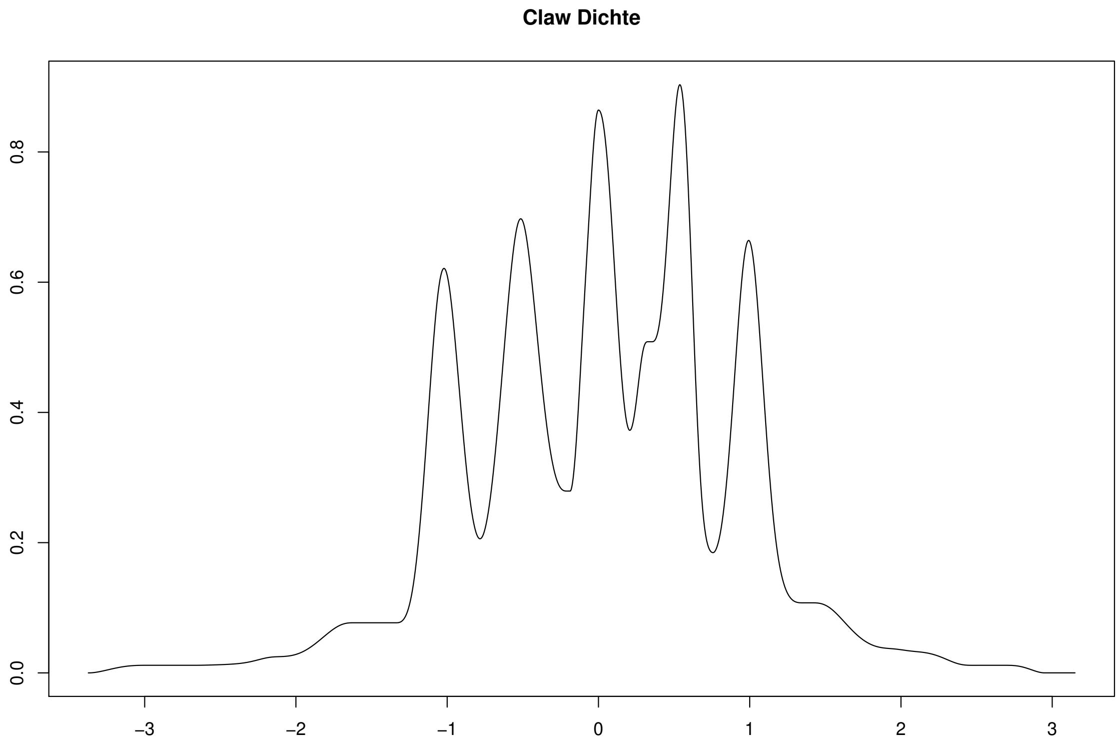


Abbildung 6.7: Das Resultat der Glättungsfunktion angewendet auf den Claw-Datensatz.

Kapitel 7

Weiterführende Gedanken

Die Zukunft hat viele Namen.

Für die Schwachen ist sie das Unerreichbare.

Für die Furchtsamen ist sie das Unbekannte.

Für die Tapferen ist sie die Chance. **VICTOR HUGO**

Es gibt zahlreiche Gedanken, die man im Anschluss an diese Arbeit verfolgen könnte. Hier gebe ich einige Anregungen für zukünftige Arbeiten:

- KAPITEL 2: Die Grundidee der Taut-String-Methode, also die Berücksichtigung der *Modalität* in der Regression, könnte auf die Regression im Mehrdimensionalen übertragen werden.

- KAPITEL 3: Die Betrachtung von *Glättungsoperatoren*, die allgemeiner sind als das hier betrachtete Funktional, scheint eine kanonische Erweiterung zu sein.

- KAPITEL 4: Die Verbesserung des Speicherbedarfs für die Quadratische Optimierung im vierten Kapitel kann auf allgemeinere Optimierungsprobleme übertragen werden. Voraussetzung hierfür ist, dass sowohl die Zielfunktion, als auch die Nebenbedingungen eine *Bandstruktur* haben.

- KAPITEL 5: Die Berechnung der *Konvergenzgeschwindigkeit* für allgemeinere *Rau-*

scherkennungstransformationen könnte interessante Erkenntnisse liefern.

- KAPITEL 6: Die Verallgemeinerung der Ergebnisse der Dichteschätzung bezüglich der Dimension sowie die theoretische Behandlung der Ergebnisse wäre ebenfalls ein erstrebenswertes Ziel in diesem Bereich.

Literaturverzeichnis

[1] J.H. Ahlberg, E.N. Nilson, and J.L. Walsh. *The Theory os Splines and Their Applications*. ACADEMIC PRESS, New York and London, 1967.

[2] L. Andersson and T. Elfving. Interpolation and Approximation by Monotone Cubic Splines. *Journal of Approximation Theory*, 66:302–333, 1991.

[3] M. Atteia. Fonctions (Spline) Definies Sur un Ensemble Convex. *Numerische Mathematik*, 12:192–210, 1968.

[4] R. Barlow, D. Bartholomew, J. Bremner, and H. Brunk. *Statistical Intference under Order Restrictions*. Wiley, New York, 1972.

[5] H. Bauer. *Maß- und Integrationstheorie*. DeGruyter, Berlin, 1990.

[6] William L. Briggs. *A Multigrid Tutorial*. SIAM, New York, 1994.

[7] Paul H. Calamai and Jorge J. Moré. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39:93–116, 1987.

[8] P. Chaudhuri and J.S. Marron. SiZer for Exploration of Structure in Curves. *Journal of the American Statistical Society*, 94, 1999.

[9] R.M. Clark. Nonparametric Estimation of Smooth Regression Function. *J. Roy. Stat. Soc.*, Ser. B 39:107–113, 1977.

[10] T.F. Coleman and L.. Jiangou. An interior Newton method for quadratic programming. *Math. Program.*, Ser. A 85:491–523, 1999.

[11] P. Copley and L.L. Schumaker. on pLg-Splines. *J. Roy. Stat. Soc.*, Ser. B 39:107–113, 1977.

[12] I.Daubechies. Nonlinear Wavelet Methods for Recovery of Signals, Images and Densities From Noisy And Incomplete Data. *American Mathematical Society*, pages 173–205, 1995.

[13] P.L. Davies. Data Features. *Statistica Neerlandica*, 49:185–245, 1995.

[14] P.L. Davies and A. Kovac. Local Extremes, Runs, Strings and Multiresolution. *to appear in Annals of statistics*, 29:1–65, 2001.

[15] P.L. Davies and A. Kovac. Densities Sepectral Densities and Modality. *under review by Annals of statistics*, 2003.

[16] R. A. De Vore. Monotone Approximation by Splines. *SIAM Journal of Mathematical Analysis*, 8:891–905, 1977.

[17] P. Deuflehard and A. Hohmann. *Numerische Mathematik I*. Walter de Gruyter, Berlin - New York, 1999.

[18] D.L. Donoho and I.M. Johnstone. Ideal Spatial Adaption by Wavelet Shrinkage. *Biometrika*, 81:425–455, 1994.

[19] Richard M. Dudley. *REAL ANALYSIS and PROBABILITY*. CHAPMAN & HALL, New York - London, 1990.

[20] L. Dümbgen. New goodness-of-fit tests and their application to nonparametric confidence sets. *Annals of Statistics*, 26:288–314, 1998.

[21] L. Dümbgen. Discussion of Local Extremes, Runs, Strings and Multiresolution. *Annals of Statistics*, 29:52–53, 2001.

[22] I. Ekeland and R. Temam. *Convex Analysis and Variational Problems*. North-Holland American Elsevier, London, 1976.

[23] Randall L. Eubank. *Spline Smoothing and Nonparametric Regression*, volume 90. Dekker, MARCELL DEKKER, INC 270 Madison Avenue, New York, New York 10016, 1 edition, 1988.

[24] J. Fan and I. Gijbels. Data-Driven Bandwidth Selection in Local Polynomial Fitting: Variable Bandwidth and Spatial Adaption. *J. Roy. Statist. Soc. Ser.*, B 57:301–370, 1995.

[25] J. Fan and J. Gijbels. *Local Polynomial Modelling and Its Applications*. Chapman & Hall, London, 1996.

[26] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, New York, 1968.

[27] A. Forsgren and W. Murray. Newton Methods for Large-Scale Linear Inequality-Constrained Minimization. *SIAM J. OPTIM.*, 7, No. 1:162–176, 1997.

[28] T. Gasser and H.-G. Müller. Estimating regression functions and their derivatives by the kernel method. *Scandinavian Journal of Statistics*, pages 171–185, 1984.

[29] Th. Gasser, H.G. Müller, W. Köhler, L. Molinari, and A. Prader. Nonparametric regression Analysis of growth curves. *Annals of Statistics*, 12:210–229, 1984.

[30] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical Optimization*. Academic Press, London, 1989.

[31] E.G. Gilstein. *Konvexe Optimierung*. Akademie-Verlag., Berlin, 1973.

[32] D. Goldfarb and A. Idnani. A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, 27:1–33, 1983.

[33] A.A. Goldstein. Convex Programming in Hilbert spaces. *Bulletin of American Mathematical Society*, 70:709–710, 1964.

[34] W. Hackbush. *Multi-Grid methods and their applications*. Springer, Berlin, 1985.

[35] Wolfgang Härdle. *Smoothing Techniques.* Springer, New York, 1991.

[36] E. Herrmann. Local bandwidth choice in kernel estimation. *Journal of Computational and Graphical Statistics*, 6:35–54, 1997.

[37] U. Hornung. Numerische Berechnung monotoner Spline-Interpolierender. *ZAMM*, 59:64–65, 1979.

[38] G. Kimeldorf and G. Wahba. Some Results on Tschebycheffian Spline Functions. *Journal of Math. Anal. and Appl.*, 33:82–95, 1971.

[39] H.W. Kuhn and A.W. Tucker. Non linear programming. *Second Berkley Sympos. Math. Stat. and Prob.*, 1951.

[40] P.J. Laurent. *Constructions of a Spline Function in a Convex Set.* Academic Press, New York, 1969.

[41] E.S. Levitin and B.T. Polyak. Constrained Minimization Problems. *USSR Computational Mathematics and Mathematical Physics*, 6:1–50, 1966.

[42] M. Löwendick. On Smoothing under Bounds and Geometric Constraints. *Dissertation, Universität GH Essen*, 2000.

[43] K. Madsen, H.B. Nielsen, and M. Pinar. A Finite Continuation Algorithm for Bound Constrained Quadratic programming. *SIAM J. OPTIM.*, 9:62–83, 1998.

[44] K. Madsen, H.B. Nielsen, and M. Pinar. Approximating quadatic programming with bound and quadratic constraints. *Math. Program.*, 84:219–226, 1999.

[45] K. Madsen, H.B. Nielsen, and M. Pinar. Bound constrained quadratic programming via piecewise quadratic functions. *Math. Program.*, 85:135–156, 1999.

[46] A. Majidi and P.L. Davies. Smooth Regression. *Proceedings of Algorithms for Approximation*, Huddersfield, July 2001.

[47] E. Mammen and S. van de Geer. Locally Adaptiv Regression Splines. *Annals of Statistics*, 25:378–413, 1997.

[48] J.S. Marron and M.P. Wand. Exact mean integrated squared error. *Annals of Statistics*, 20:712–36, 1992.

[49] M. Meise. Lokale Bandbereitenwahl in der Nichtparametrischen Regression. *Examensarbeit, Universität GH Essen*, 2001.

[50] L. Metzner. Facettierte Nichtparametrische Regression. *Dissertation, Universität Essen*, 1997.

[51] E.A. Nadaraya. *Nonparametric Estimation of Probability Densities and Regression Curves*. Kluwer Academis Publisher, New York, 1989.

[52] W. Niethammer. *Numerische Lösung von Gleichungssystemen*. FU Hagen, Hagen, 1980.

[53] Jorge Nocedal and Stephen J. Wright. *Practical Optimization*. Academic Press, London, 1989.

[54] M.R. Osborne. *Finite Algorithms in Optimization and Data Analysis*. Wiley, Australian National University, Canberra, 1985.

[55] J.O. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer, New York, 1997.

[56] J.O. Ramsay, X. Wang, and Flanagan R. Locally adaptiv regression splines. *Applied Statistics*, 44:17–30, 1995.

[57] L.I. Rudin, S. Osher, and Fatemi E. Nonlinear total variation based noise removal algorithms. *PHYSICA D*, 60:259–268, 1992.

[58] I.J. Schoenberg. Spline Functions and the Problem of Graduation. *Proc. of the National Academy of Sciences od the U.S.A.*, 52:947–950, 1964.

[59] B.W. Silverman. Using Kernel Density Estimation to investigate Multimodality. *Journal of Royal Statistical Society*, 43, B:97–99, 1981.

[60] B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London, New York, 1986.

[61] P. Spellucci. An SQP method for genereal nonlinear programs using only equality constrained subproblems. *Mathematical Programming*, 82:413–448, 1998.

[62] U. Storch and H. Wiebe. *Lehrbuch der Mathematik. Band III: Analysis mehrerer Veraenderlicher*. Wissenschaftsverlag, Mannheim, 1993.

[63] Gerardo Toraldo and Jorge J. Moré. On the solution of large quadratic programming problems with bound constraints. *SIAM J.Optimization*, 1:93–113, 1991.

[64] G. Wahba. On The Minimization of a Quadratic Functional Subject to a Continuous Family of Linear Constraints. *SIAM J. CONTROL*, 11, No. 1:64–79, 1973.

[65] G. Wahba. Improper Priors, Spline Smoothing and the Problem of Guarding Against Model in Regression. *J.R.Statist.Soc. B*, 40, No. 3:364–372, 1978.

[66] G. Wahba. Bayesian 'Confidence Intervals' for the Cross-validated Smoothing Spline. *J.R.Statist.Soc. B*, 44, No.1:133–150, 1983.

[67] Grace. Wahba. *Spline Models for Observational Data*. Society for industrial and Applied Mathematics, London, 1995.

[68] M.P. Wand and M.C. Jones. *Kernel Smoothing*. Chapman and Hall, London, 1995.

[69] J. Wegman and I. W. Wright. Isotonic, Convex and related Splines. *Annals of Statistics*, 6, No.5:1023–1035, 1980.

[70] J. Wegman and I. W. Wright. Splines in Statistics. *JASA*, 1983.

[71] T. Wichmann. Formbewahrende Regression mit Splines. *Master's Thesis*, University of Essen, 1998.

Hiermit erkläre ich, dass ich diese Arbeit ausschließlich unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Die Algorithmen in dieser Arbeit habe ich in der Programmiersprache R (bzw. in C) programmiert. Der Source-Code, mit der die Abbildungen angefertigt worden sind, steht im WWW unter
`http://www.stat-math.uni-essen.de/~majidi`
zur Verfügung.

Essen, den 20. Juni 2003   Ali Majidi