

Die Methode der kleinsten absoluten Abweichungen in linearen Modellen mit latenten Klassen und Ausreißern

Eine statistische Rechtfertigung der Methode und
Lösungsansätze aus der Standortplanung

Diplomarbeit

vorgelegt von

Robert Schieweck

aus

Berlin

angefertigt

im Institut für Numerische und Angewandte Mathematik
der Georg-August-Universität zu Göttingen

2010

Inhaltsverzeichnis

1	Einleitung	5
2	Statistische Motivation der LAD-Methode	10
2.1	Einordnung des TLAD-Schätzers	10
2.2	Statistische Grundlagen	12
2.3	Maximum-Likelihood-Eigenschaft	13
2.4	Asymptotische Eigenschaften	19
2.5	Die latent-class Regression	22
3	Platzierung einer Geraden	28
3.1	Problemstellung und Eigenschaften	28
3.1.1	Kombinatorische und ordered-median-Formulierung	28
3.1.2	Existenz und Nicht-Eindeutigkeit	31
3.2	Die geometrische Dualität	34
3.3	Eine endliche Kandidatenmenge für (TLAD)	41
4	Platzierung mehrerer Geraden	52
4.1	Problemstellung	52
4.2	NP-Vollständigkeit der Geradenplatzierung	53
4.3	Spezielle Lösungen und Existenzaussagen	57
4.4	Eine endliche Kandidatenmenge für (K-TLAD)	60
5	Algorithmen	63
5.1	Formulierung als lineares und gemischt-ganzzahliges lineares Programm	63
5.2	Der FastTLE-Algorithmus	69
5.3	Der FDS-Algorithmus	73
5.4	Vergleich der Algorithmen	74
6	Ausblick	80
	Abbildungsverzeichnis	83
	Literaturverzeichnis	84
	Danksagung	87

1 Einleitung

In der Statistik ist man häufig an einem linearen Zusammenhang zwischen reellen Daten interessiert, also an einem Zusammenhang der Form

$$Y = X\beta + \epsilon \quad (1.1)$$

mit einer Designmatrix $X \in \mathbb{R}^{n \times p}$, einem Parametervektor $\beta \in \mathbb{R}^p$ und einem n -dimensionalen Zufallsvektor ϵ . Die Designmatrix $X = (x_{i,j})_{i,j=0,\dots,p-1}$ enthält in der ersten Spalte nur Einsen, also $x_{1,0} = \dots = x_{n,0} = 1$, und hat ansonsten reelle Einträge. $(x_{i,1}, \dots, x_{i,p-1})$ geben für $i = 1, \dots, n$ verschiedene Realisationen der sogenannten unabhängigen Variablen an und $Y = (Y_1, \dots, Y_n)^t$ ist die abhängige Variable, die sich aus den unabhängigen durch einen mit dem Fehler ϵ behafteten linearen Zusammenhang erklären lässt. Man nimmt an, dass der Fehler unabhängig von der Beobachtung immer die gleiche Form hat, also unabhängig identisch verteilte $\epsilon_1, \dots, \epsilon_n$, und dass man keinen systematischen Fehler macht, also $E(\epsilon) = 0$ und somit $E(Y) = X\beta$. Außerdem setzt man häufig $V(\epsilon_i) < \infty$ voraus, um mehr statistische Mittel zur Lösung des Problems zur Verfügung zu haben. Ein solches Modell nennt man multiple lineare Regression.

In dieser Arbeit wird der Einfachheit halber nur der Fall $p = 2$ betrachtet. Damit vereinfacht sich die Modellgleichung (1.1) zu

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \forall i = 1, \dots, n \quad (1.2)$$

mit zwei reellen Parametern β_0 und β_1 . Das Modell beschreibt also eine Gerade und heißt einfache lineare Regression.

Die Aufgabe der Regression besteht nun darin, zu gegebenen Realisationen (x_i, y_i) , $i = 1, \dots, n$ die unbekannt Parameter β_0 und β_1 möglichst gut zu schätzen. Das heißt, die Modellgleichung (1.2) soll für die geschätzten Parameter $\hat{\beta}_0$ und $\hat{\beta}_1$ möglichst genau erfüllt sein:

$$y_i \approx \hat{\beta}_0 + \hat{\beta}_1 x_i \quad \forall i = 1, \dots, n.$$

Ein weit verbreitetes Verfahren dafür ist die Methode der kleinsten Quadrate, die zum ersten Mal 1805 vom französischen Mathematiker Adrien-Marie Legendre publiziert wurde

[6], [18]. Dabei werden $\hat{\beta}_0$ und $\hat{\beta}_1$ so gewählt, dass die sogenannte Summe der quadratischen Residuen

$$\sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

in β_0 und β_1 minimiert wird. Den so erhaltenen Schätzer für die Regressionsgerade nennt man Kleinste-Quadrate-Schätzer (least squares estimator, LSE). Dieser Schätzer hat viele angenehme Eigenschaften (siehe Kapitel 2), was auch der Grund für seine Beliebtheit ist. Allerdings hat er auch einen entscheidenden Nachteil: Er ist extrem anfällig gegen Punkte, die nicht der Modellgleichung folgen, sogenannte Ausreißer. Daher ist für bestimmte Anwendungen, bei denen das Auftreten solcher Ausreißer absehbar ist, ein robusterer Schätzer nötig. Ein robusterer Schätzer ist der Kleinste-Absolute-Abweichungen-Schätzer (least absolute deviations estimator, LAD). Er minimiert

$$\sum_{i=1}^n |y_i - \beta_0 - \beta_1 x_i|$$

in β_0 und β_1 . Aber auch dieser reicht häufig nicht aus, um den Einfluss der Ausreißer genügend zu kontrollieren. Er fängt zwar den Einfluss einer Fehlannahme in der Fehlerverteilung gut ab, besonders bei heavy-tail Verteilungen, die viel Wahrscheinlichkeitsmasse weit entfernt vom Mittelwert aufweisen [28]. Das bedeutet eine Robustheit in der y -Koordinate der Messpunkte. Starke Abweichungen in der x -Koordinate, sogenannte Hebelpunkte (leverage points), können jedoch mit dem LAD-Schätzer nicht effektiv behandelt werden [28]. Um hier eine höhere Robustheit zu erzielen, werden die Daten getrimmt, das heißt, man nimmt an, dass eine gewisse Anzahl $n - m$ mit $0 < m \leq n$ der Daten Ausreißer sind. Welche das genau sind, muss während der Schätzung bestimmt werden. In Kombination mit dem LAD-Schätzer wird dann die Funktion

$$\sum_{i=1}^m |y_{(i)} - \beta_0 - \beta_1 x_{(i)}| \tag{1.3}$$

in β_0 und β_1 minimiert, wobei die Permutation (\cdot) von β_0 und β_1 abhängt und die Beträge aufsteigend sortiert. Diesen Schätzer nennt man TLAD-Schätzer (trimmed least absolute deviations). So wird auch eine Robustheit gegen Hebelpunkte erzielt, da sie als Aufreißer identifiziert werden können. Man spricht hier auch von Ausreißer-resistenter Statistik, während der Begriff der Robustheit in der Literatur ursprünglich für Unempfindlichkeit gegen Fehlerverteilungen mit heavy-tails, also im Sinne einer Verteilungs-Robustheit, verwendet wurde [12, S. 4]. In dieser Arbeit wird der Einfachheit halber nur der Begriff robust ganz allgemein für die Unempfindlichkeit gegen Abweichungen von den Modellannahmen verwendet, seien es heavy-tailed Verteilungen oder Ausreißer. Ohnehin sind viele statistische Verfahren entweder robust und resistent oder keines von beidem [12, S. 4].

In der Anwendung ist außerdem die Modellierung durch einen einzelnen linearen Zusammenhang häufig nicht ausreichend. Es kann vorkommen, dass die Beobachtungen in verschiedene Klassen zerfallen und in jeder Klasse einem eigenen linearen Zusammenhang folgen. Dabei ist aber die Zugehörigkeit einer Beobachtung zu einer Klasse unbekannt. Ein Beispiel: An einer Personengruppe wird ein neues Medikament getestet, dessen Wirkung proportional zur Dosierung steigt. Es liegt also eine lineare Abhängigkeit der Wirkung von der Dosierung vor. Nun können aber Faktoren wie Vorerkrankungen bestimmter Testpersonen die Wirkung des Medikaments beeinflussen, vielleicht sogar umkehren. Wenn

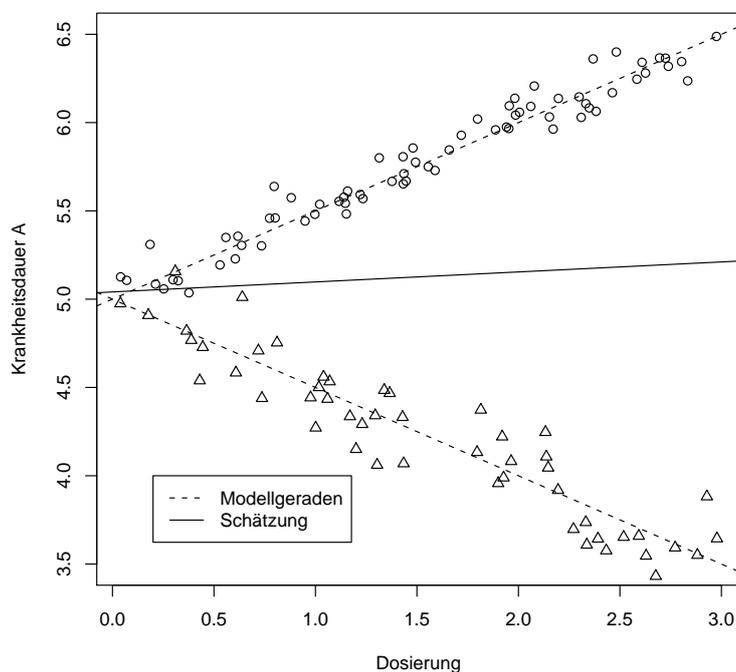


Abbildung 1.1: Kreise: 70 Patienten leiden unter Erkrankung B und das verabreichte Medikament verlängert Krankheit A;
Dreiecke: 50 Patienten haben keine Vorerkrankung, ihnen hilft das Medikament

diese Faktoren unbekannt sind, bestimmen sie die sogenannten latenten Klassen, in die die Beobachtungen zerfallen. In diesem Fall sucht man für jede der a priori unbekannt Klassen eine Regressionsgerade. Sucht man im obigen Medikamentenbeispiel nur eine einzige Gerade, könnte man einen systematischen Fehler wie in Abbildung 1.1 machen.

Hier wird die Wirkung eines Medikaments auf die Krankheitsdauer der Krankheit A untersucht. Einige der Testpersonen leiden aber auch unter Krankheit B, andere hingegen nicht. Diese Tatsache ist dem Untersuchungsteam unglücklicherweise nicht bekannt, also zerfallen die Patienten in die Klassen „hat Krankheit B“ und „hat Krankheit B nicht“, die für das Untersuchungsteam nicht zu erkennen, also latent sind. Mit Hilfe der linearen Regression wird festgestellt, dass das Medikament keine sonderlich große Wirkung hat. Richtig wäre aber, mehrere Geraden zu platzieren und mit Hilfe eines *latent class models* festzustellen, dass sehr wohl eine Wirkung gegeben ist, aber in Abhängigkeit von der unbekanntem Vorerkrankung.

In dieser Arbeit soll die Bestimmung robuster Regressiongeraden bei latenten Klassen durch die Methode der kleinsten absoluten Abweichungen als Optimierungsproblem der Standortplanung untersucht werden. Es wird also eine Anzahl von K Geraden platziert, die eine gegebene Menge von n Punkten möglichst gut approximieren soll. Als Maß für die Güte der Approximation dient der vertikale Abstand $|y - \beta_0 - \beta_1 x|$ eines Punktes $(x, y) \in \mathbb{R}^2$ von der Geraden $\{(x', y') \in \mathbb{R}^2 : y' = \beta_0 + \beta_1 x'\}$. Es sollen mögliche Ausreißer berücksichtigt werden, indem die Daten getrimmt werden und nur m Punkte als Modellpunkte angesehen werden, also als Punkte, die keine Ausreißer sind. Damit ist der Ausdruck

$$\sum_{i=1}^m \min_{k=1, \dots, K} |y_{(i)} - \beta_{1k} x_{(i)} - \beta_{0k}| \quad (1.4)$$

in den Parametern β_{0k}, β_{1k} , $k = 1, \dots, K$ der Geraden zu minimieren. Die Permutation (\cdot) sortiert die Summanden aufsteigend. Dabei sollen $K \in \mathbb{N}$ und m mit $0 < m \leq n$ feste Parameter des Problems sein. Die Bestimmung einer optimalen Zahl K^{opt} an Geraden und einer optimalen Zahl m^{opt} an Modellpunkten soll nicht behandelt werden.

In Kapitel 2 wird zunächst die Verwendung der absoluten Abweichungen im Gegensatz zur verbreiteteren Verwendung der Kleinste-Quadrate-Methode gerechtfertigt. Die Begriffe LAD-Methode und LS-Methode werden zur Unterscheidung der beiden Verfahren im Hinblick auf die verwendete Abstandsfunktion benutzt, unabhängig davon, ob getrimmt wird oder nicht, und auch unabhängig vom Vorhandensein latenter Klassen (siehe auch Tabelle 1.1). Dann werden in Kapitel 3 für den Spezialfall $K = 1$ Bedingungen für die Existenz von optimalen Geraden und eine endliche Kandidatenmenge hergeleitet. Die Kandidatenmenge liefert auch einen Algorithmus zum Lösen des Problems. In Kapitel 4 werden einige dieser Ergebnisse auf die Platzierung mehrerer Geraden verallgemeinert, insbesondere die endliche Kandidatenmenge, die wiederum ein Lösungsverfahren liefert. Es wird aber auch die NP-Vollständigkeit des Problems gezeigt, die eine effiziente Platzierung mehrerer Geraden verhindert. Kapitel 5 beschreibt einige Lösungsverfahren für das Problem. Es werden Formulierungen als mathematische Programme zur Lösung mit einem Solver angegeben. Der FastTLE-Algorithmus von Neykov u. a. [24] wird kurz

erläutert und durch eine Simulation mit der Methode des Durchprobierens der Kandidatenmenge verglichen. Schließlich liefert Kapitel 6 einen Ausblick über einige offene Fragestellungen und Möglichkeiten, die LAD-Methode zu verbessern.

Abkürzung	Bedeutung
LAD	least absolute deviations zeigt die Verwendung des Betrages als Abstandsfunktion an z. B. (LAD): Problem der Platzierung einer Geraden mit Beträgen
LS	least squares zeigt die Verwendung der Residuenquadrate als Abstandsfunktion an
T...	trimmed zeigt an, dass getrimmt wird z. B. (TLAD) bzw. (LTS): Geradenplatzierung mit Beträgen bzw. Quadraten unter Berücksichtigung von Ausreißern
K -...	zeigt die Verwendung von mehreren, nämlich K Geraden an z. B. (K -TLAD): wie (TLAD), aber mit mehreren Geraden
LP	weist auf ein lineares Programm hin
QP	weist auf ein quadratisches Programm hin
MIP	weist auf ein mixed integer programm bzw. gemischt-ganzzahliges lineares Programm hin

Tabelle 1.1: Übersicht über die in dieser Arbeit verwendeten Abkürzungen

2 Statistische Motivation der LAD-Methode

In diesem Kapitel werden einige statistische Eigenschaften des LAD-Schätzers behandelt, um sicherzustellen, dass sich seine weitere Untersuchung lohnt, und um Einsatzgebiete aufzuzeigen, in denen er dem klassischen LSE überlegen ist. Insbesondere sind wir daran interessiert, klassische Eigenschaften der LS- auf die LAD-Methode zu übertragen und zusätzlich ihre Robustheit sicherzustellen. Zuerst ordnen wir die Methode jedoch in den historischen Kontext ein.

2.1 Einordnung des TLAD-Schätzers

Robuste Statistik wird schon seit über zwei Jahrhunderten praktiziert, so publizierte Daniel Bernoulli 1777 in den *Acta Academiae Scientiarum Petropolitanae* einen Artikel mit dem Namen DIJUDICATIO MAXIME PROBABILIS PLURIUM OBSERVATIONUM DISCREPANTIUM ATQUE VERISIMILLIMA INDUCTIO INDE FORMANDA, frei übersetzt „Die wahrscheinlichste Entscheidung unter mehreren widersprüchlichen Beobachtungen und die daraus am wahrscheinlichsten abzuleitenden Folgerungen“.

Einige Jahre später äußerte sich Legendre im Zusammenhang mit der Einführung der Methode der kleinsten Quadrate zur Regression folgendermaßen über allzu große Fehler [18, S. 74]:

Si parmi ces erreurs il s'en trouve que l'on juge trop grandes pour être admises, alors on rejettera les équations qui ont produit ces erreurs, comme venant d'expériences trop défectueuses, et on déterminera les inconnues par le moyen des équations restantes, qui alors donneront des erreurs beaucoup moindres.

Auf deutsch:

Wenn unter diesen Fehlern einige sind, die man für zu groß hält, um sie zuzulassen, wird man folglich die Gleichungen, die diese Fehler verursacht haben, verwerfen, da sie aus zu fehlerhaften Experimenten hervorgegangen sind,

und man wird die Unbekannten mithilfe der übrig gebliebenen Gleichungen bestimmen, die dann viel kleinere Fehler ergeben werden ¹.

Die Vorgehensweise von Legendre erinnert ein wenig an eine Art iterative Verbesserung: Bestimme eine Regressionsgerade aus allen Beobachtungen, verwirf dann diejenigen mit großen Fehlern. Dies könnte man nun wiederholen, um zumindest eine lokal beste Lösung des Problems zu finden. Sie deutet jedoch auch schon auf das Trimmen hin, das formal für die Regression erst in den 1980er Jahren zum Beispiel von Rousseeuw [28] durch die Einführung des LTS-Schätzers (least trimmed squares) beschrieben wurde. Da sich das Trimmen durch Sortieren der Fehler und Aufsummieren der kleinsten bewerkstelligen lässt, ist der LTS-Schätzer als Linearkombination der Ordnungsstatistik der Fehler darstellbar und fällt somit in die Kategorie der L-Schätzer.

Die Idee, den Absolutbetrag anstelle des Quadrats der Residuen zu verwenden, geht auf Edgeworth im Jahre 1887 zurück [28], der mit dem langsameren Wachsen des Betrages dessen Robustheit begründete. Diese Methode liefert besonders verteilungsrobuste Schätzer, die auch bei Fehlerverteilungen mit heavy tails gute Ergebnisse liefern. Die Verwendung des Betrags wird häufig als L1-Regression bezeichnet, die der Quadrate als L2-Regression. Beide fallen in die Klasse der M-Schätzer [11], die auch den klassischen Maximum-Likelihood-Ansatz generalisieren. Bei den M-Schätzern wird nicht die negative Loglikelihoodfunktion minimiert wie bei den Maximum-Likelihood-Schätzern, sondern ein Term der Form

$$\sum_{i=1}^n \rho(y_i - \beta_1 x_i - \beta_0),$$

wobei ρ eine symmetrische Funktion mit einem einzigen Minimum bei Null ist.

Der TLAD-Schätzer kombiniert beide Ansätze, verwendet also den robusteren Betrag als Maß für die Fehler und verwirft zusätzlich die Beobachtungen mit den größten Fehlern. Es wird also

$$\sum_{i=1}^m \min_{k=1, \dots, K} |y_{(i)} - \beta_{1k} x_{(i)} - \beta_{0k}| \quad (2.1)$$

in Achsenabschnitt und Steigung β_{0k}, β_{1k} von K Geraden minimiert. Dies ist für $K = 1$ ein sogenannter trimmed likelihood estimator (TLE) nach Neykov u. a. [24], denn er hat die Form

$$\arg \max_{\theta} \sum_{i=1}^m f_{\theta}(z_{(i)}),$$

wobei θ der Parametervektor ist, der die gesuchte Gerade beschreibt, f die von θ abhängige Wahrscheinlichkeitsdichte und (\cdot) eine – auch von θ abhängige – Permutation, die die

¹Übersetzung aus dem Französischen von Gerd Schieweck

Werte der Dichte absteigend sortiert. Die Klassifikation als MLE wird in Abschnitt 2.3 genauer begründet. Für $K > 1$ bleibt die Frage nach der Klassifikation als TLE im Rahmen dieser Arbeit leider offen.

Auch der LTS-Schätzer (least trimmed squares), der analog zum TLAD die Funktion

$$\sum_{i=1}^m \min_{k=1, \dots, K} (y_{(i)} - \beta_{1k}x_{(i)} - \beta_{0k})^2$$

minimiert, ist für $K = 1$ aus den gleichen Gründen ein TLE und für ihn ist in Neykov u. a. [24] eine Heuristik angegeben, der FastTLE-Algorithmus, der auf den EM-Algorithmus aufbaut. Das Verfahren ist auch für $K > 1$ eine heuristische Methode zur Bestimmung von ML-Schätzern unter normalverteilten Fehlern. Es wird in Abschnitt 5.2 genauer beschrieben.

Es gibt noch weitere robuste Schätzer für lineare Modelle, zum Beispiel den LMS-Schätzer (least median of squares). Eine Übersicht zur robusten Regression ist in Rousseeuw und Leroy [28] zu finden. Allgemeineres zur robusten Statistik und ihrer neueren Entwicklung kann Huber und Ronchetti [12] entnommen werden. Einen historischen Überblick über ihre Anfänge liefert Stigler [33].

2.2 Statistische Grundlagen

Zuerst benötigen wir die Begriffe des statistischen Modells und des Schätzers. Hier wird nur ein grober Überblick mit dem Nötigsten zum Verständnis dieser Arbeit gegeben.

Definition 2.1 (Statistisches Modell). *Sei Z eine Zufallsvariable oder ein Zufallsvektor mit Werten in \mathcal{X} und unbekannter Verteilung P^Z auf \mathcal{X} . Nimmt man an, dass*

$$P^Z \in \mathcal{P} = \{P_\theta : \theta \in \Theta\},$$

so nennt man das Paar $(\mathcal{X}, \mathcal{P})$ ein statistisches Modell. Es sind P_θ Verteilungen auf \mathcal{X} für alle $\theta \in \Theta$ und Θ heißt Parameterraum, ein Element $\theta \in \Theta$ heißt Parameter. \mathcal{X} ist der Stichprobenraum.

Bemerkung 2.2. Haben alle Verteilungen P_θ eine Dichte p_θ , bezeichnet man manchmal auch $(\mathcal{X}, \{p_\theta : \theta \in \Theta\})$ als statistisches Modell.

Diese Definition ist eine Vereinfachung, die für unsere Zwecke hier jedoch ausreicht. Streng genommen gehört zu einem statistischen Modell auch noch eine σ -Algebra. Hier

werden wir uns aber auf den Fall $\mathcal{X} = \mathbb{R}^n$ und $\Theta = \mathbb{R}^k$ beschränken und stillschweigend mit der Borelschen σ -Algebra arbeiten. Ziel der Schätztheorie ist nun, aufgrund von Realisationen z_1, \dots, z_n von Z den wahren Parameter θ mit $P_\theta = P^Z$ durch ein $\hat{\theta} \in \Theta$ zu schätzen.

Definition 2.3 (Schätzer, [8]). *Sei $(\mathcal{X}, \mathcal{P})$ ein statistisches Modell für die Verteilung einer Zufallsvariablen Z mit Parameterraum Θ . Sei weiterhin $g : \Theta \rightarrow \Gamma$ eine Abbildung. Dann heißt eine Abbildung $t : \mathcal{X} \rightarrow \Gamma$ bzw. die Zufallsvariable $T = t(Z)$ ein Schätzer für $g(\theta)$.*

Bemerkung 2.4. Man schätzt also im Fall $\Gamma = \Theta$ und $g = \text{id}_\Theta$ den wahren Parameter der Verteilung $P^Z = P_\theta$. Dieser Fall soll hier behandelt werden.

Da nach dieser Definition keine besonderen Eigenschaften für die Abbildung t bzw. die Zufallsvariable T gefordert werden, kann jede Abbildung ein Schätzer sein. Das Schätzen kann sogar unabhängig von den Realisationen von Z erfolgen, wenn t konstant ist. Das ist im Allgemeinen natürlich nicht wünschenswert und so ist man nur an guten Schätzern mit bestimmten Eigenschaften interessiert, die gewährleisten, dass eine Schätzung sinnvoll ist. Diese werden im Folgenden behandelt und für den LAD-Schätzer beschrieben.

Für weitere Information zum Schätzen und zur Statistik im Allgemeinen sei auf Dehling und Haupt [6] verwiesen, da eine ausführlichere Behandlung der Grundlagen im Rahmen dieser Arbeit zu weit führen würde.

2.3 Maximum-Likelihood-Eigenschaft

Eine wichtige Eigenschaft von Schätzern ist die Maximum-Likelihood-Eigenschaft oder kurz ML-Eigenschaft. Solche ML-Schätzer sind für gegebene Beobachtungen besonders wahrscheinlich und man kann daher hoffen, mit ihnen das Modell möglichst gut zu beschreiben. Formal sind sie folgendermaßen definiert.

Definition 2.5 (Maximum Likelihood, [8]). *Gegeben sei ein statistisches Modell $(\mathcal{X}, \mathcal{P})$ mit einer Familie von Verteilungen $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ für die Verteilung P^Z einer Zufallsvariablen Z . Sei p_θ die Dichte von P_θ für alle $\theta \in \Theta$. Dann ist die Likelihoodfunktion*

$$L_z : \theta \rightarrow \mathbb{R}, \quad L_z(\theta) = p_\theta(z)$$

und ein Schätzer $\hat{\theta} : \mathcal{X} \rightarrow \Theta$ heißt Maximum-Likelihood-Schätzer, wenn $L_z(\theta)$ für alle $z \in \mathcal{X}$ durch Einsetzen von $\hat{\theta}(z)$ maximal wird.

Bemerkung 2.6. Auf einem diskreten Raum ist p_θ die Wahrscheinlichkeitsfunktion. Dieser Fall tritt hier aber nicht auf. Außerdem verwendet man häufig auch die Loglikelihoodfunktion

$$l_z : \theta \rightarrow \mathbb{R}, \quad l_z(\theta) = \log L_z(\theta) = \log p_\theta(z)$$

anstelle der Likelihoodfunktion. Da der Logarithmus streng monoton wächst, haben L_z und l_z dieselben Maximierer.

Wir betrachten nun die Modellgleichung (1.2) aus der Einleitung genauer: Gegeben seien Beobachtungen $(x_1, y_1), \dots, (x_n, y_n)$, von denen angenommen wird, dass sie einem einfachen linearen Modell der Form

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \text{für } i = 1, \dots, n$$

mit unbekanntem Parametern (β_0, β_1) folgen. Die ϵ_i seien unabhängig identisch verteilte Fehler mit Erwartungswert $E(\epsilon_i) = 0$ für $i = 1, \dots, n$ und unbekannter Varianz $V(\epsilon_1) = \dots = V(\epsilon_n) < \infty$.

Als Erstes leiten wir aus der Modellgleichung das Modell ab. Der Stichprobenraum ist $\mathcal{X} = \mathbb{R}^n$, denn nur die y_i werden als zufällig erzeugt angesehen, während die x_i vorher festgelegte Messstellen sind. Die unbekanntem Parameter sind β_0, β_1 und σ , also ist

$$\Theta = \{\theta = (\beta_0, \beta_1, \sigma) : \beta_0, \beta_1, \sigma \in \mathbb{R}, \sigma > 0\}$$

der Parameterraum. Die Menge \mathcal{P} der möglichen Verteilungen und damit die Form des gesamten Modells $(\mathcal{X}, \mathcal{P})$ hängt nun von der Verteilung der Fehler $\epsilon_1, \dots, \epsilon_n$ ab. Häufig nimmt man $\epsilon_1, \dots, \epsilon_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$ an. Dann sind die möglichen Verteilungen diejenigen, die eine Dichte der Form

$$\prod_{i=1}^n \phi_\theta(y_i | x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\sigma}} \exp\left(-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma}\right)$$

haben.

Für diesen Fall sind die ML-Schätzer bekannt und gegeben durch

$$\hat{\beta}_1^{\text{LSE}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{und} \quad \hat{\beta}_0^{\text{LSE}} = \bar{y} - \hat{\beta}_1^{\text{LSE}} \bar{x}.$$

Dabei sind

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{bzw.} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

die Mittelwerte der Beobachtungen in der x - bzw. y -Koordinate. Diese Schätzer existieren immer, wenn $n \geq 2$ und es $i, j \in \{1, \dots, n\}$ gibt mit $x_i \neq x_j$. Im Falle ihrer Existenz sind die Schätzer eindeutig ([8]). Man nennt $(\hat{\beta}_0^{\text{LSE}}, \hat{\beta}_1^{\text{LSE}})$ auch den Kleinste-Quadrate-Schätzer oder LSE (least squares estimator), da er die Summe der quadratischen Abstände

$$\sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$

minimiert. Er lässt sich offenbar in linearer Zeit, also sehr schnell berechnen. Der zugehörige Streuparameter σ ergibt sich im Fall normalverteilter Fehler nach der ML-Methode zu

$$(\hat{\sigma}^{\text{LSE}})^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{\beta}_0^{\text{LSE}} - \hat{\beta}_1^{\text{LSE}} x_i)^2.$$

Der LAD-Schätzer ist jedoch unter bestimmten Voraussetzungen auch ein ML-Schätzer, und zwar wenn die Fehler unabhängig identisch Laplace-verteilt sind [6], also wenn $\epsilon_1, \dots, \epsilon_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{L}(0, 2\sigma^2)$. Die Dichte der stetigen Laplace- oder Doppelsexponentialverteilung $\mathcal{L}(\mu, 2\sigma^2)$ auf \mathbb{R} ist gegeben durch

$$f(t) = \frac{1}{2\sigma} \exp\left(-\frac{|t - \mu|}{\sigma}\right)$$

mit Parametern $\mu \in \mathbb{R}$ und $\sigma > 0$.

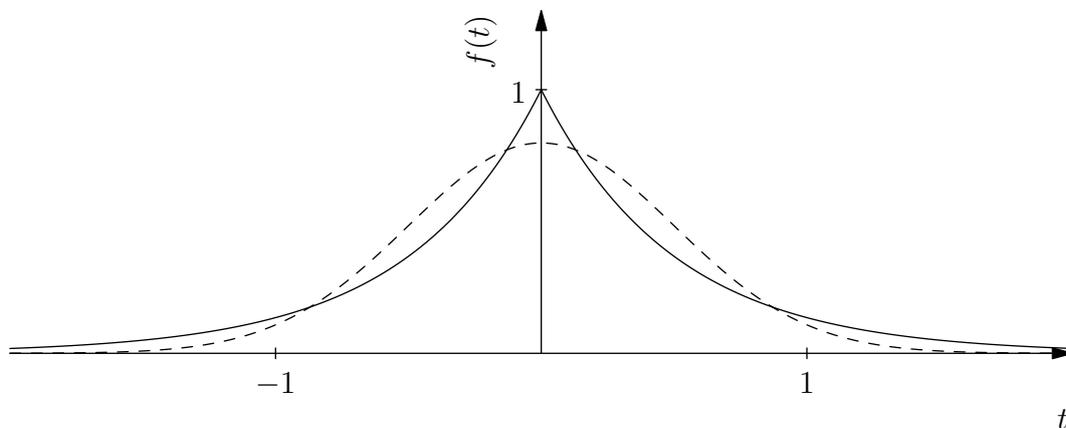


Abbildung 2.1: Dichte der Laplace-Verteilung $\mathcal{L}(0, 2\sigma^2)$ (durchgezogen) und der Normalverteilung $\mathcal{N}(0, \sigma^2)$ (gestrichelt) mit $\sigma = \frac{1}{2}$

Sie ist nicht zu verwechseln mit der Gleichverteilung auf einer endlichen Menge, die auch als Laplace-Verteilung bezeichnet wird. Eine Zufallsvariable $X \sim \mathcal{L}(\mu, 2\sigma^2)$ hat einen Erwartungswert von $E(X) = \mu$ und eine Varianz von $V(X) = 2\sigma^2$.

Lemma 2.7 (ML-Eigenschaft des LAD-Schätzers). *Für Beobachtungen $(x_1, y_1), \dots, (x_n, y_n)$ eines linearen Modells*

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \text{für } i = 1, \dots, n$$

mit Fehlern $\epsilon_1, \dots, \epsilon_n \stackrel{i.i.d.}{\sim} \mathcal{L}(0, 2\sigma^2)$ und unbekanntem Parametern $\beta_0, \beta_1, \sigma^2$ sind die ML-Schätzer $\hat{\beta}_0^{\text{LAD}}$ bzw. $\hat{\beta}_1^{\text{LAD}}$ für β_0 bzw. β_1 genau die Minimierer von

$$(LAD) \quad \min_{\beta_0, \beta_1 \in \mathbb{R}} \sum_{i=1}^n |y_i - \beta_0 - \beta_1 x_i|.$$

Diese Schätzer heißen LAD-Schätzer.

Beweis. Da die $\epsilon_1, \dots, \epsilon_n$ unabhängig sind, sind auch die Y_1, \dots, Y_n unabhängig. Damit hat ihre gemeinsame Dichte Produktform:

$$\begin{aligned} f_{\beta_0, \beta_1, \sigma}(y_1, \dots, y_n | x_1, \dots, x_n) &= \prod_{i=1}^n \frac{1}{2\sigma} \exp\left(-\frac{|y_i - \beta_0 - \beta_1 x_i|}{\sigma}\right) \\ &= \frac{1}{2^n \sigma^n} \exp\left(-\frac{1}{\sigma} \sum_{i=1}^n |y_i - \beta_0 - \beta_1 x_i|\right) \end{aligned}$$

Die Loglikelihood-Funktion ist dann gegeben durch

$$\log(f_{\beta_0, \beta_1, \sigma}(y_1, \dots, y_n | x_1, \dots, x_n)) = -\frac{1}{\sigma} \sum_{i=1}^n (|y_i - \beta_0 - \beta_1 x_i|) - n \log(2\sigma). \quad (2.2)$$

Dieser Ausdruck wird für alle $\sigma > 0$ genau dann maximal, wenn $\sum_{i=1}^n |y_i - \beta_0 - \beta_1 x_i|$ minimal wird. \square

Dieses Lemma ist eine Verallgemeinerung einer Aussage aus [6, S. 239].

Durch Einsetzen der ML-Schätzer $\hat{\beta}_0^{\text{LAD}}, \hat{\beta}_1^{\text{LAD}}$ in die Loglikelihoodfunktion (2.2) und Nullsetzen der Ableitung nach σ erhält man auch für diesen Parameter den ML-Schätzer:

$$\begin{aligned} 0 &= (\hat{\sigma}^{\text{LAD}})^{-2} \sum_{i=1}^n \left(|y_i - \hat{\beta}_0^{\text{LAD}} - \hat{\beta}_1^{\text{LAD}} x_i| \right) - \frac{n}{\hat{\sigma}^{\text{LAD}}} \\ \Leftrightarrow \hat{\sigma}^{\text{LAD}} &= \frac{1}{n} \sum_{i=1}^n |y_i - \hat{\beta}_0^{\text{LAD}} - \hat{\beta}_1^{\text{LAD}} x_i| \end{aligned} \quad (2.3)$$

Dass das tatsächlich ein Maximum ist, sieht man, indem man sich klar macht, dass die Loglikelihoodfunktion für $\sigma \rightarrow 0$ und $\sigma \rightarrow \infty$ nach unten unbeschränkt ist.

Es ist an Abbildung 2.1 schon erkennbar, dass die Annahme Laplace-verteilter Fehler robustere Schätzer hervorbringt als die Annahme normalverteilter Fehler (siehe Basset und Koenker [2]), denn die Laplaceverteilung konzentriert weniger Wahrscheinlichkeitsmasse direkt im Zentrum als die Normalverteilung (siehe Abb. 2.2).

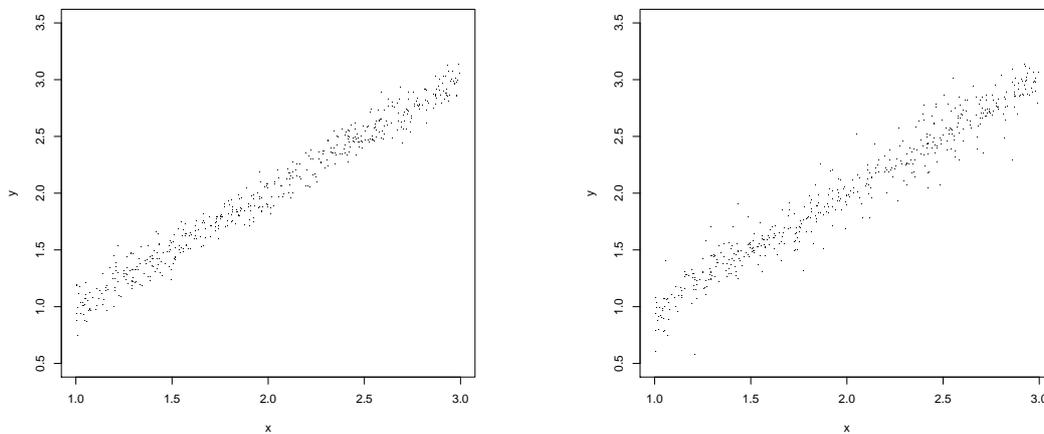


Abbildung 2.2: Lineare Modelle mit der Modellgleichung $Y_i = 0 + x_i + \epsilon_i$ für $i = 1, \dots, 500$;
 links: $\epsilon_1, \dots, \epsilon_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$; rechts: $\epsilon_1, \dots, \epsilon_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{L}(0, 2\sigma^2)$;
 in beiden Fällen: $\sigma = 0.1$

Genauer fällt sie nach außen nur in der Größenordnung $\exp(-|x|)$ ab und damit nicht so schnell wie die Normalverteilung mit der Größenordnung $\exp(-x^2)$. Das bedeutet, dass schon im Modell die Wahrscheinlichkeit der Existenz von Punkten in größerer Entfernung vom Erwartungswert durch die Verwendung Laplace-verteilter Fehler steigt. Damit haben die Ausreißer einen weniger großen Einfluss auf die Likelihoodfunktion und beeinflussen die Schätzung nicht so stark. Das ist in Abbildung 2.3 zu erkennen: Der LAD-Schätzer folgt viel besser der Geraden, die durch die Modellpunkte beschrieben wird, als der LSE, der durch die Ausreißer abgelenkt wird.

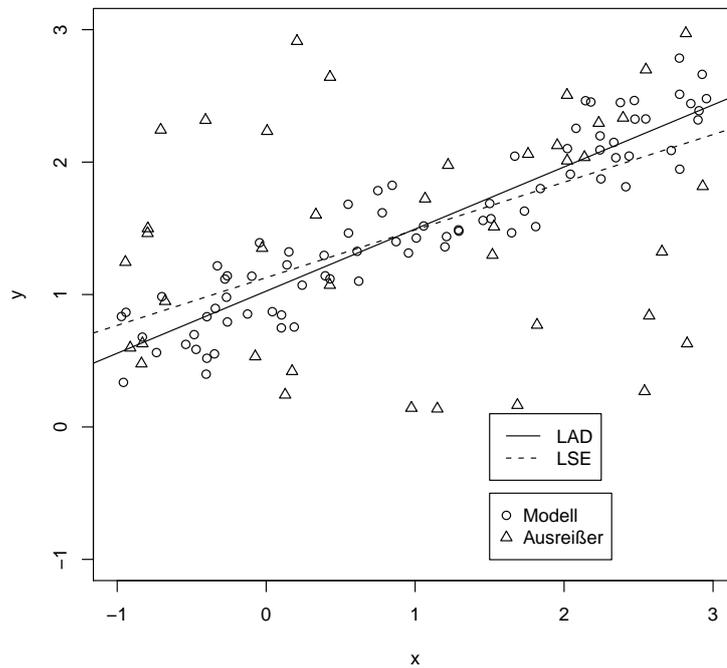


Abbildung 2.3: Kreise: 80 Punkte der Modellgleichung $Y_i = 1 + \frac{1}{2}x_i + \epsilon_i$ mit $\epsilon_1, \dots, \epsilon_n \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 0.2^2)$;
 Dreiecke: 40 Ausreißer auf $[-1, 3] \times [0, 3]$ gleichverteilt.

2.4 Asymptotische Eigenschaften

Da einzelne Beobachtungen in der Statistik aufgrund ihrer stochastischen Natur starken Schwankungen unterliegen können, sollten statistische Analysen auf einer ausreichend großen Datengrundlage aufbauen. Dies gilt auch für Schätzer wie den LAD. In diesem Abschnitt wird gezeigt, dass er zwei wichtige Eigenschaften hat, die gewährleisten, dass er bei großer Datenmenge – also einer großen Anzahl Beobachtungen (x_i, y_i) – tatsächlich gute Ergebnisse liefert, vorausgesetzt natürlich, das Modell ist korrekt. Definition 2.8 gibt eine der Eigenschaften an.

Definition 2.8 (Schwache Konsistenz, [8]). *Sei $(T_n)_{n \in \mathbb{N}}$ eine Folge von Schätzern für statistische Modelle $(\mathcal{X}_n, \mathcal{P}_n)$, die alle denselben normierten Parameterraum Θ haben und $\mathcal{P}_n = \{\otimes_{i=1}^n P_\theta, \theta \in \Theta\}$. Die Folge der Schätzer heißt konsistent für θ , falls*

$$\|T_n - \theta\| \xrightarrow{P} 0 \quad \forall \theta \in \Theta$$

in Wahrscheinlichkeit konvergiert, also

$$P(\|T_n - \theta\| \leq \epsilon) \xrightarrow{n \rightarrow \infty} 1 \quad \forall \theta \in \Theta.$$

Es ist wie in Definition 2.3 auch die Definition von Konsistenz für Funktionen $g(\theta)$ anstatt nur θ möglich, das wird hier aber nicht gebraucht.

Die Konsistenz des LAD-Schätzers folgern wir aus einer stärkeren Eigenschaft, die ebenfalls wichtig für Schätzer ist, der asymptotischen Normalität. Diese besagt, dass sich die Verteilung der absoluten Abweichungen der geschätzten Parameter $\hat{\beta}_0$ bzw. $\hat{\beta}_1$ von den wahren Werten β_0 bzw. β_1 bei Multiplikation mit einem geeigneten Faktor für große Stichproben einer Normalverteilung annähert.

Satz 2.9 (Asymptotische Normalität des LAD-Schätzers Basset und Koenker [2]). *Sei ein Modell der Form*

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \forall i = 1, 2, 3, \dots$$

gegeben mit wahren Parametern $\beta = (\beta_0, \beta_1)$. Sei $\hat{\beta}_n = (\hat{\beta}_0^n, \hat{\beta}_1^n)$ der LAD-Schätzer für Y_1, \dots, Y_n und die Fehler $\epsilon_1, \epsilon_2, \epsilon_3, \dots$ seien unabhängig identisch verteilt mit stetiger Verteilungsfunktion F , die eine in 0 positive und stetige Dichte f besitzt und deren Median verschwindet, also $F^{-1}(\frac{1}{2}) = 0$. Die Stellen x_1, x_2, x_3, \dots seien fest und X_n sei die Designmatrix für $i = 1, \dots, n$ wie in (1.1). Weiterhin gelte

$$\lim_{n \rightarrow \infty} \frac{1}{n} X_n^t X_n = Q$$

mit einer positiv definiten Matrix $Q \in \mathbb{R}^{2 \times 2}$. Dann konvergiert

$$2f(0)\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{D} Z \sim \mathcal{N}(0, Q^{-1})$$

in Verteilung gegen einen 2-dimensional normalverteilten Zufallsvektor.

Bemerkung. Konvergenz in Verteilung bei Zufallsvariablen, in Formeln $Z_n \xrightarrow{D} Z$ für $n \rightarrow \infty$, bedeutet

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad \forall x : F(x) \text{ ist stetig in } x \quad (2.4)$$

mit den Verteilungsfunktionen F_n von Z_n für alle $n \in \mathbb{N}$ und F von Z [8].

Dieser Satz wurde von Basset und Koenker in [2] in einer allgemeineren Form auch für die multiple lineare Regression und nicht nur für die einfache bewiesen. Eine Zusammenfassung dieser Ergebnisse und weiterer Voraussetzungen für diese Asymptotik finden sich in Pollard [26]. Es ist noch zu bemerken, dass Satz 2.9 für Regressionsgeraden durch den Ursprung schon 1818 von Laplace bewiesen wurde [17].

Verteilungen, die die Voraussetzungen des Satzes erfüllen, sind zum Beispiel die Normal- und die Laplaceverteilung, da sie eine auf den ganzen reellen Zahlen stetige Dichte und Median Null haben. Das heißt, obwohl der LSE bei normalverteilten Fehlern der ML-Schätzer für die Regression ist, ist dem LAD-Schätzer bei der Anwesenheit von Ausreißern der Vorzug zu geben, falls die Stichprobe groß genug ist. Denn dann liefert er aufgrund der asymptotischen Normalität und der Konsistenz aus dem nächsten Korollar gute Ergebnisse.

Korollar 2.10 (Schwache Konsistenz des LAD-Schätzers). *Für ein Modell wie in Satz 2.9 ist der LAD-Schätzer schwach konsistent.*

Beweis. Sei $\hat{\beta}_n = (\hat{\beta}_0^n, \hat{\beta}_1^n)$ die Folge der LAD-Schätzer. Wir zeigen, dass die Folge $\hat{\beta}_0^n$ konsistent für β_0 ist. Der Beweis geht dann analog für β_1 und man hat zusammen die Konsistenz von $\hat{\beta}_n$. Dazu wird Definition 2.8 geprüft, indem zu einem vorgegebenen $\delta > 0$ und $\varepsilon > 0$ ein N angegeben wird, so dass

$$P(|\hat{\beta}_0^n - \beta_0| \leq \varepsilon) \geq 1 - \delta \quad \forall n \geq N.$$

Sei F_n die Verteilungsfunktion von $\sqrt{n}|\hat{\beta}_0^n - \beta_0|$. Die Zufallsvariablen $\sqrt{n}|\hat{\beta}_0^n - \beta_0|$ folgen einer Marginalverteilung von $\|\hat{\beta}_n - \beta\|$ und sind daher nach Satz 2.9 asymptotisch $\mathcal{N}(0, \sigma^2)$ -normalverteilt mit geeignetem σ nach [8]. Die zugehörige Verteilungsfunktion sei mit Φ bezeichnet. Seien nun δ und ε gegeben. Wähle dann t so, dass

$$\Phi(t) - \Phi(-t) = 2\Phi(t) - 1 = 1 - \frac{\delta}{2}. \quad (2.5)$$

Das ist möglich, da Φ stetig und eine Verteilungsfunktion ist. Dann gilt für $\sqrt{n}\varepsilon \geq t$ bzw. $n \geq \frac{t^2}{\varepsilon^2}$

$$\begin{aligned}
 P(|\hat{\beta}_0^n - \beta_0| \leq \varepsilon) &= P(\sqrt{n}|\hat{\beta}_0^n - \beta_0| \leq \sqrt{n}\varepsilon) & (2.6) \\
 &= P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) \leq \sqrt{n}\varepsilon) - P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) < -\sqrt{n}\varepsilon) \\
 &\geq P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) \leq t) - P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) < -t) \\
 &\geq P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) \leq t) - P(\sqrt{n}(\hat{\beta}_0^n - \beta_0) \leq -t) \\
 &= F_n(t) - F_n(-t) \quad \forall n \geq \lceil t^2/\varepsilon^2 \rceil.
 \end{aligned}$$

Außerdem kann man nach Satz 2.9 N_1, N_2 wegen der Konvergenz in Verteilung nach (2.4) groß genug wählen, so dass

$$\begin{aligned}
 |F_n(t) - \Phi(t)| &\leq \frac{\delta}{4} \quad \forall n \geq N_1, & (2.7) \\
 |F_n(-t) - \Phi(-t)| &\leq \frac{\delta}{4} \quad \forall n \geq N_2.
 \end{aligned}$$

Mit $N = \max(\lceil t^2/\varepsilon^2 \rceil, N_1, N_2)$ hat man dann für alle $n \geq N$

$$P(|\hat{\beta}_0^n - \beta_0| \leq \varepsilon) \stackrel{(2.6)}{\geq} F_n(t) - F_n(-t) \stackrel{(2.7)}{\geq} \Phi(t) - \frac{\delta}{4} - \Phi(-t) - \frac{\delta}{4} \stackrel{(2.5)}{=} 1 - \delta$$

und somit die schwache Konsistenz

$$\lim_{n \rightarrow \infty} P(|\hat{\beta}_0^n - \beta_0| \leq \varepsilon) = 1 \quad \forall \varepsilon > 0.$$

□

Es bleibt noch zu klären, wie stark die Voraussetzungen von Satz 2.9 sind. Insbesondere die Frage nach der positiven Definitheit von $Q = \lim_{n \rightarrow \infty} \frac{1}{n} X_n^t X_n$ ist zu untersuchen. Da man in der Praxis meist endliche Stichproben betrachtet, lässt sich das auf die Aufgabe reduzieren, gegebene x_1, \dots, x_n so zu einer Folge $(x_i)_{i \in \mathbb{N}}$ fortzusetzen, dass

$$\lim_{i \rightarrow \infty} \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_i \end{pmatrix}}_{=X_i^t} \underbrace{\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_i \end{pmatrix}}_{X_i} = Q$$

mit einer positiv definiten Matrix Q . Da die x_i hier deterministisch sind, können sie als frei wählbare Messstellen angesehen werden. Betrachte nun

$$X_i^t X_i = \begin{pmatrix} i & \sum_{j=1}^i x_j \\ \sum_{j=1}^i x_j & \sum_{j=1}^i x_j^2 \end{pmatrix}.$$

Diese Matrix ist genau dann positiv definit, wenn ihre Hauptminoren

$$i > 0 \quad \text{und} \quad i \sum_{j=1}^i x_j^2 - \left(\sum_{j=1}^i x_j \right)^2 > 0$$

sind. Die erste Ungleichung ist für alle i erfüllt, für die zweite muss x_1, \dots, x_n für $i > n$ geeignet fortgesetzt werden. Ist die Ungleichung für $i = n$ erfüllt, setze $x_j = 0$ für alle $j > n$, ansonsten gilt $n \sum_{j=1}^n x_j^2 - \left(\sum_{j=1}^n x_j \right)^2 + C > 0$ mit einer Konstanten C , ohne Einschränkung sei $C > 1$. Wähle dann zum Beispiel $x_{n+1} = C$, $x_{n+2} = -C$ und $x_j = 0$ für alle $j > n + 2$. Dann ist

$$i \sum_{j=1}^i x_j^2 - \left(\sum_{j=1}^i x_j \right)^2 = i \sum_{j=1}^{n+2} x_j^2 - \left(\sum_{j=1}^{n+2} x_j \right)^2 > n \sum_{j=1}^n x_j^2 + 2C^2 - \left(\sum_{j=1}^n x_j \right)^2 > 0$$

für alle $i > n + 2$. Damit hat die Matrix Q nur positive Hauptminoren und ist positiv definit. Die Voraussetzungen des Satzes 2.9 an die x_i wird damit zu einer technischen Voraussetzung zur Führung des Beweises, die sich in der Praxis nicht auswirkt.

2.5 Die latent-class Regression

Wie schon in der Einleitung beschrieben, folgen die Beobachtungen bei einer latent-class Regression nicht einer einzelnen Regressiongeraden, sondern zerfallen in K Klassen, die alle ihre eigene Regressiongerade haben. Damit wird die Modellgleichung zu

$$Y_{ik} = \beta_{0k} + \beta_{1k} x_{ik} + \epsilon_{ik} \quad \forall i = 1, \dots, n_k \text{ und } \forall k = 1, \dots, K,$$

wobei sich die die Gesamtzahl der Beobachtungen zu $n = n_1 + \dots + n_K$ summiert und die Parameter β_{0k} und β_{1k} innerhalb einer Klasse k gleich sind. Ebenso ist die Verteilung der Fehler innerhalb einer Klasse gleich:

$$\epsilon_{1k}, \dots, \epsilon_{n_k k} \sim F_k \quad \forall k = 1, \dots, K.$$

Außerdem nimmt man wieder an, dass die Fehler die gleichen Voraussetzungen wie bei der normalen Regression erfüllen. Die Zugehörigkeit der Beobachtungen zu den Klassen ist a priori unbekannt und die Klassen haben jeweils eine bestimmte Wahrscheinlichkeit. Innerhalb einer Klasse hat man dieselben Annahmen wie bei der einfachen linearen Regression. Hier müssen also nicht nur die beiden Parameter einer Geraden, sondern die $2K$ Parameter von K nicht-vertikalen Geraden geschätzt werden und zusätzlich noch die Klasse jeder Beobachtung und die Mischungsverhältnisse der Klassen. Man kann sogar K als unbekannt annehmen und ebenfalls schätzen, das ist aber nicht Gegenstand dieser Arbeit.

Man führt nun Variablen π_{ik} für $i = 1, \dots, n$ und $k = 1, \dots, K$ ein (vgl. [24]), die angeben, ob Beobachtung i zu Klasse k gehört ($\pi_{ik} = 1$) oder nicht ($\pi_{ik} = 0$). Da jede Beobachtung zu genau einer Klasse gehören muss, gilt

$$\sum_{k=1}^K \pi_{ik} = 1 \quad \forall i = 1, \dots, n. \quad (2.8)$$

Man fasst die Vektoren $\vec{\pi}_i = (\pi_{i1}, \dots, \pi_{iK})$ als sogenannte fehlende Information auf, die zu den Beobachtungen gehört, aber nicht explizit beobachtet werden kann und daher geschätzt werden muss. Eine Beobachtung $(x_i, y_i, \vec{\pi}_i)$ besteht nun also aus mehr Komponenten als in der einfachen Regression, wo nur die Punkte (x_i, y_i) selbst relevant waren.

Weitere Variablen $\pi_k \in [0, 1]$ für $k = 1, \dots, K$ geben das Mischungsverhältnis der verschiedenen Klassen an, das heißt, die Wahrscheinlichkeit, dass eine Beobachtung zu Klasse k gehört, ist π_k . Daher gilt wieder

$$\sum_{k=1}^K \pi_k = 1. \quad (2.9)$$

Mit Hilfe dieser neuen Variablen können wir nun die Dichte des zugrunde liegenden Modells bestimmen. Wir nehmen an, dass die Dichte einer jeden Klasse k gegeben ist durch $p_{\theta_k}(y|x)$ mit den Parametern θ_k der Klasse k und unter der Annahme, dass x deterministische Messstellen sind. Die vollständige Dichte des Modells erhält man, indem man mit bedingten Wahrscheinlichkeiten arbeitet. Die Dichte einer jeden Beobachtung i ist bedingt durch ihre Klassenzugehörigkeit $\vec{\pi}_i$ und ist genau $p_{\theta_k}(y_i|x_i)$, wenn $\pi_{ik} = 1$. Anders geschrieben ergibt sich

$$p_{\theta}(y_i, \vec{\pi}_i|x_i) = p_{\theta}(y_i|x_i, \vec{\pi}_i) \cdot p(\vec{\pi}_i) = \prod_{k=1}^K p_{\theta_k}(y_i|x_i)^{\pi_{ik}} \pi_k^{\pi_{ik}}$$

mit $\theta = (\theta_1, \dots, \theta_K)$. Da die einzelnen Beobachtungen unabhängig sein sollen, ist die

Dichte des Modells

$$p_{\theta}(y_1, \vec{\pi}_1, \dots, y_n, \vec{\pi}_n | x_1, \dots, x_n) = \prod_{i=1}^n \prod_{k=1}^K p_{\theta_k}(y_i | x_i)^{\pi_{ik}} \pi_k^{\pi_{ik}}. \quad (2.10)$$

Als Funktion von θ und den π_k und $\vec{\pi}_i$ betrachtet, ist das die Likelihoodfunktion des Modells.

Wir betrachten diese nun im Falle Laplace-verteilter Fehler genauer. Die Likelihoodfunktion ist dann

$$L_{x_1, y_1, \dots, x_n, y_n}(\theta, \pi_1, \dots, \pi_K, \vec{\pi}_1, \dots, \vec{\pi}_n) = \prod_{i=1}^n \prod_{k=1}^K \pi_k^{\pi_{ik}} \frac{1}{(2\sigma_k)^{\pi_{ik}}} \exp\left(-\frac{|y_i - \beta_{1k}x_i - \beta_{0k}|}{\sigma_k}\right)^{\pi_{ik}}$$

wegen

$$p_{\theta_k}(y_i | x_i) = \frac{1}{2\sigma_k} \exp\left(-\frac{|y_i - \beta_{1k}x_i - \beta_{0k}|}{\sigma_k}\right)$$

mit $\theta_k = (\beta_{0k}, \beta_{1k}, \sigma_k)$. Logarithmieren ergibt die Loglikelihoodfunktion

$$l_{x_1, y_1, \dots, x_n, y_n}(\theta, \pi_1, \dots, \pi_K, \vec{\pi}_1, \dots, \vec{\pi}_n) = \sum_{i=1}^n \sum_{k=1}^K \left(\pi_{ik} (\ln(\pi_k) - \ln(2\sigma_k)) - \pi_{ik} \frac{|y_i - \beta_{1k}x_i - \beta_{0k}|}{\sigma_k} \right).$$

Man führt nun die Variablen $n_k := \sum_{i=1}^n \pi_{ik}$ für alle $k = 1, \dots, K$ ein, die die Anzahl der Beobachtungen in Klasse k angeben. Damit wird die Loglikelihoodfunktion zu

$$\sum_{k=1}^K \left(n_k (\ln(\pi_k) - \ln(2\sigma_k)) - \sum_{i=1}^n \pi_{ik} \frac{|y_i - \beta_{1k}x_i - \beta_{0k}|}{\sigma_k} \right),$$

wobei $\sum_{k=1}^K \pi_k = 1$ und $\sum_{k=1}^K n_k = n$ gelten müssen. Wir maximieren diese nun, indem wir einige Parameter festhalten und in Abhängigkeit von diesen Bedingungen an die anderen Parameter finden, die im Maximum gelten müssen. Zunächst seien nur die π_k variabel. Da diese der Einschränkung $\sum_{k=1}^K \pi_k = 1$ unterliegen und die Loglikelihoodfunktion in ihnen differenzierbar ist, verwenden wir Lagrange'sche Multiplikatoren [9, S. 109f]. Die partiellen Ableitungen nach π_k sind

$$\frac{\partial l}{\partial \pi_k} = \frac{n_k}{\pi_k} \quad \forall k = 1, \dots, K.$$

Die Einschränkungen an π_k werden durch

$$g(\pi_1, \dots, \pi_K) = \sum_{k=1}^K \pi_k - 1 \stackrel{!}{=} 0$$

realisiert. Wegen $\frac{\partial g}{\partial \pi_k} = 1$ folgt aus den Lagrange'schen Multiplikatoren

$$\left(\frac{n_1}{\pi_1}, \dots, \frac{n_K}{\pi_K} \right) = \lambda(1, \dots, 1)$$

im Maximum. Daraus erhält man zunächst $n_k/\lambda = \pi_k$, dann

$$\frac{n}{\lambda} = \frac{\sum_{k=1}^K n_k}{\lambda} = \sum_{k=1}^K \frac{n_k}{\lambda} = \sum_{k=1}^K \pi_k = 1$$

und wegen $\lambda = n$ schließlich $n_k = n\pi_k$ für alle $k = 1, \dots, K$. Anschaulich bedeutet das, dass die Schätzungen für die π_k und die π_{ik} miteinander harmonieren: Die Verhältnisse der Klassen, die sich aus den π_{ik} ergeben, entsprechen den geschätzten a-priori-Klassenwahrscheinlichkeiten π_k . Damit kann zur Maximierung die vereinfachte Loglikelihoodfunktion

$$\begin{aligned} & \sum_{k=1}^K n\pi_k(\log(\pi_k) - \log(2\sigma_k)) - \sum_{k=1}^K \sum_{i=1}^n \pi_{ik} \frac{|y_i - \beta_{1k}x_i - \beta_{0k}|}{\sigma_k} \\ &= \sum_{k=1}^K n\pi_k(\log(\pi_k) - \log(2\sigma_k)) - \sum_{k=1}^K \frac{1}{\sigma_k} \underbrace{\sum_{\substack{i=1 \\ \pi_{ik}=1}}^n |y_i - \beta_{1k}x_i - \beta_{0k}|}_{=: A_k} \end{aligned}$$

betrachtet werden. Nun maximieren wir in σ_k durch einfaches Differenzieren und Nullsetzen:

$$\frac{\partial l}{\partial \sigma_k} = -\frac{n\pi_k}{\sigma_k} + \frac{A_k}{\sigma_k^2} \stackrel{!}{=} 0$$

Es folgt $\sigma_k = A_k/(n\pi_k) = A_k/n_k$, was einem Maximum entspricht, weil die zweite Ableitung

$$\frac{\partial^2 l}{\partial \sigma_k^2} = \frac{n\pi_k}{\sigma_k^2} - \frac{2A_k}{\sigma_k^3}$$

beim Einsetzen von $\sigma_k = A_k/n_k$ einen echt negativen Wert von $-n_k^3/A_k^2$ liefert. Wieder kann man hierfür eine anschauliche Interpretation geben. Die ML-Schätzung

$$\frac{1}{n_k} \sum_{\substack{i=1 \\ \pi_{ik}=1}}^n |y_i - \beta_{1k}x_i - \beta_{0k}|$$

für σ_k liefert den gleichen Wert, als würde man nur die Punkte der Klasse k betrachten und nur in diesen einen ML-Schätzer suchen, siehe auch Gleichung (2.3). Die Loglike-

likelihoodfunktion vereinfacht sich weiter zu

$$\begin{aligned} & \sum_{k=1}^K n\pi_k \left(\log(\pi_k) - \log\left(\frac{2A_k}{n\pi_k}\right) \right) - \sum_{k=1}^K \frac{n_k}{A_k} \underbrace{\sum_{\substack{i=1 \\ \pi_{ik}=1}}^n |y_i - \beta_{1k}x_i - \beta_{0k}|}_{=A_k} \\ &= \sum_{k=1}^K n\pi_k \log\left(\frac{n\pi_k^2}{2A_k}\right) - n. \end{aligned}$$

Weiterhin ist klar, dass für feste π_{ik} die Ausdrücke

$$A_k = \sum_{I_k=\{i=1,\dots,n:\pi_{ik}=1\}} |y_i - \beta_{1k}x_i - \beta_{0k}| \quad k = 1, \dots, K$$

für jedes k in β_{0k}, β_{1k} getrennt maximiert werden können, da $\sum_{k=1}^K \pi_{ik} = 1$ gilt, also die Indexmengen I_1, \dots, I_K paarweise disjunkt sind. Also entspricht die Wahl der K Achsenabschnitte und K Steigungen K Regressionen ohne latente Klassen wie in Lemma 2.7.

Ebenso ist klar, dass für feste Klassenwahrscheinlichkeiten π_k die Klassenzugehörigkeiten π_{ik} zur Maximierung der Loglikelihoodfunktion so gewählt werden, dass jeder der Ausdrücke A_k minimal wird. Damit haben wir nun alle Werte Parameter $\beta_{0k}, \beta_{1k}, \sigma_k$ und π_{ik} im Maximum der Loglikelihoodfunktion in Abhängigkeit der Werte der Klassenwahrscheinlichkeiten π_k bestimmt. Das fassen wir in folgendem Satz zusammen:

Satz 2.11. *Für eine Regression mit latenten Klassen, die dem Modell*

$$Y_{ik} = \beta_{0k} + \beta_{1k}x_{ik} + \epsilon_{ik} \quad \forall i = 1, \dots, n_k \text{ und } \forall k = 1, \dots, K$$

mit Fehlern

$$\epsilon_{1k}, \dots, \epsilon_{n_k k} \stackrel{i.i.d.}{\sim} \mathcal{L}(0, 2\sigma_k^2) \quad \forall k = 1, \dots, K$$

folgt, sind die ML-Schätzer der Parameter bei bekannten Klassenwahrscheinlichkeiten π_k , $k = 1, \dots, K$ gegeben durch

$$\begin{aligned} \hat{n}_k &= n\pi_k && \text{(Klassengrößen)} \\ (\hat{\beta}_{01}, \dots, \hat{\beta}_{1K}, \hat{\pi}_{11}, \dots, \hat{\pi}_{nK}) &= \arg \min_{\substack{\pi_{ik}, \beta_{0k}, \beta_{1k}: \\ \pi_{ik} \in \{0,1\} \\ \sum_{i=1}^n \pi_{ik} = \hat{n}_k \\ \sum_{k=1}^K \pi_{ik} = 1}} \sum_{k=1}^K A_k && \text{(Geraden, Zugehörigkeiten)} \\ \hat{\sigma}_k &= \frac{1}{\hat{n}_k} \sum_{\substack{i=1 \\ \hat{\pi}_{ik}=1}}^n |y_i - \beta_{1k}x_i - \beta_{0k}| && \text{(Streuparameter)} \end{aligned}$$

für alle $k = 1, \dots, K$.

Beweis. Der Beweis ergibt sich aus den Rechnungen der vorigen Seiten. \square

Es ist noch nicht gezeigt, dass die minimierenden Geradenparameter der Likelihoodfunktion bei latenten Klassen ($K > 1$) den K -LAD-Schätzern entsprechen, die aus der Minimierung von (2.1) hervorgehen. Es ist also auch noch nicht gezeigt, dass der K -TLAD-Schätzer (K -LAD mit Trimmen) ein TLE nach Neykov u. a. [24] ist. Man kann aber hoffen, dass der Unterschied nicht allzu groß ist. Die Simulation in Abschnitt 5.4 legt zumindest nahe, dass man mit der LAD-Methode nahe bei den wahren Parametern liegt.

3 Platzierung einer Geraden

3.1 Problemstellung und Eigenschaften

3.1.1 Kombinatorische und ordered-median-Formulierung

Für die lineare Regression ohne Klasseneinteilung ist die Bestimmung eines LAD-Schätzers ein Spezialfall des allgemeinen Problems, zu gegebenen Punkten $z_1, \dots, z_n \in \mathbb{R}^2$ eine Gerade l im \mathbb{R}^2 zu finden, sodass die Summe der vertikalen Abstände von der Geraden zu den Punkten minimal wird. Für TLAD sollen jedoch für ein beliebiges m mit $0 < m \leq n$ nur m Punkte berücksichtigt werden bzw. die $n-m$ Punkte vernachlässigt werden, die die größten Abstände von der optimalen Geraden haben. Die gesuchte Gerade kann zunächst einmal vertikal sein oder nicht. Im Folgenden wird dieses Problem als mathematisches Programm formuliert und es werden Bedingungen an die Punkte z_1, \dots, z_n hergeleitet, unter denen eine optimale Gerade für das Programm immer nicht-vertikal ist, also einem TLAD-Schätzer entspricht. Zunächst vereinfachen einige Notationen die Formulierung.

Notation 3.1. *Im Folgenden bezeichne*

$$\ell(\beta_0, \beta_1) = \{(x, y) \in \mathbb{R}^2 : y = \beta_0 + \beta_1 x\}$$

eine nicht-vertikale Gerade im \mathbb{R}^2 mit Steigung β_1 und Ordinatenabschnitt β_0 . Weiter sei für eine Gerade $l = \ell(\beta_0, \beta_1)$ ihre Höhe an der Stelle $x \in \mathbb{R}$ mit $l_x = \beta_0 + \beta_1 x$ bezeichnet.

Auf diese Weise kann man alle nicht-vertikalen Geraden eindeutig darstellen. Weiterhin müssen noch die vertikalen Abstände erklärt werden. Dies geschieht in der auch intuitiv naheliegenden Weise in Notation 3.2.

Notation 3.2 (Vertikaler Abstand, [30]). *Für zwei Punkte $z_1 = (x_1, y_1)$ und $z_2 = (x_2, y_2)$ im \mathbb{R}^2 ist der vertikale Abstand von z_1 und z_2 gegeben durch*

$$d_v(z_1, z_2) = \begin{cases} |y_2 - y_1| & \text{falls } x_1 = x_2 \\ \infty & \text{falls } x_1 \neq x_2 \end{cases}.$$

Für eine Gerade l im \mathbb{R}^2 und einen Punkt $z = (x, y) \in \mathbb{R}^2$ ist ihr vertikaler Abstand gegeben durch

$$d_v(z, l) = \begin{cases} |y - (\beta_1 x + \beta_0)| & \text{falls } l = \ell(\beta_0, \beta_1) \\ 0 & \text{falls } l = \{(x', y') \in \mathbb{R}^2 : x' = x\} \\ \infty & \text{sonst} \end{cases}$$

Es werden $d_v(z, l)$ und $d_v(l, z)$ gleichwertig verwendet.

Um das Weglassen der $n - m$ Punkte mit den größten Abständen zur optimalen Gerade zu beschreiben, sei $(1), \dots, (n)$ eine Permutation von $\{1, \dots, n\}$, sodass

$$d_v(z_{(1)}, l) \leq d_v(z_{(2)}, l) \leq \dots \leq d_v(z_{(n)}, l). \quad (3.1)$$

Diese Permutation hängt natürlich von der Geraden l ab. Insbesondere ist eine solche Permutation für die optimale Gerade, die das Problem löst, unbekannt, da ja die Gerade unbekannt ist. Sie muss also auch beim Lösen des mathematischen Programms, das sich nun wie folgt formulieren lässt, erst bestimmt werden:

$$\min_{l \text{ ist Gerade im } \mathbb{R}^2} \sum_{i=1}^m d_v(z_{(i)}, l) \quad (3.2)$$

für ein gegebenes m mit $0 < m \leq n$. Da wir im Zuge der linearen Regression an einem linearen Zusammenhang zwischen den x - und y -Koordinaten der Punkte z_1, \dots, z_n interessiert sind und eine vertikale Gerade keinen solchen Zusammenhang darstellt, wollen wir jetzt Bedingungen herleiten, unter denen man sich in (3.2) auf nicht-vertikale Geraden der Form $l = \ell(\beta_0, \beta_1)$ beschränken kann. Dazu formulieren wir das Problem äquivalent in kombinatorischer Form [24]:

$$\min_{l \text{ ist Gerade im } \mathbb{R}^2} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} d_v(z_i, l) \quad (3.3)$$

Die Formulierungen in (3.2) und (3.3) sind offenbar äquivalent, denn der zulässige Bereich ist bei beiden der \mathbb{R}^2 und die Zielfunktionen sind gleich, denn für jede feste Gerade l im \mathbb{R}^2 werden im inneren kombinatorischen Optimierungsproblem in Formulierung (3.3) genau die m Indizes i_1, \dots, i_m mit kleinsten Werten $d_v(z_{i_j}, l)$ in die Menge I aufgenommen, also gilt für jede Lösung dieses inneren Problems $I = \{(1), \dots, (m)\}$, wobei die Permutation (\cdot) Gleichung (3.1) genügt.

Das nächste Lemma liefert jetzt die Bedingungen, unter denen $l = \ell(\beta_0, \beta_1)$ ohne Einschränkung angenommen werden kann.

Lemma 3.3 (Beschränkung auf nicht-vertikale Geraden). *Seien Punkte $z_1, \dots, z_n \in \mathbb{R}^2$ und $m : 2 \leq m \leq n$ vorgegeben mit $z_i = (x_i, y_i)$ für $i = 1, \dots, n$. Es gibt genau dann eine vertikale Gerade, die eine optimale Lösung des Problems in (3.2) ist, wenn mindestens m der gegebenen Punkte die gleiche x -Koordinate haben.*

Beweis. Der Beweis verwendet die kombinatorische Formulierung in (3.3).

\implies : Angenommen, es gibt weniger als m unter den gegebenen Punkten, die die gleiche x -Koordinate haben. Dann gibt es zu jeder Indexmenge $I \subset \{1, \dots, m\}$ mit $|I| = m$ zwei Indizes $i_1, i_2 \in I$ mit $x_{i_1} \neq x_{i_2}$. Daher hat jede vertikale Gerade l von einem dieser Punkte einen unendlich großen vertikalen Abstand: $d_v(z_{i_1}, l) = \infty$ oder $d_v(z_{i_2}, l) = \infty$. Daher gilt für jede solche Indexmenge

$$\sum_{i \in I} d_v(z_i, l) \geq d_v(z_{i_1}, l) + d_v(z_{i_2}, l) = \infty \quad \forall I \subset \{1, \dots, n\} \text{ mit } |I| = m$$

und jede nicht-vertikale Gerade ist eine bessere Lösung des Problems, da $d_v(z, \ell(\beta_0, \beta_1)) < \infty$ für alle $z \in \mathbb{R}^2$ und alle $\beta_0, \beta_1 \in \mathbb{R}$, also

$$\sum_{i \in I} d_v(z_i, \ell(\beta_0, \beta_1)) < \infty.$$

\impliedby : Angenommen, es gibt m Punkte z_{i_1}, \dots, z_{i_m} mit der gleichen x -Koordinate $x_{i_1} = \dots = x_{i_m}$. Dann gilt für die Indexmenge $I = \{i_1, \dots, i_m\}$ und die Gerade $l = \{(x, y) \in \mathbb{R}^2 : x = x_{i_1}\}$:

$$d_v(z_i, l) = 0 \quad \forall i \in I \quad \implies \quad \sum_{i \in I} d_v(z_i, l) = 0$$

Da die vertikalen Abstände aber immer größer oder gleich Null sind, ist l eine optimale Lösung des Problems in (3.3). \square

Im Folgenden gehen wir nun davon aus, dass unter den gegebenen Punkten weniger als m die gleiche x -Koordinate haben, und können uns somit auf nicht-vertikale Geraden bzw. die lineare Regression zurückziehen. Dies stellt keine besonders große Einschränkung dar, insbesondere für die Anwendung im Bereich der linearen Regression, da dort meistens genug Daten zur Verfügung stehen. Ein Beispiel: Bei 1000 Messpunkten mit einem vermuteten Ausreißeranteil von 30% ist $m = 700$, also dürfen höchstens 699 Punkte die gleiche x -Koordinate haben. Sind die Messpunkte zum Beispiel auch in der x -Koordinate zufällig erzeugt durch eine stetige Verteilung, verschwindet die Wahrscheinlichkeit gleicher x -Koordinaten ohnehin. Sind die x -Koordinaten im Experiment festlegbar, können sie

geeignet gewählt werden. Sollten doch einmal zu viele Punkte mit gleicher x -Koordinate auftreten, sollte man sich die Frage stellen, ob eine lineare Regression zur Modellierung der Daten geeignet ist.

Wir müssen nun für gegebene Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ und gegebenes m mit $0 < m \leq n$ das Problem

$$(TLAD) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{i=1}^m d_v(z_{(i)}, \ell(\beta_0, \beta_1)) \quad (3.4)$$

lösen. Es fällt in die Klasse der ordered-median-Probleme, für die eine allgemeine Theorie in Nickel und Puerto [25] beschrieben ist. Das kombinatorische Äquivalent dazu ist

$$(TLAD) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} d_v(z_i, \ell(\beta_0, \beta_1)). \quad (3.5)$$

TLAD steht für trimmed least absolute deviation, also getrimmte minimale absolute Abweichung. Die Bezeichnung stammt aus der Statistik. Im Falle $m = n$ muss nicht getrimmt werden und das Problem wird dann mit (LAD) bezeichnet. Es ist zu bemerken, dass eine Lösung von (TLAD) nicht nur aus Werten für β_0 und β_1 besteht, sondern auch aus einer Permutation (\cdot) von $\{1, \dots, n\}$ bzw. einer Indexmenge $I \subset \{1, \dots, n\}$ mit $|I| = m$, für die das Minimum angenommen wird und durch die die Ausreißer identifiziert werden. Dies wird besonders an der kombinatorischen Formulierung (3.5) deutlich, da hier tatsächlich zwei verschachtelte Minimierungsprobleme vorliegen. Daher wird im Folgenden sowohl eine optimale Gerade l^{opt} als auch ein Paar $(l^{\text{opt}}, I^{\text{opt}})$ aus einer optimalen Geraden und zugehöriger Indexmenge als optimale Lösung von (TLAD) bezeichnet. Außerdem wird der kombinatorischen Formulierung aufgrund ihrer intuitiven Beschreibung der Punkte, die dem Modell folgen, nämlich durch eine Teilmenge anstelle einer Permutation, bei der Herleitung von Eigenschaften des (TLAD) häufig der Vorzug gegeben.

3.1.2 Existenz und Nicht-Eindeutigkeit

Zuerst ist man natürlich an Existenz- und Eindeutigkeitsaussagen für dieses Optimierungsproblem interessiert. Diese werden im Folgenden formuliert.

Lemma 3.4. *Für zwei Punkte $z_1 = (x_1, y_1)$ und $z_2 = (x_2, y_2)$ mit $x_1 \neq x_2$ gilt*

$$d_v(z_1, \ell(\beta_0, \beta_1)) + d_v(z_2, \ell(\beta_0, \beta_1)) \xrightarrow{\|(\beta_0, \beta_1)\|_{2 \rightarrow \infty} \rightarrow \infty} \infty.$$

Beweis. Angenommen, das wäre nicht so. Dann gibt es eine Folge $(\beta_{0n}, \beta_{1n})_{n \in \mathbb{N}}$ mit $\|(\beta_{0n}, \beta_{1n})\|_2 \rightarrow \infty$ für $n \rightarrow \infty$, so dass

$$d_v(z_1, \ell(\beta_{0n}, \beta_{1n})) + d_v(z_2, \ell(\beta_{0n}, \beta_{1n})) \leq C \quad \forall n \in \mathbb{N}$$

und ein festes $C \in \mathbb{R}$. Außerdem gilt

$$\|(\beta_{0n}, \beta_{1n})\|_2 \rightarrow \infty \quad \implies \quad |\beta_{0n}| \rightarrow \infty \text{ oder } |\beta_{1n}| \rightarrow \infty. \quad (3.6)$$

Da die vertikalen Abstände nicht-negativ sind, gelten dann auch

$$\begin{aligned} |y_1 - \beta_{1n}x_1 - \beta_{0n}| &= d_v(z_1, \ell(\beta_{0n}, \beta_{1n})) \leq C & \forall n \in \mathbb{N} \\ |y_2 - \beta_{1n}x_2 - \beta_{0n}| &= d_v(z_2, \ell(\beta_{0n}, \beta_{1n})) \leq C & \forall n \in \mathbb{N}. \end{aligned}$$

Da $x_1 \neq x_2$, ist eines der beiden, ohne Einschränkung x_1 , von Null verschieden. Dann ist $\beta_{1n}x_1 + \beta_{0n} \in [y_1 - C, y_1 + C]$ und es muss

$$\beta_{0n} \rightarrow \pm\infty \quad \implies \quad \beta_{1n} \rightarrow \mp \operatorname{sgn}(x_1)\infty$$

gelten oder vereinfacht $|\beta_{0n}| \rightarrow \infty \Rightarrow |\beta_{1n}| \rightarrow \infty$. Mit (3.6) ergibt sich

$$\|(\beta_{0n}, \beta_{1n})\|_2 \rightarrow \infty \quad \implies \quad |\beta_{1n}| \rightarrow \infty \quad (3.7)$$

und man hat den Widerspruch

$$|y_2 - \beta_{1n}x_2 - \beta_{0n}| = \underbrace{|y_2 - \beta_{1n}x_1 - \beta_{0n}|}_{\text{beschränkt}} - \underbrace{(x_2 - x_1)}_{\neq 0} |\beta_{1n}| \xrightarrow{(3.7)} \infty.$$

□

Satz 3.5 (Existenz optimaler Lösungen). *Seien Punkte $z_1, \dots, z_n \in \mathbb{R}^2$ und m mit $0 < m \leq n$ vorgegeben. Dann hat (TLAD) eine optimale Lösung und falls $m \geq 2$ und unter den Punkten keine m mit gleicher x -Koordinate sind, ist jede Optimallösung endlich in Bezug auf β_0 und β_1 .*

Beweis. Der Beweis orientiert sich an einem Beweis in [25, S. 27] und verwendet die kombinatorische Variante aus (3.5). Wir betrachten dann die Probleme

$$(P(I)) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{i \in I} d_v(z_i, \ell(\beta_0, \beta_1))$$

mit $I \subset \{1, \dots, n\}$ und $|I| = m$. Von diesen gibt es nur $\binom{n}{m}$, also endlich viele, und die optimalen Lösungen der Probleme (P(I)) mit kleinstem Zielfunktionswert sind genau die optimalen Lösungen von (TLAD).

Fall 1: Für jedes I , das mindestens zwei verschiedene Indizes $i_1, i_2 \in I$ enthält, so dass $x_{i_1} \neq x_{i_2}$, gilt

$$\begin{aligned} \sum_{i \in I} d_v(z_i, \ell(\beta_0, \beta_1)) &= \sum_{i \in I} |y_i - (\beta_0 + \beta_1 x_i)| \\ &\geq |y_{i_1} - \beta_0 - \beta_1 x_{i_1}| + |y_{i_2} - \beta_0 - \beta_1 x_{i_2}| \xrightarrow{\|(\beta_0, \beta_1)\|_2 \rightarrow \infty} \infty \end{aligned}$$

nach Lemma 3.4. Also kann man für solche Indexmengen ohne Einschränkung $(\beta_0, \beta_1) \in Q_I$ für eine kompakte Menge Q_I fordern. Die stetige Zielfunktion (Komposition von linearer Funktion, Betrag, Sortierung, Summe) nimmt auf dieser Menge ihr Minimum an. Also gibt es für jedes $(P(I))$ mit solcher Indexmenge I eine optimale Lösung und alle optimalen Lösungen sind endlich bezüglich β_0 und β_1 .

Fall 2: Für jede Indexmenge I , so dass $x_i = x$ für alle $i \in I$, ist jede Gerade, die durch (x, y) verläuft, optimal für $(P(I))$ mit dem Median y der y_i , $i \in I$, denn die Zielfunktion wird zu

$$\sum_{i \in I} d_v(z_i, \ell(\beta_0, \beta_1)) = \sum_{i \in I} |y_i - \beta_1 x - \beta_0|$$

und diese wird minimal, wenn $\beta_1 x + \beta_0 = y$ nach [6, S. 239]. Es existiert also eine optimale Lösung, aber $\|(\beta_0, \beta_1)\|_2$ kann beliebig groß werden.

Zusammen ergibt sich, dass in jedem Fall eine optimale Lösung von (TLAD) existiert. Falls $m \geq 2$ und es keine m gegebenen Punkte mit gleicher x -Koordinate gibt, tritt nur der erste Fall ein und jede optimale Lösung ist endlich bezüglich β_0 und β_1 . \square

Allerdings muss die Lösung nicht eindeutig sein. Für $n = m = 4$, $z_1 = (1, 1)$, $z_2 = (1, 3)$, $z_3 = (5, 1)$ und $z_4 = (5, 3)$ gilt

$$\begin{aligned} \sum_{i=1}^k d_v(z_{(i)}, \ell(\beta_0, \beta_1)) &= \sum_{i=1}^4 |y_i - \beta_0 - \beta_1 x_i| \\ &= |1 - \beta_0 - \beta_1| + |3 - \beta_0 - \beta_1| + |1 - \beta_0 - 5\beta_1| + |3 - \beta_0 - 5\beta_1| \\ &= (\beta_0 + \beta_1 - 1) + (3 - \beta_0 - \beta_1) + (\beta_0 + 5\beta_1 - 1) + (3 - \beta_0 - 5\beta_1) \\ &= 2 + 2 = 4. \end{aligned}$$

für alle (β_0, β_1) mit $\beta_0 + \beta_1 \in [1, 3]$ und $\beta_0 + 5\beta_1 \in [1, 3]$ (siehe Abb. 3.1 links). Um diese Bedingungen zu erfüllen, gibt es für (β_0, β_1) unendlich viele Möglichkeiten, z.B.

$\beta_0 \in [1, 3]$ und $\beta_1 = 0$. Nach der Dreiecksungleichung gilt aber

$$\begin{aligned} & |1 - \beta_0 - \beta_1| + |3 - \beta_0 - \beta_1| + |1 - \beta_0 - 5\beta_1| + |3 - \beta_0 - 5\beta_1| \\ & \left(|\beta_0 + \beta_1 - 1| + |3 - \beta_0 - \beta_1| \right) + \left(|\beta_0 + 5\beta_1 - 1| + |3 - \beta_0 - 5\beta_1| \right) \\ \geq & \left| \beta_0 + \beta_1 - 1 + 3 - \beta_0 - \beta_1 \right| + \left| \beta_0 + 5\beta_1 - 1 + 3 - \beta_0 - 5\beta_1 \right| \\ = & 2 + 2 = 4. \end{aligned}$$

Also sind alle (β_0, β_1) , die obige Bedingungen erfüllen, optimale Lösungen. Es kann je nach Wahl der Punkte aber auch eine eindeutige Lösung geben, z.B. wenn alle Punkte auf einer Geraden liegen (siehe Abb. 3.1 rechts).

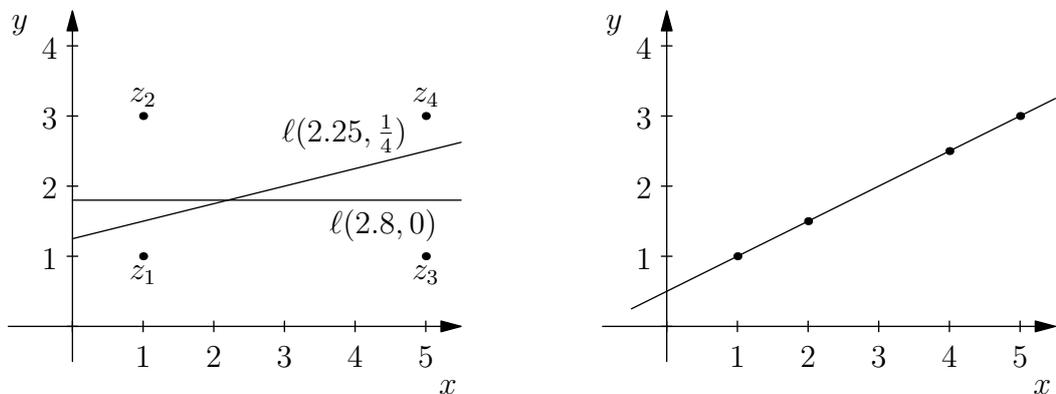


Abbildung 3.1: Zwei (TLAD)-Probleme mit $n = m = 4$. Links: unendlich viele Lösungen; rechts: eindeutige Lösung

3.2 Die geometrische Dualität

In diesem Abschnitt führen wir eine duale Transformation der Ebene ein, die die Herleitung einer endlichen Kandidatenmenge für das Problem (TLAD) im nächsten Abschnitt 3.3 erleichtert. Diese Transformation trägt der Beobachtung Rechnung, dass die Platzierung einer nicht-vertikalen Geraden durch die Bestimmung der beiden Parameter β_0 und β_1 für den Achsenabschnitt und die Steigung geschieht. Diese können als Wert jeder reellen Zahl annehmen und daher liegt der Gedanke nahe, das Problem so umzuschreiben, dass die Lösung als ein Punkt im \mathbb{R}^2 zu suchen ist und nicht mehr als Gerade. Um das zu erreichen, ist also eine Transformation $*$ nötig, die eine Gerade l in einen Punkt l^* abbildet und die gegebenen Punkte z_1, \dots, z_n in Objekte z_1^*, \dots, z_n^* umwandelt, für die

wieder ein vertikaler Abstand $d_v(l^*, z_i^*)$ definiert werden kann, der gleich dem ursprünglichen Abstand $d_v(z_i, l)$ ist. Dies leistet die Transformation in Definition 3.6 und Lemma 3.8 hält die wichtigsten Eigenschaften dieser Transformation, z. B. Bijektivität, fest, die nötig sind, um (TLAD) äquivalent zu transformieren.

Definition 3.6 (Duale Transformation, [20]). *Die duale Transformation $*$ ordnet jedem Punkt $z = (x, y) \in \mathbb{R}^2$ die nicht-vertikale Gerade*

$$z^* = \ell(-y, x) = \{(s, t) \in \mathbb{R}^2 : t = xs - y\}$$

und jeder nicht-vertikalen Gerade $l = \ell(\beta_0, \beta_1)$ den Punkt

$$l^* = (\beta_1, -\beta_0) \in \mathbb{R}^2$$

zu. Man nennt z^ eine duale Gerade und l^* einen dualen Punkt.*

Für die einfachere Beschreibung einiger Zusammenhänge ist außerdem die Notation der vertikalen Residuen zweckmäßig. Sie entsprechen den Residuen der linearen Regression in der Statistik.

Notation 3.7. *Für einen Punkt $z = (x, y) \in \mathbb{R}^2$ und eine nicht-vertikale Gerade $l = \ell(\beta_0, \beta_1)$ ist das vertikale Residuum gegeben durch*

$$r_v(z, l) = y - (\beta_0 + \beta_1 x).$$

Es gilt also $d_v(z, l) = |r_v(z, l)|$. Außerdem liegt z über l genau dann, wenn $r_v(z, l) > 0$, und z liegt unter l genau dann, wenn $r_v(z, l) < 0$. $r_v(z, l) = 0$ schließlich ist gleichbedeutend mit $d_v(z, l) = 0$ und $z \in l$.

Um mit der dualen Transformation arbeiten zu können, werden im folgenden Lemma wichtige Eigenschaften gezeigt.

Lemma 3.8 (Eigenschaften der dualen Transformation, [20]). *Sei $z = (x, y) \in \mathbb{R}^2$ ein beliebiger Punkt und $l = \ell(\beta_0, \beta_1)$ eine beliebige nicht-vertikale Gerade im \mathbb{R}^2 . Dann gelten*

1. *Die Transformation $*$ ist idempotent, also $z^{**} = z$ und $l^{**} = l$.*
2. *Die Transformation $*$ erhält die vertikalen Residuen zwischen Punkten und Geraden, also $r_v(z, l) = r_v(l^*, z^*)$.*
3. *Die Transformation $*$ erhält den vertikalen Abstand zwischen Punkten und Geraden, also $d_v(z, l) = d_v(l^*, z^*)$.*

4. Es gilt $z \in l$ genau dann, wenn $l^* \in z^*$.
5. z liegt über l genau dann, wenn l^* über z^* liegt. Die gleiche Aussage gilt für das „Darunterliegen“.
6. Die Transformation $*$ bildet die Punkte des \mathbb{R}^2 bijektiv auf die nicht-vertikalen dualen Geraden im \mathbb{R}^2 ab.

Beweis. ad 1. Nach der Definition der dualen Transformation (Def. 3.6) rechnet man

$$\begin{aligned} z^{**} &= (x, y)^{**} = \ell(-y, x)^* = (x, y) = z \\ l^{**} &= \ell(\beta_0, \beta_1)^{**} = (\beta_1, -\beta_0)^* = \ell(\beta_0, \beta_1) = l. \end{aligned}$$

ad 2. Es folgt aus Notation 3.7

$$r_v(z, l) = y - (\beta_1 x + \beta_0) = -\beta_0 - (x\beta_1 - y) = r_v((\beta_1, -\beta_0), \ell(-y, x)) = r_v(l^*, z^*).$$

ad 3. Folgt aus Punkt 2 mit $d_v(z, l) = |r_v(z, l)| = |r_v(l^*, z^*)| = d_v(l^*, z^*)$.

ad 4. Aus $z \in l$ folgt $y = \beta_1 x + \beta_0$, also $d_v(z, l) = 0$. Da unter Punkt 3 gezeigt wurde, dass $*$ die Abstände erhält, gilt auch $d_v(l^*, z^*) = 0$ mit $l^* = (\beta_1, -\beta_0)$ und $z^* = \ell(-y, x)$, also $-\beta_0 = -y\beta_1 + x$ und daher $l^* \in z^*$. Die Rückrichtung folgt aus der Idempotenz von $*$.

ad 5. Liege z über l , also $0 < r_v(z, l) = r_v(l^*, z^*)$. Damit liegt auch l^* über z^* . Die Idempotenz von $*$ ergibt die Rückrichtung und der Beweis für das „Darunterliegen“ geht analog.

ad 6. Seien $z_1 = (x_1, y_1) \neq z_2 = (x_2, y_2)$ zwei Punkte, d.h. $x_1 \neq x_2$ oder $y_1 \neq y_2$. Dann sind ihre dualen Geraden $z_1^* = \ell(-y_1, x_1)$ bzw. $z_2^* = \ell(-y_2, x_2)$ verschieden, da sie entweder verschiedene Achsenabschnitte oder verschiedene Steigungen haben. Also ist $*$ injektiv. Sei nun eine Gerade $l = \ell(\beta_0, \beta_1)$ vorgegeben. Dann gilt mit $z = l^* = (\beta_1, -\beta_0)$ nach der Idempotenz aus Punkt 1 $z^* = l^{**} = l$ und l hat den Punkt z als Urbild. Damit ist $*$ auch surjektiv und zusammen bijektiv. \square

Die Bijektivität der dualen Transformation $*$ ermöglicht es nun, für nicht-vertikale Geraden z^* zu schreiben und stillschweigend anzunehmen, dass ihr ein Punkt z zugrunde liegt. Ebenso kann man für einen Punkt l^* schreiben mit einer nicht-vertikalen Geraden l .

Besonders interessant ist die zweite Eigenschaft aus Lemma 3.8. Sie erlaubt, die Zielfunktion von (TLAD) aus (3.4) folgendermaßen umzuschreiben:

$$\sum_{i=1}^m d_v(\ell(\beta_0, \beta_1), z_{(i)}) = \sum_{i=1}^m d_v(\ell(\beta_0, \beta_1)^*, z_{(i)}^*) = \sum_{i=1}^m d_v((\beta_1, -\beta_0), \ell(-y_{(i)}, x_{(i)})).$$

Also ist das Programm

$$\begin{aligned} \text{(dual-TLAD)} \quad & \min_{l^*=(s,t) \in \mathbb{R}^2} \sum_{i=1}^m d_v(l^*, z_{(i)}^*) \\ & = \min_{l^*=(s,t) \in \mathbb{R}^2} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} d_v(l^*, z_i^*) \end{aligned} \quad (3.8)$$

mit einem $m : 0 < m \leq n$ und einer von $l^* \in \mathbb{R}^2$ abhängigen Permutation (\cdot) mit

$$d_v(l^*, z_{(1)}^*) \leq d_v(l^*, z_{(2)}^*) \leq \dots \leq d_v(l^*, z_{(n)}^*)$$

äquivalent zu (TLAD), da die sechste Eigenschaft aus Lemma 3.8 die eindeutige Zuordnung der Lösungen von (TLAD) und (dual-TLAD) ermöglicht.

Jetzt wollen wir die duale Transformation ausnutzen, um das Lösen des (TLAD) bzw. (dual-TLAD) zu erleichtern. Dafür werden Punkte im Dualraum im Folgenden mit $l^* = (s, t) \in \mathbb{R}^2$ oder in ähnlicher Weise bezeichnet, um Verwechslungen mit Punkten $z = (x, y) \in \mathbb{R}^2$ im primalen Raum zu vermeiden. Wir stellen zunächst anhand eines Beispiels fest, dass die gegebenen Punkte z_1, \dots, z_n durch ihre Transformation zu Geraden im Dualraum diesen in mehrere Gebiete zerlegen, ihn also parkettieren. Seien dazu Punkte

$$\begin{aligned} z_1 &= \left(-\frac{3}{4}, -1\right) & z_2 &= \left(-\frac{1}{2}, \frac{1}{4}\right) & z_3 &= \left(\frac{1}{2}, -\frac{1}{2}\right) \\ z_4 &= \left(\frac{1}{2}, 0\right) & z_5 &= (2, -1) \end{aligned} \quad (3.9)$$

gegeben. Dann sind ihre dualen Geraden

$$\begin{aligned} z_1^* &= \ell\left(1, -\frac{3}{4}\right) & z_2^* &= \ell\left(-\frac{1}{4}, -\frac{1}{2}\right) & z_3^* &= \ell\left(\frac{1}{2}, \frac{1}{2}\right) \\ z_4^* &= \ell\left(0, \frac{1}{2}\right) & z_5^* &= \ell(1, 2). \end{aligned} \quad (3.10)$$

Abbildung 3.2 zeigt oben die Punkte z_1, \dots, z_5 und unten die Parkettierung, die durch ihre dualen Geraden beschrieben wird.

Um diese Parkettierung zu beschreiben, seien z_1, \dots, z_n o.B.d.A. paarweise verschieden, um die Schreibweise $P = \{z_1, \dots, z_n\}$ zu ermöglichen.

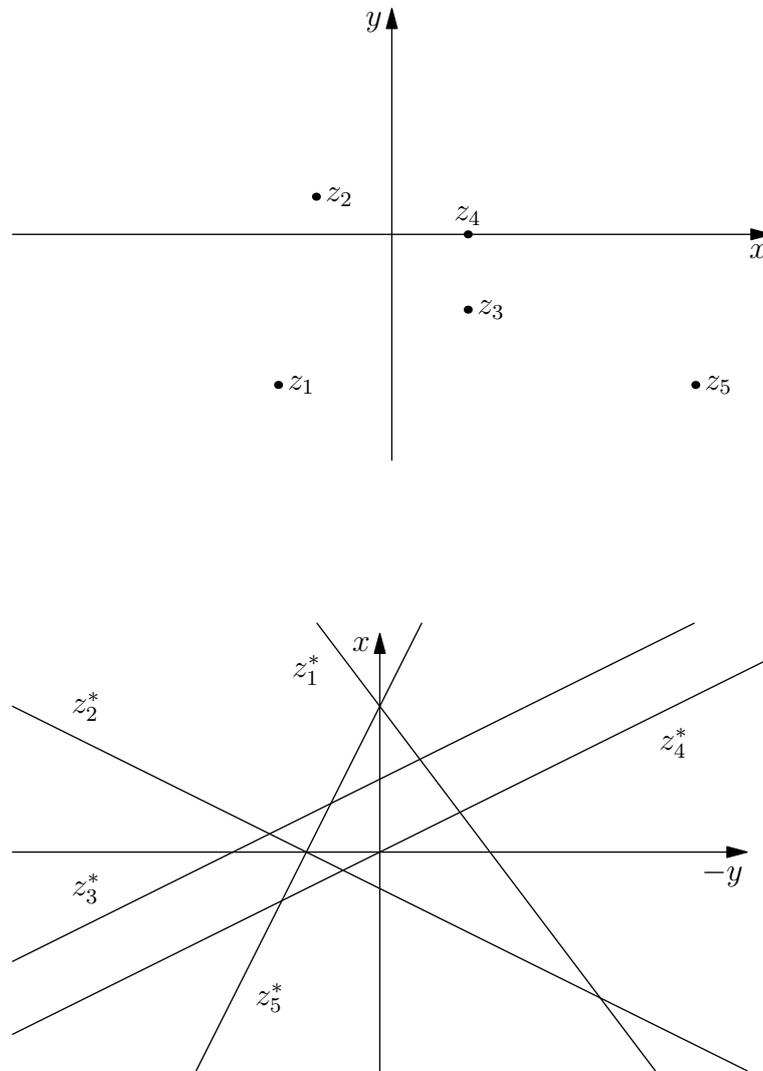


Abbildung 3.2: oben: die primalen Punkte aus (3.9) ;
 unten: die zugehörigen dualen Geraden aus (3.10)

Definition 3.9 (Duale Zelle). Zu einer gegebenen Menge an Punkten $P = \{z_1, \dots, z_n\} \subset \mathbb{R}^2$ und festen Vorzeichen $\sigma \in \{-1, 1\}^n$ ist eine duale Zelle gegeben durch

$$C(\sigma) = \{l^* \in \mathbb{R}^2 : \sigma_i \cdot r_v(l^*, z_i^*) \geq 0 \quad \forall i = 1, \dots, n\},$$

falls diese Menge mehr als ein Element enthält.

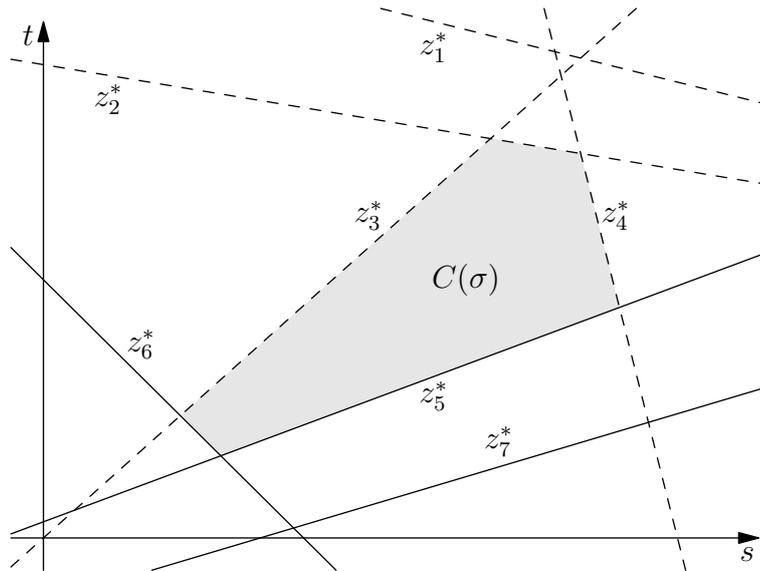


Abbildung 3.3: Beispiel für eine duale Zelle $C(\sigma)$ nach Definition 3.9

Abbildung 3.3 zeigt eine duale Zelle $C(\sigma)$ mit $P = \{z_1, \dots, z_7\}$, $\sigma_1 = \dots = \sigma_4 = -1$ und $\sigma_5 = \dots = \sigma_7 = 1$. Mit dieser Abbildung ist die anschauliche Bedeutung der σ_i leicht zu erklären, denn es gilt offenbar für alle $i = 1, \dots, n$

$$\begin{aligned} \sigma_i = 1 & \iff r_v(l^*, z_i^*) \geq 0 \quad \forall l^* \in C(\sigma) \\ \sigma_i = -1 & \iff r_v(l^*, z_i^*) \leq 0 \quad \forall l^* \in C(\sigma) \end{aligned} \quad (3.11)$$

und $\sigma_i = 1$ gilt genau dann, wenn der duale Punkt l^* über der dualen Geraden z_i^* liegt (in Abb. 3.3 die durchgezogenen Linien) bzw. wenn der Punkt z_i über der Geraden l liegt (siehe Lemma 3.8). Analog gilt $\sigma_i = -1$ genau dann, wenn der duale Punkt l^* unter der dualen Geraden z_i^* liegt (in Abb. 3.3 die gestrichelten Linien) bzw. wenn der Punkt z_i unter der Geraden l liegt. Hiermit kann man insbesondere den vertikalen Abstand anders schreiben:

$$d_v(z_i, l) = d_v(l^*, z_i^*) = |r_v(l^*, z_i^*)| = \sigma_i r_v(l^*, z_i^*) = \sigma_i r_v(z_i, l) \quad \forall i = 1, \dots, n \quad (3.12)$$

Die dualen Zellen sind offenbar Polyeder, denn sie sind ein Schnitt von endlich vielen Halbräumen. Mit

$$\sigma_i r_v(l^*, z_i^*) \geq 0 \iff \sigma_i(t - x_i s + y_i) \geq 0 \iff \sigma_i x_i s - \sigma_i t \leq \sigma_i y_i$$

kann man $C(\sigma)$ in der Form

$$C(\sigma) = \{l^* = (s, t) \in \mathbb{R}^2 : \begin{pmatrix} \sigma_1 x_1 & -\sigma_1 \\ \vdots & \vdots \\ \sigma_n x_n & -\sigma_n \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} \leq \begin{pmatrix} \sigma_1 y_1 \\ \vdots \\ \sigma_n y_n \end{pmatrix}\} \quad (3.13)$$

schreiben. Diese Polyeder können Polytope oder unbeschränkt sein. Sie haben allerdings immer ein nichtleeres Inneres. Denn anderenfalls wäre eine solche Zelle die leere Menge, ein einzelner Punkt oder ein möglicherweise unbeschränkter Geradenabschnitt. Die ersten beiden Fälle sind durch $|C(\sigma)| > 1$ ausgeschlossen, der dritte Fall tritt nicht ein, da in P jeder Punkt nur einmal vorkommen kann. Als Polyeder sind die Zellen insbesondere konvex und abgeschlossen.

Weiter hat jede Zelle $C(\sigma)$ mindestens eine Ecke, wenn P mindestens zwei Punkte mit verschiedenen x -Koordinaten enthält, denn dann haben ihre dualen Geraden unterschiedliche Steigung. Daher wird jede Gerade $z_{i_0}^*$, die $C(\sigma)$ begrenzt, also

$$\sigma_{i_0} r_v(l^*, z_{i_0}^*) \geq 0 \quad \forall l^* \in C(\sigma) \quad \text{und} \quad \sigma_{i_0} r_v(l_0^*, z_{i_0}^*) = 0 \quad \text{für ein } l_0^* \in C(\sigma),$$

von einer der Geraden z_i^* geschnitten. Bezeichne den Schnittpunkt mit l_1^* . Falls $l_1^* \in C$ ist, ist l_1^* die gesuchte Ecke von $C(\sigma)$. Sonst muss es eine weitere Gerade $z_{i_1}^*$ mit $z_{i_1} \in P$ geben, die l_1^* von $C(\sigma)$ trennt, also

$$\sigma_{i_1} r_v(l_1^*, z_{i_1}^*) < 0 \quad \text{und} \quad \sigma_{i_1} r_v(l^*, z_{i_1}^*) \geq 0 \quad \forall l^* \in C(\sigma). \quad (3.14)$$

Es können $z_{i_0}^*$ und $z_{i_1}^*$ nicht parallel sein, weil $l_0^* \in z_{i_0}^*$ und $l_1^* \in z_{i_0}^*$. Das bedeutet aber

$$r_v(l_0^*, z'^*) = r_v(l_1^*, z'^*) \quad \forall z'^* \text{ mit } z'^* \parallel z_{i_0}^*$$

und ist ein Widerspruch zu (3.14), da $l_0^* \in C(\sigma)$. Also haben $z_{i_0}^*$ und $z_{i_1}^*$ einen Schnittpunkt l_2^* , der in $C(\sigma)$ liegen kann oder nicht. Liegt er in $C(\sigma)$, ist er die gesuchte Ecke, wenn nicht, kann er wie eben von $C(\sigma)$ getrennt werden. Dies kann man aber nicht unendlich lange fortführen, da $P^* = \{z_1^*, \dots, z_n^*\}$ endlich ist. Also muss einer der so konstruierten Schnittpunkte eine Ecke von $C(\sigma)$ sein.

Definition 3.10 (Duale Parkettierung [20]). *Zu einer gegebenen Menge an Punkten $P = \{z_1, \dots, z_n\} \subset \mathbb{R}^2$ ist die duale Parkettierung $DT(P)$ (dual tessellation) gegeben durch die Menge aller dualen Zellen $C(\sigma)$ wie in Definition 3.9:*

$$DT(P) = \{C(\sigma) : \sigma \in \{0, 1\}^n, |C(\sigma)| > 1\}$$

Es gilt offenbar

$$\bigcup_{C \in DT(P)} C = \mathbb{R}^2, \quad (3.15)$$

weil für jeden Punkt $(s, t) \in \mathbb{R}^2$ und jedes $i \in \{1, \dots, n\}$ immer mindestens eine der Ungleichungen

$$r_v((s, t), z_i^*) \leq 0 \quad , \quad r_v((s, t), z_i^*) \geq 0$$

erfüllt ist. Daher kann σ_i für jedes i so gewählt werden, dass $\sigma_i \cdot r_v((s, t), z_i^*) \geq 0$. Also ist

$$(s, t) \in C(\sigma) \in DT(P).$$

3.3 Eine endliche Kandidatenmenge für (TLAD)

Nun können wir mit Hilfe der dualen Parkettierung die Linearitätsbereiche der Zielfunktion von (LAD) als duale Zellen identifizieren und dann den Hauptsatz der linearen Optimierung anwenden, um für jede dieser Zellen eine Ecke anzugeben, in der die Zielfunktion ihr Minimum auf dieser Zelle annimmt, so dass nur noch die Ecken aller dualen Zellen nach einem Optimum für (dual-LAD) abgesucht werden müssen. Da jede Ecke ein Schnittpunkt von zwei oder mehr dualen Geraden ist, erhält man durch Rücktransformation in den ursprünglichen Raum eine optimale Gerade, die mindestens durch zwei der gegebenen Punkte z_1, \dots, z_n geht. Diese Eigenschaft heißt *weak incidence property* und gilt auch für die Platzierung einer Hyperebene im \mathbb{R}^n nach Schöbel [30].

Lemma 3.11 ([20]). *Die Zielfunktion*

$$\sum_{i=1}^n d_v(\cdot, z_i^*) : \mathbb{R}^2 \rightarrow \mathbb{R}$$

des (dual-LAD) ist affin linear auf jeder Zelle $C(\sigma) \in DT(P)$, wobei $P = \{z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)\}$.

Beweis. Auf jeder dualen Zelle $C(\sigma) \in DT(P)$ gilt für alle $l^* \in C(\sigma)$

$$\begin{aligned} \sum_{i=1}^n d_v(l^*, z_i^*) &= \sum_{i=1}^n \sigma_i r_v(l^*, z_i^*) = \sum_{i=1}^n \sigma_i (t - x_i s + y_i) \\ &= t \sum_{i=1}^n \sigma_i - s \sum_{i=1}^n \sigma_i x_i + \sum_{i=1}^n \sigma_i y_i. \end{aligned} \quad (3.16)$$

Das ist offenbar eine in $l^* = (s, t)$ affin lineare Funktion, weil die x_i und y_i fest sind und σ auf jeder dualen Zelle $C(\sigma)$ fest ist. \square

Lemma 3.12. *Die Zielfunktion*

$$f := \sum_{i=1}^n d_v(\cdot, z_i^*) : \mathbb{R}^2 \rightarrow \mathbb{R}$$

des (dual-LAD) ist konvex.

Beweis. Seien $l_1^* = (s_1, t_1), l_2^* = (s_2, t_2) \in \mathbb{R}^2$ und $\lambda \in [0, 1]$. Dann gilt

$$\begin{aligned} f(\lambda l_1^* + (1 - \lambda)l_2^*) &= \sum_{i=1}^n |(\lambda t_1 + (1 - \lambda)t_2) - x_i(\lambda s_1 + (1 - \lambda)s_2) + y_i| \\ &= \sum_{i=1}^n |\lambda t_1 - \lambda x_i s_1 + \lambda y_i + (1 - \lambda)t_2 - (1 - \lambda)x_i s_2 + (1 - \lambda)y_i| \\ &\stackrel{(1)}{\leq} \sum_{i=1}^n \left[|\lambda(t_1 - x_i s_2 + y_i)| + |(1 - \lambda)(t_2 - x_i s_2 + y_i)| \right] \\ &\stackrel{(2)}{=} \lambda \sum_{i=1}^n |t_1 - x_i s_2 + y_i| + (1 - \lambda) \sum_{i=1}^n |t_2 - x_i s_2 + y_i| \\ &= \lambda f(l_1^*) + (1 - \lambda)f(l_2^*) \end{aligned}$$

nach der Dreiecksungleichung (1) und weil $\lambda \geq 0$ und $1 - \lambda \geq 0$ (2). □

Damit reicht es, nach lokalen Minima zu suchen, da lokale Minima konvexer Funktionen auch deren globale Minima sind [4, S. 287]. Man kann zur Bestimmung eines globalen Minimums also Verfahren des steilsten Abstiegs oder andere Methoden zur Berechnung lokaler Minima anwenden, ohne sich mit einem nicht-globalen Minimum begnügen zu müssen. Leider ist die Zielfunktion von (dual-TLAD) im Allgemeinen nicht einmal quasikonvex, wie ein Beispiel zeigt. Gegeben seien die Punkte

$$\begin{array}{lll} z_1 = (1, 0) & z_2 = (3, 1) & z_3 = (5, 2) \\ z_4 = (2, 3) & z_5 = (4, 4) & z_6 = (6, 5) \end{array} \quad (3.17)$$

und $m = 3$. Dann sind die Geraden $l_1 = \ell(-1/2, 1/2)$ und $l_2 = \ell(2, 1/2)$ optimale Lösungen für (TLAD), da sie Zielfunktionswert Null haben. Die Gerade $l_3 = \ell(3/2, 1/2)$ allerdings ist keine optimale Lösung, denn sie hat den Zielfunktionswert $3/2$. Es ist aber l_3^* eine Konvexkombination von l_1^* und l_2^* und die Zielfunktion kann nicht quasikonvex sein, da die Menge der Minima quasikonvexer Funktionen konvex ist [31], siehe Abbildung 3.4.

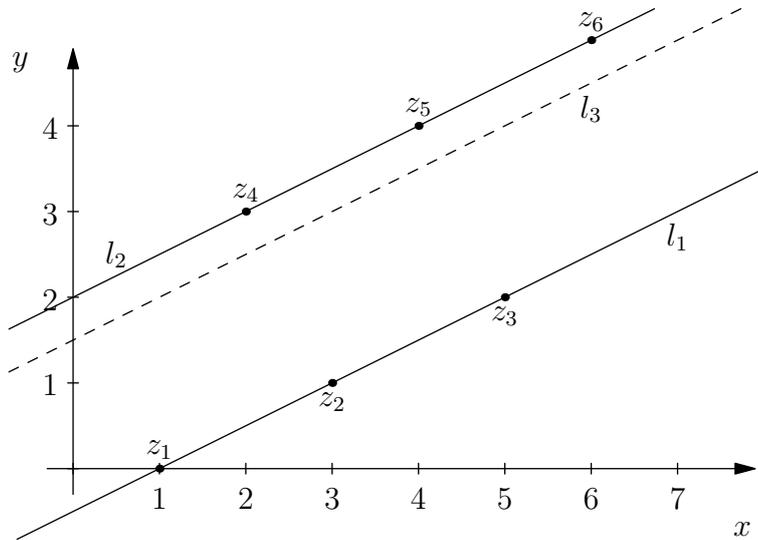


Abbildung 3.4: Die Zielfunktion von (TLAD) ist im Allgemeinen nicht quasikonvex.

Satz 3.13 (Starker Punkt). *Unter gegebenen Punkten $z_i = (x_i, y_i)$, $i = 1, \dots, n$ mit mindestens zwei verschiedenen x -Koordinaten gibt es einen Punkt z_j , so dass jede für (LAD) optimale Gerade durch diesen Punkt verläuft, falls eine der folgenden Bedingungen erfüllt ist:*

1. n ist eine ungerade Zahl.
2. n ist eine gerade Zahl und es gibt keine Partitionierung I_1, I_2 der Indizes $\{1, \dots, n\}$ mit $|I_1| = |I_2| = n/2$ und

$$\sum_{i \in I_1} x_i = \sum_{i \in I_2} x_i. \tag{3.18}$$

Beweis. Nach Lemma 3.5 hat (LAD) zu den gegebenen Punkten eine endliche Optimallösung. Sei $P = \{z \in \mathbb{R}^2 : z = z_i \text{ für ein } i \in \{1, \dots, n\}\}$. Wir verwenden das zu (LAD) äquivalente Programm (dual-LAD). Der Beweis gliedert sich in drei Schritte:

1. Es gibt keine optimale Lösung, die im Inneren einer dualen Zelle $C(\sigma)$ liegt.
2. Zu zwei optimalen Lösungen gibt es einen Punkt z_j , durch den beide verlaufen.
3. z_j ist der gesuchte Punkt.

Schritt 1: Zuerst zeigen wir, dass unter den obigen Bedingungen kein Punkt, der im Inneren einer dualen Zelle $C(\sigma) \in DT(P)$ liegt, eine optimale Lösung für (dual-LAD) sein kann. Denn dort ist die Zielfunktion

$$\sum_{i=1}^n d_v(l^*, z_i^*) = t \sum_{i=1}^n \sigma_i - s \sum_{i=1}^n \sigma_i x_i + \sum_{i=1}^n \sigma_i y_i$$

mit $l^* = (s, t)$ nach Lemma 3.11 linear und daher stetig differenzierbar. Ihr Gradient

$$\left(- \sum_{i=1}^n \sigma_i x_i, \sum_{i=1}^n \sigma_i \right)$$

müsste bei einem Minimum also verschwinden. Wegen der ersten Bedingung ist offenbar

$$\sum_{i=1}^n \sigma_i = 0$$

nicht möglich, da $\sigma \in \{-1, 1\}^n$ und der Gradient kann nicht verschwinden. Damit gibt es kein Minimum im Inneren von $C(\sigma)$.

Nun nehmen wir

$$\left(- \sum_{i=1}^n \sigma_i x_i, \sum_{i=1}^n \sigma_i \right) = (0, 0)$$

an und führen das zum Widerspruch zur zweiten Bedingung. Es folgt nämlich

$$\sum_{i=1}^n \sigma_i = 0 \implies \begin{cases} \sigma_i = +1 & \text{für } \frac{n}{2} \text{ Indizes} \\ \sigma_i = -1 & \text{für } \frac{n}{2} \text{ Indizes} \end{cases}.$$

Seien I_+ und I_- die entsprechenden Indextmengen. Für diese gilt

$$- \sum_{i \in I_+} \sigma_i x_i - \sum_{i \in I_-} \sigma_i x_i = 0 \implies \sum_{i \in I_+} x_i = \sum_{i \in I_+} \sigma_i x_i = - \sum_{i \in I_-} \sigma_i x_i = \sum_{i \in I_-} x_i.$$

Da aber $|I_+| = |I_-| = n/2$, ist dies ein Widerspruch zur zweiten Bedingung.

Damit liegt jede optimale Lösung von (dual-LAD) auf dem Rand einer dualen Zelle. D. h. es gibt für jede optimale Lösung $l^{*\text{opt}}$ von (dual-LAD) eine duale Gerade $z_{i(l^{*\text{opt}})}^*$, so dass

$$d_v(l^{*\text{opt}}, z_{i(l^{*\text{opt}})}^*) = 0.$$

Da die duale Transformation * nach Lemma 3.8 die vertikalen Abstände erhält, verläuft die für (LAD) optimale Gerade l^{opt} durch den Punkt $z_{i(l^{*\text{opt}})}$. Falls die optimale Lösung eindeutig ist, ist der Beweis hier abgeschlossen. Sonst führen wir Schritt 2 und 3 durch.

Schritt 2: Seien $l_0^{*\text{opt}}$ bzw. $l_1^{*\text{opt}}$ zwei verschiedene optimale Lösungen von (dual-LAD). Wir zeigen, dass diese durch einen gemeinsamen Punkt z_j verlaufen. Die beiden Lösungen müssen nach Schritt 1 jeweils auf dem Rand einer dualen Zelle liegen, also gilt

$$d_v(l_0^{*\text{opt}}, z_{i_0}^*) = 0 \quad \text{bzw.} \quad d_v(l_1^{*\text{opt}}, z_{i_1}^*) = 0$$

mit geeigneten Punkten z_{i_0} und z_{i_1} . Falls

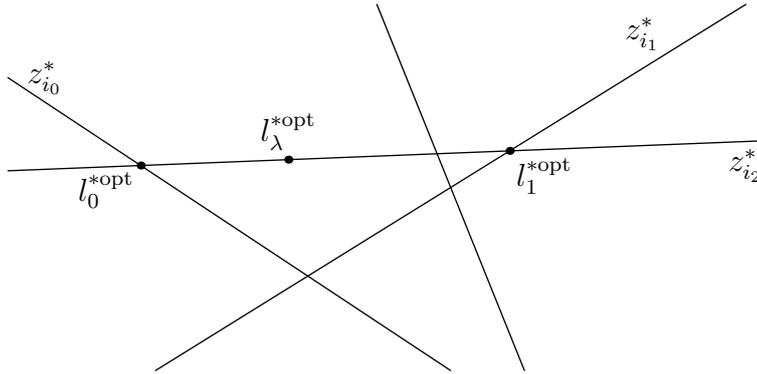


Abbildung 3.5: Erläuterung zu Schritt 2 des Beweises von Satz 3.13

$$d_v(l_0^{*\text{opt}}, z_{i_1}^*) = 0 \quad \text{oder} \quad d_v(l_1^{*\text{opt}}, z_{i_0}^*) = 0,$$

sind wir fertig, denn l_0^{opt} und l_1^{opt} verlaufen beide durch z_{i_0} oder z_{i_1} . Sonst ist auch

$$l_\lambda^{*\text{opt}} = \lambda l_1^{*\text{opt}} + (1 - \lambda) l_0^{*\text{opt}}$$

eine optimale Lösung von (dual-LAD) für jedes $\lambda \in [0, 1]$ nach Lemma 3.12, da die Menge der Minima konvexer Funktionen konvex ist [4, S. 287]. Angenommen, es gilt

$$\min_{i=1, \dots, n} d_v(l_\lambda^{*\text{opt}}, z_i^*) > 0$$

für ein $\lambda \in [0, 1]$. Dann liegt $l_\lambda^{*\text{opt}}$ im Inneren einer Zelle $C(\sigma)$ im Widerspruch zu Schritt 1. Also ist

$$\min_{i=1, \dots, n} d_v(l_\lambda^{*\text{opt}}, z_i^*) = 0$$

für jedes $\lambda \in [0, 1]$. Dann gibt es einen festen Index i_2 mit $d_v(l_\lambda^{*\text{opt}}, z_{i_2}^*) = 0$ für alle $\lambda \in [0, 1]$, denn die Menge $\{l_\lambda^{*\text{opt}} : \lambda \in [0, 1]\}$ ist eine Strecke im Dualraum. Das bedeutet aber, $d_v(l_0^{*\text{opt}}, z_{i_2}^*) = d_v(l_1^{*\text{opt}}, z_{i_2}^*) = 0$ und l_0^{opt} und l_1^{opt} verlaufen durch z_{i_2} .

Insgesamt verlaufen l_0^{opt} und l_1^{opt} beide durch einen der Punkte

$$z_j := \begin{cases} z_{i_0} \\ z_{i_1} \\ z_{i_2} \end{cases} .$$

Schritt 3: Sei z_j wie im vorigen Schritt und l^{opt} eine optimale Lösung von (LAD). Angenommen, es ist $d_v(z_j, l^{\text{opt}}) = d_v(l^{\text{opt}}, z_j^*) > 0$, das heißt, l^{opt} verläuft nicht durch z_j . Dann hat die konvexe Hülle $B := \text{conv}(l_0^{\text{opt}}, l_1^{\text{opt}}, l^{\text{opt}})$ ein nicht-leeres Inneres, da die dualen Punkte l_0^{opt} und l_1^{opt} auf der dualen Geraden z_j^* liegen, l^{opt} aber nicht, und die Punkte daher ein Dreieck beschreiben (siehe Abb. 3.6).

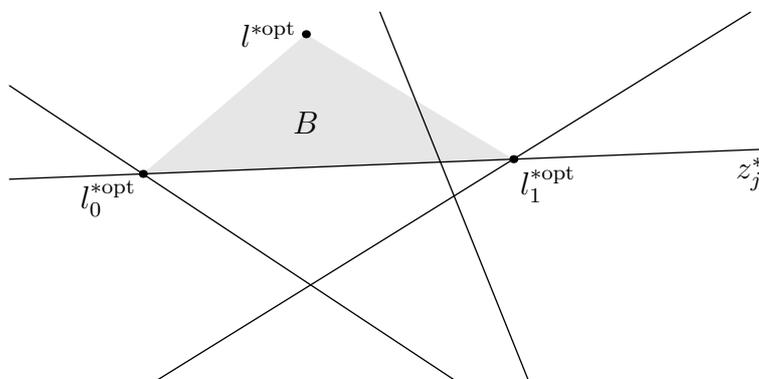


Abbildung 3.6: Widerspruch zu Schritt 1

Weil die Zielfunktion von (dual-LAD) konvex ist und die Menge der Minima konvexer Funktionen konvex ist, sind alle $l^* \in B$ optimale Lösungen für (dual-LAD). Da aber in Schritt 1 gezeigt wurde, dass jedes Minimum auf dem Rand einer dualen Zelle liegen muss, hat die Menge der Minima ein leeres Inneres. Das ist ein Widerspruch und es muss $d_v(z_j, l^{\text{opt}}) = d_v(l^{\text{opt}}, z_j^*) = 0$ gelten.

Insgesamt verlaufen also alle optimalen Geraden durch den Punkt z_j . Ein solcher Punkt z_j sei als *starker Punkt* bezeichnet. □

Ob die erste Bedingung des Satzes erfüllt ist, sieht man natürlich sofort an den Eingabedaten des LAD-Problems. Die zweite Bedingung ist jedoch nur schwer nachzuprüfen, da die Bestimmung einer solchen Partition NP-vollständig ist. Das sieht man durch polynomielle Reduzierung des folgenden Partitionierungs-Problems auf das Problem in (3.18):

Partitionierungs-Problem. Finde zu gegebenen Zahlen $c_1, \dots, c_n \in \mathbb{Z}$ eine Indexmenge $I \subset \{1, \dots, n\}$, so dass

$$\sum_{i \in I} c_i = \sum_{i \notin I} c_i.$$

Dieses Problem ist nach Karp [14] NP-vollständig.

Lemma 3.14. Das Problem, für gerades n und zu gegebenen Zahlen $x_1, \dots, x_n \in \mathbb{R}$ Indextmengen $I_1, I_2 \subset \{1, \dots, n\}$ zu finden mit $|I_1| = |I_2| = n/2$, $I_1 \cap I_2 = \emptyset$ und

$$\sum_{i \in I_1} x_i = \sum_{i \in I_2} x_i,$$

ist NP-vollständig bezüglich der Eingabegröße n .

Beweis. Angenommen, man könnte das Problem doch in polynomialer Zeit $p(n)$ lösen. Dann könnte man das gleiche Problem für die Zahlen $x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}$ auch in polynomialer Zeit lösen, da auch $p(2n)$ ein Polynom in n ist. Wähle dann $x_1, \dots, x_n \in \mathbb{Z}$ beliebig und $x_{n+1} = \dots = x_{2n} = 0$. Seien I_1, I_2 die Indextmengen, die das Problem für die $2n$, mit Nullen ergänzten Zahlen lösen. Setze dann $I = \{i \in I_1 : i \leq n\}$. So hat man wegen

$$\sum_{i \in I} x_i = \sum_{\substack{i \in I_1 \\ i \leq n}} x_i = \sum_{i \in I_1} x_i = \sum_{i \in I_2} x_i = \sum_{\substack{i \in I_2 \\ i \leq n}} x_i = \sum_{i \notin I} x_i$$

eine Lösung des Partitionierungs-Problems gefunden. Dies ist in polynomialer Zeit $p(2n)$ geschehen, was ein Widerspruch zur NP-Vollständigkeit dieses Problems ist. \square

Das nächste Lemma liefert die Aussage, dass es eine für (LAD) optimale Gerade gibt, die durch mindestens zwei der gegebenen Punkte verläuft. Damit lässt sich die Geradenplatzierung durch eine endliche Kandidatenmenge auf ein kombinatorisches Optimierungsproblem reduzieren [30]. Diese Eigenschaft wird in 3.16 und Satz 4.7 noch auf die Geradenplatzierung mit Ausreißern und die Platzierung mehrerer Geraden mit Ausreißern verallgemeinert.

Lemma 3.15 (weak incidence property für (LAD), [30]). Seien $z_1, \dots, z_n \in \mathbb{R}^2$ gegeben, so dass es darunter mindestens zwei Punkte mit verschiedenen x -Koordinaten gibt. Dann enthält die Menge

$$L = \{l = \ell(\beta_0, \beta_1) : d_v(l, z_{i_1}) = d_v(l, z_{i_2}) = 0, i_1 \neq i_2, i_1, i_2 \in \{1, \dots, n\}\}$$

eine optimale Lösung von (LAD). Das heißt, es gibt eine optimale Gerade für (LAD), die durch mindestens zwei der gegebenen Punkte verläuft.

Beweis. Wir betrachten dazu das zu (LAD) äquivalente Programm

$$(\text{dual-LAD}) \quad \min_{(s,t) \in \mathbb{R}^2} \sum_{i=1}^n d_v((s,t), z_i^*).$$

Dann ist die Aussage des Lemmas äquivalent dazu, dass

$$L = \{l = \ell(\beta_0, \beta_1) : d_v(l^*, z_{i_1}^*) = d_v(l^*, z_{i_2}^*) = 0, i_1 \neq i_2, i_1, i_2 \in \{1, \dots, n\}\}$$

eine optimale Lösung von (dual-LAD) enthält. Definiere nun die Menge

$$P = \{z \in \mathbb{R}^2 : z = z_i \text{ für ein } i \in \{1, \dots, n\}\}.$$

Nach Lemma 3.11 ist dann für jedes $C \in DT(P)$ das Programm

$$(\text{dual-LAD}(C)) \quad \min_{l^* \in \mathbb{R}^2} \sum_{i=1}^n d_v(l^*, z_i^*)$$

s.d. $l^* \in C$

ein lineares Programm, da C nach (3.13) ein Polyeder ist und die Zielfunktion affin linear nach Lemma 3.11. Weiterhin hat jedes dieser Optimierungsprobleme eine endliche Optimallösung, weil die Zielfunktion für $\|l^*\|_2 \rightarrow \infty$ unendlich groß wird (analog zum Beweis von Lemma 3.5), aber ein Minimierungsproblem vorliegt. Außerdem hat jede Zelle C eine Ecke und nach dem Hauptsatz der linearen Optimierung [29, S. 76] wird das Minimum dann auch in einer Ecke l_C^* von C angenommen. Eine solche optimale Ecke von (dual-LAD(C)) ist aber der Schnittpunkt von zwei dualen Geraden $z_{i_1}^*, z_{i_2}^* \in P, i_1 \neq i_2$, also $d_v(l_C^*, z_{i_1}^*) = d_v(l_C^*, z_{i_2}^*) = 0$ und $l_C^* \in L^*$.

Weiterhin haben wir in Gleichung (3.15) gesehen, dass

$$\bigcup_{C \in DT(P)} C = \mathbb{R}^2.$$

Damit ist eine Optimallösung von (dual-LAD) gegeben durch

$$l_C^{\text{opt}} = \underset{\substack{l_C^* \text{ ist optimale} \\ \text{Ecke von (dual-LAD}(C))}}{\arg \min} \sum_{i=1}^n d_v(l_C^*, z_i^*) \in L^*,$$

da $|DT(P)| < \infty$. Also ist $l_C^{\text{opt}} \in L$ eine Optimallösung von (LAD). □

Diese Eigenschaft gilt nicht nur für die Geradenplatzierung, die der Entfernungsmessung die vertikalen Abstände zugrunde legt, sondern nach Schöbel [30] auch für allgemeine norminduzierte Abstände und in höheren Dimensionen. Formal heißt das, dass es für das Problem, eine Hyperebene $H \subset \mathbb{R}^p$ so zu platzieren, dass die Summe der Abstände

$$\sum_{i=1}^n d(z_i, H) = \sum_{i=1}^n \min_{z \in H} \|z_i - z\| \quad (3.19)$$

zu gegebenen Punkten $z_1, \dots, z_n \in \mathbb{R}^p$ bei Vorgabe einer beliebigen Norm $\|\cdot\|$ minimiert wird, immer eine optimale Hyperebene gibt, die durch eine p -elementige, affin unabhängige Teilmenge der gegebenen Punkte verläuft. Diese Eigenschaft gilt auch, wenn der Abstand ein gerichteter Abstand der Form

$$d_t(z, H) = \min_{\lambda} \{|\lambda| : z + \lambda t \in H\} \quad (3.20)$$

mit einem $t \in \mathbb{R}^p$ ist [30]. Der hier behandelte vertikale Abstand ist ein Spezialfall dieser Abstände mit $p = 2$ und $t = (0, 1)^t$.

Weiterhin erfüllen alle optimalen Hyperebenen, die bezüglich der hier genannten Entfernungen platziert werden, die pseudo-halving property:

$$\sum_{i: z_i \in H^+} 1 \leq \frac{n}{2} \quad \text{und} \quad \sum_{i: z_i \in H^-} 1 \leq \frac{n}{2},$$

wobei H^+ der Teil des \mathbb{R}^p ist, der über H liegt, H^- der darunter liegende. Falls zusätzlich jedem Punkt z_i ein Gewicht w_i zugeordnet ist, wird diese Eigenschaft zu

$$\sum_{i: z_i \in H^+} w_i \leq \frac{n}{2} \quad \text{und} \quad \sum_{i: z_i \in H^-} w_i \leq \frac{n}{2}.$$

Es ist anzumerken, dass die üblicherweise in der Regression verwendeten quadratischen Residuen in keine der hier angegebenen Klassen von Entfernungsmaßen fallen und tatsächlich gilt dort weder die weak incidence noch die pseudo-halving Eigenschaft.

Die weak incidence property lässt sich auf das TLAD-Problem verallgemeinern, indem man zeigt, dass eine optimale Lösung $(l^{\text{opt}}, I^{\text{opt}})$ eines (TLAD) zu Punkten z_1, \dots, z_n einer optimalen Lösung l^{opt} des (LAD) zu den Punkten z_i , $i \in I^{\text{opt}}$ entspricht. Daher gilt sie nicht nur für die absoluten Abweichungen, sondern auch für die oben angegebenen Abstände, denn der Beweis verwendet im Wesentlichen Lemma 3.15, das ja auch für Abstände wie in (3.19) und (3.20) gilt.

Satz 3.16 (weak incidence property für (TLAD)). *Seien $z_1, \dots, z_n \in \mathbb{R}^2$ gegeben, so dass darunter keine m Punkte mit gleicher x -Koordinate sind für ein gegebenes m mit $2 \leq m \leq n$. Dann enthält die Menge*

$$L = \{l = \ell(\beta_0, \beta_1) : d_v(l, z_{i_1}) = d_v(l, z_{i_2}) = 0, i_1 \neq i_2, i_1, i_2 \in \{1, \dots, n\}\}$$

eine optimale Lösung von (TLAD), das heißt, es gibt eine optimale Gerade für (TLAD), die durch mindestens zwei der gegebenen Punkte verläuft.

Beweis. Nach Lemma 3.5 gibt es eine optimale Lösung $(l^{\text{opt}}, I^{\text{opt}})$ von (TLAD). Damit muss l^{opt} eine optimale Lösung des Programms

$$(\text{LAD}(I^{\text{opt}})) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{i \in I^{\text{opt}}} d_v(z_i, \ell(\beta_0, \beta_1))$$

sein, denn wäre l' eine optimale Lösung von $(\text{LAD}(I^{\text{opt}}))$ mit

$$\sum_{i \in I^{\text{opt}}} d_v(z_i, l') < \sum_{i \in I^{\text{opt}}} d_v(z_i, l^{\text{opt}}),$$

dann wäre (l', I^{opt}) eine bessere Lösung für (TLAD) im Widerspruch zur Optimalität von $(l^{\text{opt}}, I^{\text{opt}})$. Aber $(\text{LAD}(I^{\text{opt}}))$ ist ein (LAD)-Programm für die Punkte $z_i, i \in I^{\text{opt}}$. Von diesen Punkten haben mindestens zwei eine unterschiedliche x -Koordinate, denn sonst wären alle $x_i, i \in I^{\text{opt}}$ gleich und es gäbe im Widerspruch zur Voraussetzung m Punkte mit gleicher x -Koordinate unter den z_1, \dots, z_n . Nach Lemma 3.15 enthält L eine optimale Lösung l_0^{opt} von $(\text{LAD}(I^{\text{opt}}))$. Für diese Lösung gilt

$$\sum_{i \in I^{\text{opt}}} d_v(z_i, l_0^{\text{opt}}) = \sum_{i \in I^{\text{opt}}} d_v(z_i, l^{\text{opt}})$$

und damit ist $(l_0^{\text{opt}}, I^{\text{opt}})$ auch eine optimale Lösung von (TLAD). □

Mit dieser Aussage kann eine Lösung von (TLAD) bestimmt werden, indem aus den vorgegebenen Punkten z_1, \dots, z_n alle $\binom{n}{2} = O(n^2)$ Paare von Punkten bestimmt werden und die Zielfunktionswerte der dadurch bestimmten Geraden ausgewertet werden. Für das Auswerten des Zielfunktionswerts für jedes Paar von Punkten ist erstens die Berechnung der Abstände aller Punkte von der zugehörigen Gerade in einer Zeit von $O(n)$ nötig und zweitens die Bestimmung der m kleinsten Abstände durch eine Sortierung der n Abstände. Die Sortierung benötigt eine Zeit von $O(n \log n)$ [10]. Die Auswertung der Zielfunktion geschieht also $O(n^2)$ mal in einer Zeit von $O(n \log n)$, so dass sich insgesamt eine Zeitkomplexität von $O(n^3 \log n)$ zur Bestimmung aller möglichen Zielfunktionswerte

ergibt. Die Bestimmung des kleinsten Wertes in $O(n^2)$ lässt die Zeitkomplexität unverändert.

Außerdem kann man aus Satz 3.16 ein kompaktes Gebiet $Q \subset \mathbb{R}^2$ bestimmen, in dem die Geradenparameter (β_0, β_1) liegen müssen, wenn man nur an einer optimalen Lösung interessiert ist und nicht unbedingt an allen. Zuerst schränken wir die Steigung β_1 auf ein Kompaktum Q_1 ein. Da es eine optimale Lösung l gibt, die durch zwei der gegebenen Punkte, z_i und z_j , verläuft, ist die Steigung dieser Lösung $\beta_1 = \frac{y_j - y_i}{x_j - x_i}$. Daher könnte man β_1 sinnvoll einschränken auf

$$\beta_1 \in \left[\min_{i,j=1,\dots,n} \frac{y_j - y_i}{x_j - x_i}, \quad \max_{i,j=1,\dots,n} \frac{y_j - y_i}{x_j - x_i} \right] =: [\underline{q}_1, \bar{q}_1] =: Q_1. \quad (3.21)$$

Die Berechnung dieser Schranken benötigt offenbar eine Zeit von $O(n^2)$, da Minimum und Maximum über zwei Indizes von 1 bis n gebildet werden müssen.

Nun muss noch β_0 auf ein Kompaktum Q_0 beschränkt werden. Wir nutzen wieder aus, dass es eine optimale Gerade $l = \ell(\beta_0, \beta_1)$ gibt, die durch einen Punkt z_i verläuft, also $y_i = \beta_0 + \beta_1 x_i$. Bei bekannter Steigung β_1 dieser Geraden ergibt sich der Achsenabschnitt zu $\beta_0 = y_i - \beta_1 x_i$. Wir hatten aber durch Q_1 die Steigung schon beschränkt. Also ist der kleinstmögliche Achsenabschnitt

$$\begin{aligned} \min_{\substack{i=1,\dots,n \\ \beta_1 \in Q_1}} y_i - \beta_1 x_i &\geq \min_{i=1,\dots,n} y_i - \max_{\beta_1 \in Q_1} \beta_1 x_i \\ &= y_{\min} - \max_{q=\underline{q}_1, \bar{q}_1} \{q \cdot x_{\min}, q \cdot x_{\max}\} =: \underline{q}_0. \end{aligned} \quad (3.22)$$

In analoger Weise lässt sich auch \bar{q}_0 bestimmen. Diese Schranken können offenbar in linearer Zeit bestimmt werden, so dass man nun die Geradenparameter in einer Zeit von $O(n^2)$ auf ein Kompaktum $Q = Q_0 \times Q_1$ beschränken kann.

4 Platzierung mehrerer Geraden

4.1 Problemstellung

Eine natürliche Erweiterung der Fragestellung aus dem vorigen Kapitel ist die Platzierung mehrerer Geraden zur Beschreibung der gegebenen Punkte. Das Problem besteht hier also darin, die Punkte z_1, \dots, z_n durch $K \in \mathbb{N}$ Geraden l_1, \dots, l_K anzunähern, wobei wie schon bei der Platzierung einer Geraden nur die m Punkte berücksichtigt werden sollen, die den optimalen Geraden am nächsten sind. Es soll hierbei ein Punkt z_i als gut angenähert bezeichnet werden, wenn er von einer der Geraden l_k einen kleinen Abstand hat, also $d_v(z_i, l_k)$ klein ist. Die Abstände dieses Punktes von den anderen Geraden sind dann nicht mehr von Belang, der Punkt wird also der Geraden fest zugeordnet, die ihn am besten approximiert. Man kann daher mit

$$\min_{k=1, \dots, K} d_v(z_i, l_k)$$

den Abstand des Punktes z_i von den Geraden l_1, \dots, l_K bezeichnen. Die Geraden sind dann so zu bestimmen, dass die Summe dieser Abstände für alle m relevanten Punkte minimal wird. Als mathematisches Programm erhält man

$$\min_{\substack{l_1, \dots, l_K \\ \text{Geraden im } \mathbb{R}^2}} \sum_{i=1}^m \min_{k=1, \dots, K} d_v(z_i, l_k) \quad (4.1)$$

mit einer Permutation (\cdot) , die analog zu (3.1) zu wählen ist und von den Geraden l_1, \dots, l_K abhängt:

$$\min_{k=1, \dots, K} d_v(z_{(1)}, l_k) \leq \dots \leq \min_{k=1, \dots, K} d_v(z_{(n)}, l_k) \quad (4.2)$$

Das lässt sich wie in (3.3) auch wieder kombinatorisch formulieren:

$$\min_{\substack{l_1, \dots, l_K \\ \text{Geraden im } \mathbb{R}^2}} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} \min_{k=1, \dots, K} d_v(z_i, l_k) \quad (4.3)$$

Es ist klar, dass die Punkte immer besser beschrieben werden können, je mehr Geraden verwendet werden. Im Extremfall kann man $K = \lceil m/2 \rceil$ Geraden platzieren und eine

Gerade durch je zwei Punkte legen (Für ungerade m ist die letzte Gerade dann nicht eindeutig bestimmt, da sie nur durch einen Punkt verlaufen muss). Für diese Geraden wäre der Zielfunktionswert gleich Null und damit wären sie optimal.

Dieses Programm erhält also als Eingangsdaten Punkte z_1, \dots, z_n , eine Anzahl m mit $0 < m \leq n$ von Punkten, die dem Modell folgen, und eine Anzahl K von Geraden, die zur Modellierung verwendet werden sollen. K sollte kleiner als $\lceil m/2 \rceil$ sein, damit das Programm nicht trivial wird. Eine Lösung des Programms besteht aus Geraden l_1, \dots, l_k und Indextmengen I_1, \dots, I_K , wobei $i \in I_k$ nur, wenn

$$d_v(z_i, l_k) = \min_{k'=1, \dots, K} d_v(z_i, l_{k'}).$$

Sollte das Minimum für mehrere Indizes angenommen werden, wird i nur einer der Indextmengen zugeordnet, so dass diese disjunkt sind. Diese Zuordnung kann willkürlich geschehen. Außerdem sollen die Indextmengen keine Ausreißer enthalten, also nur Indizes i mit $(i) \leq m$ in (4.2) werden aufgenommen, d. h. $|I_1 \cup \dots \cup I_K| = m$ und alle Punkte, die in keiner dieser Mengen enthalten sind, werden als Ausreißer betrachtet. Bei Mehrdeutigkeit kann auch hier die Auswahl willkürlich geschehen. Eine solche Lösung notieren wir als $(l_1, I_1, \dots, l_K, I_K)$. Hier muss man sich klarmachen, dass die Indextmengen keine freien Parameter sind, wie es auf den ersten Blick scheinen könnte, denn die Wahl der Geraden legt die Indextmengen schon fest, zumindest bis auf die oben erwähnten Uneindeutigkeiten.

Als Erstes soll das Problem wieder auf nicht-vertikale Geraden eingeschränkt werden, um für die lineare Regression mit latenten Klassen von Nutzen zu sein. Dann wird das Problem aus (4.1) zu

$$(K\text{-TLAD}) \quad \min_{(\beta_{01}, \beta_{11}), \dots, (\beta_{0K}, \beta_{1K}) \in \mathbb{R}^2} \sum_{i=1}^m \min_{k=1, \dots, K} d_v(z_{(i)}, \ell(\beta_{0k}, \beta_{1k})) \quad (4.4)$$

oder zu

$$(K\text{-TLAD}) \quad \min_{(\beta_{01}, \beta_{11}), \dots, (\beta_{0K}, \beta_{1K}) \in \mathbb{R}^2} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} \min_{k=1, \dots, K} d_v(z_i, \ell(\beta_{0k}, \beta_{1k})). \quad (4.5)$$

4.2 NP-Vollständigkeit der Geradenplatzierung

Bevor weitere Eigenschaften des K -TLAD-Problems untersucht werden, soll zunächst gezeigt werden, dass eine algorithmische Lösung sehr aufwendig ist. Dazu wird die NP-Vollständigkeit von (K -TLAD) nachgewiesen, indem es auf eine leicht abgewandelte Form

des point covering Problems aus Megiddo und Tamir [23] reduziert wird. Unter dem Gesichtspunkt der NP-Vollständigkeit des Problems erscheint auch die Komplexität des in Abschnitt 4.4 angegebenen Lösungsverfahrens in einem anderen Licht: Trotz der langen Laufzeit des Verfahrens wirft die NP-Vollständigkeit des Problems die berechnete Frage auf, ob es überhaupt deutlich schnellere Algorithmen gibt – unter der Annahme $\mathcal{P} \neq \mathcal{NP}$.

Wir beginnen mit der Angabe des ursprünglichen point covering Problems und einer direkten Folgerung für das Problem aus 4.1.

Das point covering Problem (PC). Seien Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ in der Ebene gegeben und eine Zahl K . Entscheide, ob es K Geraden gibt, so dass diese die gegebenen Punkte überdecken. Das heißt, jeder der Punkte muss mindestens auf einer der Geraden liegen.

Satz 4.1 (NP-Vollständigkeit des point covering Problems, [23]). *Das point covering Problem (PC) ist NP-vollständig bezüglich der Eingabegröße K .*

Die NP-Vollständigkeit wurde in [23] sogar für rationale Punkte x_i, y_i gezeigt. Damit ist das Problem (4.1) bzw. (4.3) NP-vollständig, denn wenn es in polynomieller Zeit lösbar wäre, könnte man in dieser Zeit ermitteln, ob dessen Zielfunktionswert verschwindet oder größer als Null ist. Im ersten Fall hat man K Geraden gefunden, so dass $\min_{k=1, \dots, K} d_v(z_i, l_k) = 0$ für alle $i = 1, \dots, n$, also liegt jeder Punkt auf mindestens einer der Geraden. Im zweiten Fall hingegen muss es mindestens einen Punkt z_i geben, für den $\min_{k=1, \dots, K} d_v(z_i, l_k) > 0$ gilt. Damit ist das Entscheidungsproblem gelöst.

Allerdings ist (K -TLAD) ein Spezialfall von (4.1), so dass dessen NP-Vollständigkeit noch nicht direkt gefolgert werden kann. Dazu benötigen wir zuerst eine Folgerung aus Satz 4.1, die das point covering Problem einschränkt, aber die NP-Vollständigkeit erhält.

Das eingeschränkte point covering Problem (PC'). Seien Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ in der Ebene gegeben und eine Zahl K . Entscheide, ob es K nicht-vertikale Geraden gibt, so dass diese die gegebenen Punkte überdecken. Das heißt, jeder der Punkte muss mindestens auf einer der Geraden liegen.

Korollar 4.2 (NP-Vollständigkeit des eingeschränkten point covering Problems). *Das eingeschränkte point covering Problem (PC') ist NP-vollständig bezüglich der Eingabegröße K .*

Beweis. Wir nehmen an, dass auf nicht-vertikale Geraden beschränkte point covering

Problem wäre in polynomieller Zeit $p(K)$ lösbar, und erzeugen einen Widerspruch durch polynomielle Reduktion wie in Lemma 3.14.

Zunächst wählt man $K + 1$ paarweise verschiedene Winkel $\varphi'_1, \dots, \varphi'_{K+1} \in [0, \pi]$. Dann sei T_φ eine Drehung der Ebene um φ . Löse dann die $K + 1$ point covering Probleme mit Einschränkung auf nicht-vertikale Geraden für jeweils die Punkte

$$T_{-\varphi'_k}(z_1), \dots, T_{-\varphi'_k}(z_n)$$

in jeweils polynomieller Zeit $p(K)$, also insgesamt in polynomieller Zeit $(K + 1)p(K)$. Jetzt können zwei Fälle auftreten.

Fall 1. Bei mindestens einem der transformierten Probleme – ohne Einschränkung bei dem mit der Transformation $T_{-\varphi'_1}$ – gibt es K Geraden l_1, \dots, l_K , so dass jeder der transformierten Punkte auf mindestens einer dieser Geraden liegt. Dann liegt auch jeder der zurücktransformierten Punkte $z_i = T_{\varphi'_1}(T_{-\varphi'_1}(z_i))$, $i = 1, \dots, n$ auf mindestens einer der Geraden $T_{\varphi'_1}(l_1), \dots, T_{\varphi'_1}(l_K)$, da Drehungen den euklidischen Abstand erhalten. Das ursprüngliche point covering Problem ist positiv beantwortet.

Fall 2. Bei keinem der transformierten Probleme gibt es K Geraden, so dass jeder der transformierten Punkte auf mindestens einer der Geraden liegt. Angenommen, das ursprüngliche point covering Problem hätte aber K solche überdeckenden Geraden l_1, \dots, l_K . Dann bildet jede dieser Geraden genau einen Winkel mit der y -Achse im mathematisch positiven Sinne gemessen. Also bilden die K Geraden höchstens K verschiedene solche Winkel $\varphi_1, \dots, \varphi_K$, da sie auch parallel sein können. Die Winkel liegen alle in $[0, \pi]$. Jetzt gilt aber $\varphi'_{k_0} \neq \varphi_k$ für alle $k = 1, \dots, K$ und einen festen Index $k_0 \in \{1, \dots, K + 1\}$. Das hieße aber, dass alle Geraden $T_{-\varphi_{k_0}}(l_1), \dots, T_{-\varphi_{k_0}}(l_K)$ nicht-vertikal sind und die Punkte $T_{-\varphi_{k_0}}(z_1), \dots, T_{-\varphi_{k_0}}(z_n)$ überdecken. Das ist ein Widerspruch und das ursprüngliche point covering Problem ist negativ beantwortet.

Da die Transformation der Punkte in konstanter Zeit bezüglich K möglich ist, könnte man so das point covering Problem in polynomieller Zeit lösen, wenn man das point covering Problem mit Einschränkung auf nicht-vertikale Geraden in polynomieller Zeit lösen könnte. Das ist ein Widerspruch zur NP-Vollständigkeit des point covering Problems nach Satz 4.1. \square

Bemerkung 4.3. Der Beweis funktioniert auch, wenn man nicht die vertikalen Geraden ausschließt, sondern zum Beispiel die horizontalen. Dann sind lediglich die Winkel $\varphi_1, \dots, \varphi_K$ anders. Dieselbe Beweisidee führt ebenso zum Ziel, wenn man alle Geraden, die einen bestimmten Winkel mit der x -Achse bilden, verbietet.

Man kann sogar eine feste Anzahl W von Winkeln $\tilde{\varphi}_1, \dots, \tilde{\varphi}_W$ vorgeben und verlangen,

dass Geraden, die diesen Winkel mit der x -Achse bilden, nicht zum Überdecken verwendet werden dürfen. Dann muss man nicht $K + 1$ transformierte Probleme lösen, sondern $WK + 1$. Wenn man auch $W = 0$ als leere Bedingung an die Steigungswinkel der Geraden zulässt, kann man das point covering Problem sehr allgemein formulieren.

Das allgemeine point covering Problem (PC’). Seien Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ in der Ebene gegeben und eine Zahl K . Entscheide, ob es K Geraden gibt, so dass diese die gegebenen Punkte überdecken. Das heißt, jeder der Punkte muss mindestens auf einer der Geraden liegen. Dabei darf keine der Geraden einen Winkel $\varphi_1, \dots, \varphi_W$ mit der x -Achse bilden (gemessen im mathematisch positiven Sinne).

Damit hat man wie oben unmittelbar die NP-Vollständigkeit einer ganzen Klasse von Geradenplatzierungsproblemen.

Satz 4.4 (NP-Vollständigkeit der Geradenplatzierung). *Seien Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ in der Ebene und zwei Zahlen K und m gegeben. Außerdem seien W Winkel $\tilde{\varphi}_1, \dots, \tilde{\varphi}_W$ gegeben und d eine Abstandsfunktion, die den Abstand $d(z, l)$ von einem Punkt z zu einer Geraden l angibt. d erfülle*

1. $d(z, l) \geq 0$
2. $d(z, l) = 0 \iff z \in l$

für alle Punkte z und Geraden l in der Ebene. Dann ist das Problem, K Geraden l_1, \dots, l_K zu finden, die die Zielfunktion

$$\min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} \min_{k=1, \dots, K} d(z_i, l_k) \tag{4.6}$$

minimieren, sodass keine der Geraden einen Winkel $\tilde{\varphi}_1, \dots, \tilde{\varphi}_W$ mit der x -Achse bildet, NP-vollständig in Bezug auf die Eingabegröße K . Dies gilt auch, wenn $W = 0$, also kein Winkel verboten ist.

Beweis. Wir betrachten nur den Spezialfall $m = n$ und daraus folgt die NP-Vollständigkeit des allgemeineren Problems für beliebiges m . Angenommen, das Geradenplatzierungsproblem wäre nicht NP-vollständig. Dann kann man in polynomieller Zeit den optimalen Zielfunktionswert für (4.6) ermitteln. Dieser kann verschwinden oder größer als Null sein wegen der ersten Bedingung an die Abstandsfunktion d .

Fall 1. Angenommen, die Zielfunktion verschwindet. Dann liegen wegen der zweiten Bedingung an d alle gegebenen Punkte auf mindestens einer der optimalen Geraden.

Würden keine Einschränkungen an die Steigungswinkel der Geraden gestellt, wäre damit (PC) positiv beantwortet. Würden die Steigungswinkel eingeschränkt, wurde (PC'') positiv beantwortet.

Fall 2. Wenn der Zielfunktionswert echt größer als Null ist, dann muss mindestens ein Punkt z_j unter den gegebenen existieren, der $\min_{k=1,\dots,K} d(z_j, l_k) > 0$ für die optimalen Geraden l_1, \dots, l_K erfüllt. Dieser liegt dann auf keiner der Geraden wegen der zweiten Bedingung an d . Damit ist (PC) bzw. (PC'') negativ beantwortet.

Insgesamt könnte also (PC) bzw. (PC'') in polynomieller Zeit gelöst werden im Widerspruch zu deren NP-Vollständigkeit nach Satz 4.1 bzw. Korollar 4.2 und Bemerkung 4.3. \square

Der vertikale Abstand erfüllt natürlich die beiden Bedingungen von Satz 4.4 (siehe dazu auch Lemma 3.8) und im Fall von (K -TLAD) ist $W = 1$ und der verbotene Winkel $\varphi_1 = \pi/2$. Somit ist (K -TLAD) NP-Vollständig. Es ist allerdings zu bemerken, dass auch die Verwendung der LS-Methode mit den quadrierten Residuen als Abstandsfunktion die Voraussetzungen des Satzes erfüllt und die latent-class Regression unter ihrer Verwendung ebenfalls NP-vollständig ist. Damit verliert die LS-Methode bei mehreren Geraden einen ihrer größten Vorzüge, den sie bei der Platzierung einer Geraden gegenüber der LAD-Methode aufweist: ihre Effizienz.

4.3 Spezielle Lösungen und Existenzaussagen

Nachdem wir gezeigt haben, dass aufgrund der NP-Vollständigkeit der Geradenplatzierung die Lösung von (K -TLAD) sehr rechenintensiv sein wird, suchen wir in diesem Abschnitt nach Bedingungen an die gegebenen Punkte $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$, die sicherstellen, dass man guten Gewissens nach nicht-vertikalen Geraden zu deren Approximation suchen kann und so statt des Problems aus (4.1) nach Lösungen von (K -TLAD) suchen kann. Außerdem sind wir an Existenzaussagen zu dem Problem interessiert.

Lemma 4.5. *Seien Punkte $z_1, \dots, z_n \in \mathbb{R}^2$, $m : 2K \leq m \leq n$ und $K \in \mathbb{N}$ vorgegeben mit $z_i = (x_i, y_i)$ für $i = 1, \dots, n$. Seien x_1, \dots, x_n paarweise verschieden. Dann lässt sich aus jeder Lösung $(l_1, I_1, \dots, l_K, I_K)$ des Problems in (4.1) eine Lösung $(l'_1, I'_1, \dots, l'_K, I'_K)$ mit gleichem oder besserem Zielfunktionswert konstruieren, wobei die Geraden $l'_k = \ell(\beta'_{0k}, \beta'_{1k})$ nicht-vertikal sind und jeder Geraden mindestens zwei Punkte zugeordnet sind, also $|I'_k| \geq 2$ für alle $k = 1, \dots, K$.*

Beweis. Wir nehmen an, dass es einen Index k_0 gibt, so dass l_{k_0} vertikal ist oder $I_{k_0} = \emptyset$ oder $I_{k_0} = \{i_0\}$ für einen Index i_0 , denn sonst ist nichts zu zeigen. Für jeden der drei Fälle geben wir eine Vorschrift an, die $l_{k_0} \mapsto l'_{k_0}$ bzw. $I_{k_0} \mapsto I'_{k_0}$ abbildet, so dass für den Index k_0 in der neuen Lösung die gewünschten Eigenschaften erzeugt werden, für die anderen Klassen $k \in \{1, \dots, K\} \setminus \{k_0\}$ die gewünschten Eigenschaften erhalten bleiben und der Zielfunktionswert sich nicht verschlechtert.

Falls $I_{k_0} = \{i_0\}$, gibt es k_1 mit $|I_{k_1}| \geq 3$, denn sonst wäre $|I_1 \cup \dots \cup I_K| < 2K$. Setze dann $I'_{k_0} := I_{k_0} \cup \{i_1\}$ mit einem $i_1 \in I_{k_1}$ und $I'_{k_1} := I_{k_1} \setminus \{i_1\}$, so dass $|I'_{k_0}| = 2$ und $|I'_{k_1}| \geq 2$, und wähle l'_{k_0} als die Gerade, die durch z_{i_0} und z_{i_1} verläuft. Sie ist nicht-vertikal, da $x_{i_0} \neq x_{i_1}$ nach Voraussetzung. Setze $l'_k := l_k$ für alle k außer k_0 und $I'_k := I_k$ für alle k außer k_0 und k_1 . Für die Zielfunktionswerte der beiden Lösungen gilt dann

$$\begin{aligned}
 \sum_{k=1}^K \sum_{i \in I_k} d_v(z_i, l_k) &= \sum_{\substack{k=1 \\ k \neq k_0 \\ k \neq k_1}}^K \sum_{i \in I_k} d_v(z_i, l_k) + d_v(z_{i_0}, l_{k_0}) + \sum_{\substack{i \in I_{k_1} \\ i \neq i_1}} d_v(z_i, l_{k_1}) + d_v(z_{i_1}, l_{k_1}) \\
 &= \sum_{\substack{k=1 \\ k \neq k_0 \\ k \neq k_1}}^K \sum_{i \in I'_k} d_v(z_i, l'_k) + \sum_{i \in I'_{k_1}} d_v(z_i, l'_{k_1}) + \underbrace{d_v(z_{i_0}, l_{k_0})}_{\geq 0} + \underbrace{d_v(z_{i_1}, l_{k_1})}_{\geq 0} \\
 &\geq \sum_{\substack{k=1 \\ k \neq k_0}}^K \sum_{i \in I'_k} d_v(z_i, l'_k) + \underbrace{d_v(z_{i_0}, l'_{k_0})}_{=0} + \underbrace{d_v(z_{i_1}, l'_{k_1})}_{=0} \\
 &= \sum_{k=1}^K \sum_{i \in I'_k} d_v(z_i, l'_k).
 \end{aligned}$$

Falls $I_{k_0} = \emptyset$, gibt es k_1 mit $|I_{k_1}| \geq 4$ oder k_1 und k_2 mit $|I_{k_1}| \geq 3$ und $|I_{k_2}| \geq 3$, denn sonst wäre $|I_1 \cup \dots \cup I_K| < 2K$. Im ersten Fall setze $I'_{k_0} := \{i_1, i_2\}$ mit zwei verschiedenen Indizes $i_1, i_2 \in I_{k_1}$ und $I'_{k_1} := I_{k_1} \setminus \{i_1, i_2\}$. Im zweiten Fall setze $I'_{k_0} := \{i_1, i_2\}$ mit $i_1 \in I_{k_1}$ und $i_2 \in I_{k_2}$ sowie $I'_{k_1} := I_{k_1} \setminus \{i_1\}$ und $I'_{k_2} := I_{k_2} \setminus \{i_2\}$. Setze weiter l'_{k_0} als die Gerade, die durch z_{i_1} und z_{i_2} verläuft und damit nicht-vertikal ist. Außerdem $l'_k := l_k$ für alle k außer k_0 und $I'_k := I_k$ für alle k außer k_1 bzw. k_1 und k_2 . Wie im Fall $I_{k_0} = \{i_0\}$ überprüft man, dass die Lösung $(l'_1, I'_1, \dots, l'_K, I'_K)$ keinen schlechteren Zielfunktionswert hat.

Falls l_{k_0} vertikal ist, gilt $d_v(z_{i_0}, l_{k_0}) < \infty$ für höchstens ein $i_0 \in \{1, \dots, n\}$, da die x_i nach Voraussetzung paarweise verschieden sind. Damit ist $d_v(z_i, l_{k_0}) = \infty$ für alle anderen i . Verschiebe daher alle i mit $i \in I_{k_0}$ und $d_v(z_i, l_{k_0}) = \infty$ in eine beliebige Indexmenge I_{k_1} ,

$k_1 \neq k_0$, da in jedem Fall $d_v(z_i, l_{k_1}) \leq \infty$ gilt. Gehe dann zum ersten oder zweiten Fall.

Jetzt kann man diese drei Vorschriften so lange wiederholen, bis man eine Lösung mit nur nicht-vertikalen Geraden erhält, denen jeweils mindestens zwei Punkte zugeordnet sind. \square

Durch das Voraussetzen von $2K \leq m \leq n$ und paarweise verschiedenen x -Koordinaten kann also nicht ausgeschlossen werden, dass es eine vertikale Gerade gibt, die in einer optimalen Lösung für (4.1) enthalten ist, allerdings gibt es im Falle der Existenz einer Optimallösung auch immer eine, die nur nicht-vertikale Geraden enthält. Damit kann man sich auf das Problem (K -TLAD) aus (4.4) bzw. (4.5) zurückziehen, wenn man nur irgendeine optimale Lösung sucht und nicht alle. Wie schon auf Seite 30 für (TLAD) argumentiert, stellen diese Voraussetzungen keine besonders große Einschränkung für die Anwendung des K -TLAD-Problems im Bereich der Statistik dar. Bei wählbaren x -Koordinaten können diese paarweise verschieden gesetzt werden, bei zufällig bestimmten x -Koordinaten verschwindet die Wahrscheinlichkeit zweier gleicher x -Koordinate für stetige Verteilungen.

Nun müssen wir noch die Existenz einer optimalen Lösung sicherstellen (siehe Lemma 4.6). Die Frage der Nicht-Eindeutigkeit konnte ja schon für (TLAD) geklärt werden (Seite 33) und erübrigt sich somit für (K -TLAD), da es eine Verallgemeinerung von (TLAD) ist.

Lemma 4.6. *Für Punkte $z_1, \dots, z_n \in \mathbb{R}^2$ mit $z_i = (x_i, y_i)$ für $i = 1, \dots, n$ und paarweise verschiedenen x_1, \dots, x_n sowie m mit $2K \leq m \leq n$ hat (K -TLAD) eine optimale Lösung.*

Beweis. Wegen Lemma 4.5 können wir uns auf Lösungen $(l_1, I_1, \dots, l_K, I_K)$ mit $l_k = \ell(\beta_{0k}, \beta_{1k})$ und $|I_k| \geq 2$ für alle $k = 1, \dots, K$ beschränken. Sei nun J_k für jedes k eine feste Indexmenge mit mindestens zwei Elementen, die die Punkte z_j , $j \in J_k$ der Geraden l_k zuordnet. Die J_k seien paarweise disjunkt und es gelte $|J_1 \cup \dots \cup J_K| = m$. Dann gilt für jedes k und jedes J_k

$$\begin{aligned} \sum_{j \in J_k} d_v(z_j, l_k) &= \sum_{j \in J_k} |y_j - \beta_{0k} - \beta_{1k}x_j| \\ &\geq |y_{j_1} - \beta_{0k} - \beta_{1k}x_{j_1}| + |y_{j_2} - \beta_{0k} - \beta_{1k}x_{j_2}| \xrightarrow{\|(\beta_{0k}, \beta_{1k})\|_2 \rightarrow \infty} \infty \end{aligned}$$

für zwei verschiedene Indizes $j_1, j_2 \in J_k$. Aber aus

$$\|(\beta_{01}, \beta_{11}, \dots, \beta_{0K}, \beta_{1K})\|_2 \rightarrow \infty$$

folgt $\|(\beta_{0\bar{k}}, \beta_{1\bar{k}})\|_2 \rightarrow \infty$ für mindestens ein $\bar{k} \in \{1, \dots, K\}$. Daher gilt

$$\sum_{k=1}^K \sum_{j \in J_k} d_v(z_i, l_k) \geq \sum_{j \in J_{\bar{k}}} d_v(z_i, l_{\bar{k}}) \xrightarrow{\|(\beta_{01}, \beta_{11}, \dots, \beta_{0K}, \beta_{1K})\|_2 \rightarrow \infty} \infty$$

für jede Zuordnung J_1, \dots, J_K der Punkte zu den Geraden. Also kann man die Geradenparameter $(\beta_{01}, \beta_{11}, \dots, \beta_{0K}, \beta_{1K})$ auf ein Kompaktum $Q \subset \mathbb{R}^{2K}$ beschränken, da ein Minimierungsproblem vorliegt. Die Zielfunktion von (K -TLAD) ist als Komposition stetiger Funktionen (lineare Funktion, Betrag, Sortierung, Bildung eines Minimums, Summe) stetig und nimmt auf Q ihr Minimum an. \square

Wir haben schon gesehen, dass (LAD) eine konvexe Zielfunktion hat und (TLAD) nicht einmal eine quasikonvexe. Daher hat natürlich auch (K -TLAD) als Verallgemeinerung von (TLAD) keine quasikonvexe Zielfunktion. Man könnte nun vermuten, dass aber (K -LAD) als Verallgemeinerung von (LAD) noch eine konvexe oder zumindest quasikonvexe Zielfunktion besitzt. Das ist aber im Allgemeinen nicht der Fall, wie die Betrachtung des Beispiels aus Abbildung 3.4 zeigt. Sollen nämlich zwei Geraden ohne Trimmen zu den Punkten aus (3.17) platziert werden, sind die Paare (l_1, l_2) und (l_2, l_1) mit $l_1 = \ell(-1/2, 1/2)$, $l_2 = \ell(2, 1/2)$ und Zielfunktionswert Null optimal für das zugehörige (K -LAD). Aber (l_3, l_3) mit $l_3 = \ell(3/2, 1/2)$ ist natürlich keine optimale Lösung, obwohl es eine Konvexkombination der ersten beiden Lösungen ist. Damit kann die Zielfunktion nicht quasikonvex sein. Diese Beobachtung steht in engem Zusammenhang mit der Nicht-Identifizierbarkeit im statistischen Modell der latent-class Regression.

4.4 Eine endliche Kandidatenmenge für (K -TLAD)

Nun haben wir alle Mittel beisammen, um die weak incidence property, die wir schon für (TLAD) in Satz 4.7 hatten, auch auf die Platzierung mehrerer Geraden mit (K -TLAD) zu verallgemeinern. Zuerst zeigen wir, dass für eine optimale Lösung $(l_1^{\text{opt}}, I_1^{\text{opt}}, \dots, l_K^{\text{opt}}, I_K^{\text{opt}})$ von (K -TLAD) jede Gerade l_k^{opt} eine optimale Lösung von (LAD) für die Punkte z_i , $i \in I_k^{\text{opt}}$ ist, und wenden dann Lemma 3.15 an.

Satz 4.7 (weak incidence property für (K -TLAD)). *Seien Punkte $z_1, \dots, z_n \in \mathbb{R}^2$ gegeben sowie eine Anzahl $K \in \mathbb{N}$ von Geraden und eine Anzahl m mit $2K \leq m \leq n$. Dann enthält die Menge*

$$L = \{(l_1, \dots, l_K) : l_k \text{ verläuft durch mindestens zwei der Punkte } z_1, \dots, z_n \quad \forall k = 1, \dots, K\}$$

eine optimale Lösung von (K -TLAD), falls x_1, \dots, x_n paarweise verschieden sind.

Beweis. Nach Lemma 4.6 gibt es eine optimale Lösung $(l_1^{\text{opt}}, I_1^{\text{opt}}, \dots, l_K^{\text{opt}}, I_K^{\text{opt}})$ von (K -TLAD). Da die Voraussetzungen von Lemma 4.5 erfüllt sind, nehmen wir ohne Einschränkung an, dass alle Geraden $l_1^{\text{opt}}, \dots, l_K^{\text{opt}}$ nicht-vertikal sind und $|I_k^{\text{opt}}| \geq 2$ für alle $k = 1, \dots, K$. Damit ist l_k^{opt} eine optimale Lösung von

$$(\text{LAD}(I_k^{\text{opt}})) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{i \in I_k^{\text{opt}}} d_v(z_i, \ell(\beta_0, \beta_1))$$

für jedes $k = 1, \dots, K$.

Angenommen, das wäre für ein k' nicht so. Dann gibt es nach Lemma 3.5 eine optimale Lösung $l_{k'}^0$ von $(\text{LAD}(I_{k'}^{\text{opt}}))$ und für den Zielfunktionswert von (K -TLAD) gilt

$$\sum_{k=1}^K \sum_{i \in I_k^{\text{opt}}} d_v(z_i, l_k^{\text{opt}}) > \sum_{\substack{k=1 \\ k \neq k'}}^K \sum_{i \in I_k^{\text{opt}}} d_v(z_i, l_k^{\text{opt}}) + \sum_{i \in I_{k'}^{\text{opt}}} d_v(z_i, l_{k'}^0)$$

im Widerspruch zur Optimalität von $(l_1^{\text{opt}}, I_1^{\text{opt}}, \dots, l_K^{\text{opt}}, I_K^{\text{opt}})$. Jetzt gibt es nach Lemma 3.15 eine optimale Lösung l_k^L für $(\text{LAD}(I_k^{\text{opt}}))$, die durch mindestens zwei der Punkte z_i , $i \in I_k^{\text{opt}}$ verläuft. Damit ist $(l_1^L, I_1^{\text{opt}}, \dots, l_K^L, I_K^{\text{opt}})$ optimal für (K -TLAD) und (l_1^L, \dots, l_K^L) liegt in L . \square

Hieraus kann man man (K -TLAD) auf ein rein kombinatorisches Problem zurückführen, denn man kann es lösen, indem man für alle Kombinationen von K Geraden, die durch mindestens zwei der gegebenen Punkte verlaufen, die Zielfunktion auswertet. Nun zeigen wir, dass die Zeitkomplexität dieses Lösungsverfahrens $O(n^{2K+1} \log n)$ ist.

Um den Zielfunktionswert gegebener Geraden l_1, \dots, l_K auszuwerten, müssen zuerst die vertikalen Abstände

$$d_v(z_i, (l_1, \dots, l_k)) := \min_{k=1, \dots, K} d_v(z_i, l_k) \quad (4.7)$$

jedes der n Punkte von den Geraden bestimmt werden. Das geschieht für jeden Punkt in $O(K)$, also für alle Punkte in $O(nK) = O(n \log n)$, wenn man $K \leq \log n$ annimmt, was sicher gerechtfertigt ist, da es bei einer zu großen Zahl von Geraden zu einem Overfitting kommt. Dann müssen die m kleinsten der n Werte in (4.7) gefunden werden, zum Beispiel durch eine Sortierung in $O(n \log n)$ [10], was damit auch die Zeitkomplexität der gesamten Auswertung ist. Jetzt muss noch bestimmt werden, wie oft eine solche Auswertung geschehen muss. Um die erste Gerade festzulegen, stehen $\binom{n}{2}$ Paare von Punkten zur Verfügung. Für die zweite hat man nur noch $\binom{n-2}{2}$ Paare, da die Punkte, durch die

schon die erste Gerade verläuft, nicht mehr angenähert werden müssen. Damit ergibt sich eine Anzahl

$$\binom{n}{2} \cdot \binom{n-2}{2} \cdot \dots \cdot \binom{n-2K+2}{2} = O(n^{2K})$$

von Geradentupeln. Damit hat die Auswertung der Zielfunktion für alle diese Tupel eine Zeitkomplexität von $O(n^{2K+1} \log n)$. Die Bestimmung des Minimums dieser Werte ist in $O(n^{2K})$ möglich und ändert die gesamte Zeitkomplexität nicht mehr.

Bei vielen Geraden, also wenn K groß ist, wird diese Zeit allerdings enorm groß. Will man zum Beispiel 2000 Punkte durch vier Geraden annähern, benötigt man hierfür Zeiteinheiten in der Größenordnung von $2000^9 \cdot \log 2000 \approx 3,891 \cdot 10^{30}$. Wenn jede Zeiteinheit nur eine Nanosekunde lang ist, sind das ca. $1,2 \cdot 10^{14}$ Jahre. Man kann aber wie schon auf Seite 51 in einer Zeit von $O(n^2)$ ein Kompaktum $Q \subset \mathbb{R}^2$ bestimmen, auf das man die Parameter (β_{0k}, β_{1k}) jeder Gerade beschränken kann. Im obigen Beispiel sind das $4 \cdot 10^6$ Zeiteinheiten, also vier Millisekunden.

5 Algorithmen

In diesem Kapitel sollen Verfahren beschrieben werden, um die Probleme (LAD), (TLAD) und (K -TLAD) möglichst schnell und gut zu lösen. Mit Hilfe der weak incidence property (Satz 3.16 und Satz 4.7) haben wir schon eine endliche Kandidatenmenge hergeleitet, die nach einer optimalen Lösung abgesucht werden kann, was aber nur für kleine K in akzeptabler Zeit machbar ist. Andere Verfahren wie der FastTLE-Algorithmus von Neykov u. a. [24] sind zwar deutlich schneller, liefern allerdings nur lokal optimale Lösungen.

Die numerischen Ergebnisse, die in diesem Abschnitt präsentiert werden, wurden mit R [27] berechnet. Der Arbeit liegt ein Datenträger mit allen Quelltexten bei und an entsprechender Stelle wird auf Dateien dieses Datenträgers verwiesen.

5.1 Formulierung als lineares und gemischt-ganzzahliges lineares Programm

Bevor wir Algorithmen zur Lösung von (K -TLAD) angeben, formulieren wir das Problem zunächst als gemischt-ganzzahliges Programm und im Falle des (LAD) sogar als lineares Programm, um zum Beispiel die Möglichkeit zu haben, diese Probleme mit Hilfe eines Solvers wie Xpress zu lösen. Wir beginnen mit diesem einfachen Fall. Zur Erinnerung: Zu gegebenen Punkten $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n)$ war (LAD) gegeben als

$$(LAD) \quad \min_{(\beta_0, \beta_1) \in \mathbb{R}^2} \sum_{i=1}^n d_v(z_i, \ell(\beta_0, \beta_1)) \quad (5.1)$$

nach (3.4) bzw. mit ausgeschriebener Zielfunktion als

$$(LAD) \quad \min \sum_{i=1}^n |y_i - x_i \beta_1 - \beta_0|$$

s. d. $\beta_0, \beta_1 \in \mathbb{R}$

Um dies als lineares Programm zu formulieren, muss die Nicht-Linearität der Zielfunktion beseitigt werden. Dazu definiert man

$$d_i := |y_i - x_i \beta_1 - \beta_0| \quad \forall i = 1, \dots, n \quad (5.2)$$

und erhält

$$\begin{aligned}
 \text{(LAD)} \quad & \min \sum_{i=1}^n d_i \\
 \text{s. d.} \quad & d_i = |y_i - x_i\beta_1 - \beta_0| \quad \forall i = 1, \dots, n \\
 & \beta_0, \beta_1 \in \mathbb{R}, \quad d_i \geq 0 \quad \forall i = 1, \dots, n,
 \end{aligned}$$

wobei die neuen Nebenbedingungen sicherstellen, dass (5.2) erfüllt ist und $d_i \geq 0$ natürlich ohne Einschränkung gefordert werden kann. Nun ist die Zielfunktion linear, die Nebenbedingungen jedoch nicht. Mit Hilfe der Beobachtung $|x| = \max\{x, -x\}$ kann man

$$d_i \geq |y_i - x_i\beta_1 - \beta_0| \quad \iff \quad \begin{cases} d_i \geq y_i - x_i\beta_1 - \beta_0 \\ \text{und} \\ d_i \geq x_i\beta_1 + \beta_0 - y_i \end{cases}$$

statt (5.2) fordern. Da minimiert wird, wird damit im Falle einer optimalen Lösung $(d^{\text{opt}}, \beta_0^{\text{opt}}, \beta_1^{\text{opt}})$ immer

$$d_i^{\text{opt}} = |y_i - x_i\beta_1^{\text{opt}} - \beta_0^{\text{opt}}| \quad \forall i = 1, \dots, n$$

gelten. Damit erhält man die optimalen Lösungen von (LAD) durch das lineare Programm

$$\begin{aligned}
 \text{(LP-LAD)} \quad & \min \sum_{i=1}^n d_i \\
 \text{s. d.} \quad & -x_i\beta_1 - \beta_0 - d_i \leq -y_i \quad \forall i = 1, \dots, n \\
 & x_i\beta_1 + \beta_0 - d_i \leq y_i \quad \forall i = 1, \dots, n \\
 & \beta_0, \beta_1 \in \mathbb{R}, \quad d_i \geq 0 \quad \forall i = 1, \dots, n.
 \end{aligned}$$

(LAD) und (LP-LAD) sind nicht äquivalent, liefern aber dieselben optimalen Lösungen. Man kann den LAD-Schätzer also durch ein Simplex- oder Innere-Punkte-Verfahren schnell bestimmen. Das Simplex-Verfahren nach Dantzig [5] ist zwar im Mittel effizient, hat aber im schlimmsten Fall exponentielle Laufzeit [16]. Grundsätzlich ist die Lösung von (LP-LAD) in polynomialer Zeit möglich [15], [13] – polynomial in n und Länge der binären Kodierung des Problems. Die tatsächliche Laufzeit hängt jedoch von der Art der binären Kodierung ab und ist sogar bei festem n unbeschränkt, da die Koeffizienten x_i und y_i völlig unabhängig von n sind und daher in n unbeschränkt [21]. Damit führt die Ausnutzung der weak incidence property aus Satz 3.15 zum schnellsten der hier vorgestellten Lösungsverfahren für (LAD) mit einer Zeitkomplexität $O(n^3 \log n)$ (siehe Abschnitt 5.3). Die lineare Lösbarkeit linearer Programme nach Megiddo [22] ist hier nicht anwendbar, da die Voraussetzung der festen Dimension verletzt ist.

In ganz ähnlicher Weise kann auch (K -TLAD) umformuliert werden. Das Programm ist allerdings ganzzahlig und zunächst quadratisch. Für gegebene Punkte z_1, \dots, z_n und Parameter $m : 0 < m \leq n$ und $K \in \mathbb{N}$ war

$$(K\text{-TLAD}) \quad \min_{(\beta_{01}, \beta_{11}), \dots, (\beta_{0K}, \beta_{1K}) \in \mathbb{R}^2} \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} \min_{k=1, \dots, K} d_v(z_i, \ell(\beta_{0k}, \beta_{1k}))$$

nach (4.5) bzw.

$$(K\text{-TLAD}) \quad \min_{\substack{I \subset \{1, \dots, n\} \\ |I|=m}} \sum_{i \in I} \min_{k=1, \dots, K} |y_i - \beta_{1k}x_i - \beta_{0k}|$$

s. d. $\beta_{0k}, \beta_{1k} \in \mathbb{R} \quad \forall k = 1, \dots, K.$

Man braucht nun Entscheidungsvariablen $\pi_{ik} \in \{0, 1\}$, die angeben, ob ein Punkt z_i durch die Gerade $\ell(\beta_{0k}, \beta_{1k})$ am besten angenähert wird, also

$$\pi_{ik} = 1 \quad \implies \quad k = \arg \min_{k'=1, \dots, K} |y_i - \beta_{1k'}x_i - \beta_{0k'}|.$$

Da jeder Punkt nur von einer Geraden approximiert werden soll, darf für jedes i nur ein $\pi_{ik} = 1$ sein, siehe (5.3d). Außerdem lässt sich mit diesen Variablen auch das Trimmen modellieren, indem man fordert, dass genau m Punkte überhaupt einer Geraden zugeordnet werden, siehe (5.3e). Wie schon für (LP-LAD) führt man Variablen d_{ik} ein, die größer sind als der Abstand von Punkt z_i zur Geraden $\ell(\beta_{0k}, \beta_{1k})$:

$$d_{ik} \geq |y_i - x_i\beta_{1k} - \beta_{0k}| \quad \forall i = 1, \dots, n \quad \forall k = 1, \dots, K,$$

siehe (5.3b) und (5.3c). Dann hat man im Falle einer Optimallösung wieder Gleichheit, wenn $\pi_{ik} = 1$, da minimiert wird. In der Zielfunktion wird dann über alle d_{ik} summiert, wenn der Punkt z_i der Geraden $\ell(\beta_{0k}, \beta_{1k})$ zugeordnet ist, siehe (5.3a). Mit diesen Mitteln erhält man das Programm

$$(Q\text{-}K\text{-TLAD}) \quad \min \sum_{k=1}^K \sum_{i=1}^n \pi_{ik} d_{ik} \tag{5.3a}$$

$$\text{s. d. } y_i - x_i\beta_{1k} - \beta_{0k} \leq d_{ik} \quad \forall i \tag{5.3b}$$

$$x_i\beta_{1k} + \beta_{0k} - y_i \leq d_{ik} \quad \forall i \tag{5.3c}$$

$$\sum_{k=1}^K \pi_{ik} \leq 1 \quad \forall i \tag{5.3d}$$

$$\sum_{k=1}^K \sum_{i=1}^n \pi_{ik} = m \tag{5.3e}$$

$$\beta_{0k}, \beta_{1k} \in \mathbb{R} \quad \forall k$$

$$\pi_{ik} \in \{0, 1\}, \quad d_{ik} \geq 0 \quad \forall i, k.$$

Es ist nicht äquivalent zu (K -TLAD), hat aber nach der gleichen Argumentation wie für (LP-LAD) dieselben Optimallösungen. Es ist außerdem wegen der Entscheidungsvariablen π_{ik} teilweise ganzzahlig und in der Zielfunktion quadratisch.

Durch Einführen neuer Variablen und Nebenbedingungen wollen wir die Zielfunktion nun linearisieren (nach A. Schöbel). Dazu betrachten wir den Term $\pi \cdot x$ für eine Entscheidungsvariable $\pi \in \{0, 1\}$ und eine nicht-negative und beschränkte kontinuierliche Variable $x \in [0, c]$. Dann führt man die neue Variable z ein und stellt durch Nebenbedingungen sicher, dass

$$z = \pi \cdot x = \begin{cases} 0 & \text{falls } \pi = 0 \\ x & \text{falls } \pi = 1 \end{cases}.$$

Die Nebenbedingungen sind

$$z \leq \pi \cdot c \tag{5.4a}$$

$$z \leq x \tag{5.4b}$$

$$z \geq x - c(1 - \pi) \tag{5.4c}$$

$$z \geq 0. \tag{5.4d}$$

Wenn nämlich $\pi = 0$, erhält man aus (5.4a) und (5.4d) $z = 0$, während (5.4b) erfüllt ist, da $x \geq 0$ und (5.4c) zu $z \geq x - c$ wird und mit $z = 0$ auch erfüllt ist, da $x \leq c$. Ist hingegen $\pi = 1$, wird (5.4c) zu $z \geq x$ und erzwingt zusammen mit (5.4b) $z = x$. Außerdem sind (5.4a) und (5.4b) erfüllt, da $x \in [0, c]$ gilt.

Jetzt sind wir in der Lage, (Q- K -TLAD) als gemischt-ganzzahliges Programm (mixed integer program, MIP) zu formulieren, indem wir diese Beobachtung auf die $\pi_{ik}d_{ik}$ anwen-

den (Xpress-Implementierung in Datei MIP-K-TLAD.mos auf beiliegenden Datenträger):

$$\begin{aligned}
 \text{(MIP-K-TLAD)} \quad & \min \sum_{k=1}^K \sum_{i=1}^n z_{ik} \\
 \text{s. d.} \quad & y_i - x_i \beta_{1k} - \beta_{0k} \leq d_{ik} && \forall i, k \\
 & x_i \beta_{1k} + \beta_{0k} - y_i \leq d_{ik} && \forall i, k \\
 & \sum_{k=1}^K \pi_{ik} \leq 1 && \forall i \\
 & \sum_{k=1}^K \sum_{i=1}^n \pi_{ik} = m \\
 & z_{ik} \leq \pi_{ik} \cdot c_i && \forall i, k \\
 & z_{ik} \leq d_{ik} && \forall i, k \\
 & z_{ik} \geq d_{ik} - c_i(1 - \pi_{ik}) && \forall i, k \\
 & z_{ik} \geq 0 && \forall i, k \\
 & \beta_{0k}, \beta_{1k} \in Q && \forall k \\
 & \pi_{ik} \in \{0, 1\}, \quad d_{ik} \geq 0 && \forall i, k
 \end{aligned}$$

mit oberen Schranken c_i für alle d_{i1}, \dots, d_{iK} , die noch bestimmt werden müssen. Wir wissen schon, dass für jede optimale Lösung

$$d_{ik}^{\text{opt}} = |y_i - \beta_{1k}^{\text{opt}} x_i - \beta_{0k}^{\text{opt}}| \quad \forall i = 1, \dots, n \quad \forall k = 1, \dots, K \text{ mit } \pi_{ik}^{\text{opt}} = 1$$

gilt. Wenn $\pi_{ik}^{\text{opt}} = 0$, ist der Wert von d_{ik}^{opt} egal, da er die Zielfunktion nicht verändert, und kann beliebig eingeschränkt werden. Also schätzen wir einfach alle d_{ik}^{opt} nach oben ab und nehmen an, dass $(\beta_{0k}, \beta_{1k}) \in Q = [\underline{q}_0, \bar{q}_0] \times [\underline{q}_1, \bar{q}_1]$ von Seite 51, daher auch die entsprechenden Nebenbedingung im Programm. So liefert (MIP-K-TLAD) zwar nicht alle optimalen Lösungen von (K-TLAD), aber mindestens eine (siehe Satz 4.7). Weil

$$d_{ik}^{\text{opt}} = |y_i - \beta_{1k}^{\text{opt}} x_i - \beta_{0k}^{\text{opt}}| = \max\{y_i - \beta_{1k}^{\text{opt}} x_i - \beta_{0k}^{\text{opt}}, \beta_{1k}^{\text{opt}} x_i + \beta_{0k}^{\text{opt}} - y_i\},$$

schätzen wir die beiden Argumente des Maximums getrennt ab durch

$$\begin{aligned}
 y_i - \beta_{1k}^{\text{opt}} x_i - \beta_{0k}^{\text{opt}} &\leq \max_{(\beta_{0k}, \beta_{1k}) \in Q} \{y_i - \beta_{1k} x_i - \beta_{0k}\} \\
 &\leq y_i - \min_{\beta_{1k} \in [\underline{q}_1, \bar{q}_1]} \{\beta_{1k} x_i\} - \min_{\beta_{0k} \in [\underline{q}_0, \bar{q}_0]} \{\beta_{0k}\} \\
 &= \begin{cases} y_i - \underline{q}_1 x_i - \underline{q}_0 & \text{falls } x_i \geq 0 \\ y_i - \bar{q}_1 x_i - \underline{q}_0 & \text{falls } x_i \leq 0 \end{cases} =: c_i^{(1)}
 \end{aligned}$$

bzw.

$$\begin{aligned}
 \beta_{1k}^{\text{opt}} x_i + \beta_{0k}^{\text{opt}} - y_i &\leq \max_{(\beta_{0k}, \beta_{1k}) \in Q} \{\beta_{1k} x_i + \beta_{0k} - y_i\} \\
 &\leq \max_{\beta_{1k} \in [\underline{q}_1, \bar{q}_1]} \{\beta_{1k} x_i\} + \max_{\beta_{0k} \in [\underline{q}_0, \bar{q}_0]} \{\beta_{0k}\} - y_i \\
 &= \begin{cases} \bar{q}_1 x_i + \bar{q}_0 - y_i & \text{falls } x_i \geq 0 \\ \underline{q}_1 x_i + \bar{q}_0 - y_i & \text{falls } x_i \leq 0 \end{cases} =: c_i^{(2)}
 \end{aligned}$$

und können so

$$d_{ik} \leq c_i := \max\{c_i^{(1)}, c_i^{(2)}\} \quad (5.5)$$

für alle $i = 1, \dots, n$ und alle $k = 1, \dots, K$ einschränken. Wenn Q schon bekannt ist, kann man das in konstanter Zeit für jedes Paar (i, k) tun, insgesamt also in $O(nK)$.

Ein großer Nachteil dieses Programms ist seine große Anzahl an $(3n + 2)K$ Variablen, von denen nK von booleschem Typ sind, und $5nK + n + 2K + 1$ Nebenbedingungen. Außerdem sind $2K$ der Variablen frei, also nicht vorzeichenbeschränkt. Durch Einführung eines geeigneten großen M wie in [3, S. 260] durch Bertsimas und Shioda erhält man folgendes Programm (Xpress-Implementierung in Datei BS-K-TLAD.mos auf beiliegenden Datenträger):

$$\begin{aligned}
 \text{(BS-K-TLAD)} \quad \min \quad & \sum_{k=1}^K \sum_{i=1}^n d_{ik} \\
 \text{s. d.} \quad & y_i - x_i \beta_{1k} - \beta_{0k} \leq d_{ik} + M(1 - \pi_{ik}) & \forall i, k \\
 & x_i \beta_{1k} + \beta_{0k} - y_i \leq d_{ik} + M(1 - \pi_{ik}) & \forall i, k \\
 & \sum_{k=1}^K \pi_{ik} \leq 1 & \forall i \\
 & \sum_{k=1}^K \sum_{i=1}^n \pi_{ik} = m \\
 & \beta_{0k}, \beta_{1k} \in Q & \forall k \\
 & \pi_{ik} \in \{0, 1\}, \quad d_{ik} \geq 0 & \forall i, k
 \end{aligned}$$

M sei dabei fest und so groß gewählt, dass die erste und zweite Nebenbedingung für $\pi_{ik} = 0$ redundant werden. Hier hat man nur $(2n + 2)K$ Variablen, davon nK boolesch und $2K$ frei, und $2nK + n + 1$ Nebenbedingungen, wodurch diese Formulierung beim Lösen vorzuziehen ist. Die Bestimmung von M ist in linearer Zeit in n und K möglich, wenn ein Kompaktum Q wie von Seite 51 bekannt ist, denn M muss $M > |y_i - \beta_{0k} - \beta_{1k}|$ für alle $i = 1, \dots, n$ und alle $k = 1, \dots, K$ erfüllen, kann also gesetzt werden als $M := \max_{i=1, \dots, n} c_i$ mit den c_i aus (5.5).

Leider hat sich herausgestellt, dass (BS- K -TLAD) durch einen Solver wie Xpress nicht schnell lösbar ist. Schon kleine Beispiele mit $K = 2$ und $n = 50$ beanspruchten mehrere Stunden Rechenzeit. Das liegt daran, dass dieses Programm durch Xpress mit einem Branch&Bound-Verfahren gelöst wird, aber während des Lösens keine gute untere Schranke für den Zielfunktionswert gefunden wird. Dadurch muss fast jeder Ast des B&B-Baums vollständig abgesucht werden, was entsprechend lange dauert. Dass diese unteren Schranken nicht gefunden werden können liegt an der speziellen Struktur des (BS- K -TLAD). Durch das Einführen von M kann in der LP-Relaxation, die als Quelle für untere Schranken verwendet werden könnte, das m aus der letzten Nebenbedingung recht gleichmäßig auf die π_{ik} aufgeteilt werden, das heißt $\pi_{ik} \approx \frac{1}{mK}$. Wenn M und m nun entsprechend groß sind, bleibt auch $(1 - \pi_{ik})M \approx (1 - \frac{1}{mK})M$ recht groß, so dass viele – vielleicht sogar alle d_{ik} auf Null gesetzt werden können. Dann verschwindet aber der optimale Zielfunktionswert der LP-Relaxation und liefert daher eine unbrauchbare weil triviale untere Schranke. Das Duale der LP-Relaxation liefert dann ebenso einen optimalen Zielfunktionswert von Null und einfach nur zulässige Lösungen des Dualen können sogar negative Zielfunktionswerte ergeben, die wiederum als untere Schranke unbrauchbar sind. Insgesamt kann es also schwer werden, an gute Schranken zu gelangen, und die schlechte Leistungsfähigkeit des Solvers lässt sich so begründen.

Um die Anzahl an Variablen und Nebenbedingungen zu reduzieren, schlagen Bertsimas und Shioda [3] ein Preprocessing der Daten vor. Dabei werden mehrere nahe beieinander liegenden Punkt zu einem Cluster zusammengefasst um im Weiteren als Einheit behandelt. Somit entsteht ein heuristisches Verfahren. Die Autoren bezeichnen ihre Ergebnisse als ausbaufähig, sind aber zuversichtlich, dass sich das Investieren von mehr Arbeit in ihre Methode lohnt. Insbesondere sei es wichtig, die große Zahl M möglichst klein zu wählen, was aus obigem Grund sehr hilfreich wäre. Falls das gelingt und noch einige andere Probleme beseitigt werden können, sei ihre Methode (CRIO = classification and regression via integer optimization) auf einem guten Weg. Bisher hat sie besonders bei großen Datenmengen mit Laufzeitproblemen zu kämpfen.

5.2 Der FastTLE-Algorithmus

Der FastTLE-Algorithmus von Neykov u. a. [24] ist eine Heuristik zur Lösung des (K -TLAD)-Problems. Das heißt, er findet nicht unbedingt die global optimale Lösung, sondern eine lokal optimale Lösung. Das liegt unter anderem daran, dass er auf den EM-Algorithmus von Dempster u. a. [7] aufbaut, der auch eine Heuristik ist. Der EM-Algorithmus dient dazu, Maximum-Likelihood-Schätzer für statistische Modelle mit fehlender Information (incomplete data) wie die Regression mit latenten Klassen (siehe Abschnitt 2.5) zu finden. Der FastTLE-Algorithmus fügt die Funktionalität zur Behand-

lung von Daten mit Ausreißern hinzu. Das heißt, der EM-Algorithmus löst (K -LAD) heuristisch und der FastTLE-Algorithmus (K -TLAD). Die zugrunde liegende Implementierung des EM-Algorithmus (flexmix) ist von Leisch [19]. FastTLE und flexmix sind als R-Packages verfügbar [27].

Die Vorgehensweise des FastTLE besteht aus zwei Schritten. Im ersten Schritt wird eine Stichprobe der Größe k^* der Daten gewählt und auf deren Basis ein Start-Schätzer durch flexmix bestimmt. Neykov u. a. geben einen sinnvollen Wert d für k^* an. Der so erhaltene Schätzer ist also ein heuristisch bestimmter ML-Schätzer für eine Teilmenge der gegebenen Daten. In diesem Schritt kann auch ein anderer Start-Schätzer gewählt werden, zum Beispiel ein tatsächlich optimaler ML-Schätzer. FastTLE leistet das allerdings nicht. Im zweiten Schritt werden die k Datenpunkte bestimmt, die für die Start-Schätzung die größten Loglikelihood-Werte haben, und für diese k Punkte wird wieder heuristisch ein ML-Schätzer bestimmt. Laut Neykov u. a. ist $k = \lfloor (n + d + 1)/2 \rfloor$ eine empfehlenswerte Wahl. Diese beiden Schritte werden wiederholt und der beste Schätzer aller Wiederholungen wird als Ergebnis des FastTLE ausgegeben. Das Hauptproblem dieses Verfahrens

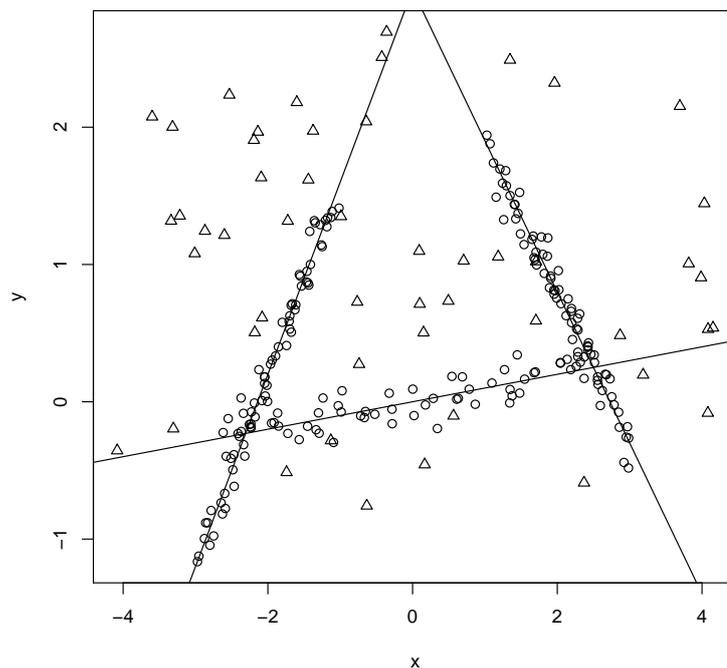


Abbildung 5.1: Lineares Modell mit drei latenten Klassen, 250 Punkte

ist, dass es nur eine Heuristik ist und nur lokal optimale Lösungen findet. Dazu geben wir ein Beispiel. Wir betrachten ein Modell mit drei Klassen:

$$\begin{aligned} Y_{i1} &= 3 + 1.4x_{i1} + \epsilon_{i1} & \forall i = 1, \dots, 70 \\ Y_{i2} &= 3 - 1.1x_{i2} + \epsilon_{i2} & \forall i = 1, \dots, 70 \\ Y_{i3} &= 0.1x_{i3} + \epsilon_{i3} & \forall i = 1, \dots, 60 \end{aligned}$$

mit Fehlern $\epsilon_{ik} \sim \mathcal{N}(0, \sigma^2)$ und $\sigma = 0, 1$. Die x -Koordinaten der Modellpunkte werden durch Gleichverteilungen auf $[-3, -1]$ für Klasse 1, $[1, 3]$ für Klasse 2 und $[-3, 3]$ für Klasse 3 erzeugt. Schließlich werden Ausreißer durch eine Gleichverteilung auf $[-4.5, 4.5] \times [-0.8, 2.8]$ hinzugefügt. Es ergibt sich ein Bild wie in Abbildung 5.1. Bestimmt man nun mit dem FastTLE-Algorithmus die Modellparameter, ergeben sich Lösungen, die offenbar nicht global optimal sind. Tabelle 5.1 zeigt die Loglikelihood-Werte für normal- und Laplace-verteilte Fehler von zwei FastTLE-Schätzungen, denen als Startwerte die korrekte Anzahl an Klassen ($K = 3$) und die korrekte Anzahl an Modellpunkten ($m = 200$) übergeben wurden (siehe auch Datei `neykov.r` auf beiliegendem Datenträger). Da für

	wahre Parameter	FastTLE Schätzung 1	FastTLE Schätzung 2
normalverteilt	-24.299	-97.784	-66.312
Laplace-verteilt	-35.293	-104.754	-75.944

Tabelle 5.1: Loglikelihood-Werte für das Modell aus Abbildung 5.1 und zugehörige FastTLE-Schätzungen (siehe auch Abbildung 5.2)

den Algorithmus im ersten Schritt eine zufällige Stichprobe der Daten für eine Start-Schätzung gewählt wird, können verschiedene Ergebnisse trotz gleicher Eingangsdaten auftreten, wie hier auch zu sehen ist. Die zweite Schätzung ist besser als die erste, aber die Loglikelihood-Werte der wahren Modellparameter sind noch einmal deutlich besser als die geschätzten. Da das globale Optimum einen möglicherweise noch besseren Loglikelihood-Wert liefert, ist klar, dass eine Heuristik hier nicht ausreicht. Diesen Eindruck verstärkt die Abbildung 5.2 noch. Hier sind die Geraden aus den beiden Schätzungen eingezeichnet. Die erste Schätzung hat nur eine der Geraden richtig platziert, die zweite Schätzung immerhin zwei. Insgesamt wurde aber die Struktur, die durch die Punkte beschrieben wird, nicht erkannt. Gerade das sollte aber die eigentliche Aufgabe sein, wenn man an Anwendungen denkt – zum Beispiel in der Medizin wie in der Einleitung. Man kann dieses Problem zu einem gewissen Grad beheben, indem man mehr FastTLE-Schätzungen durchführt und schließlich diejenige mit dem besten Zielfunktionswert verwendet. So bleibt die Hoffnung, dass man zufällig die Lösung trifft, die die Struktur der Daten tatsächlich beschreibt. Allerdings bleibt in der Praxis dann weiterhin unklar, ob wirklich ein globales Optimum gefunden wurde, weil dort kein ein Vergleich mit den wahren Parametern erfolgen kann, da diese ja unbekannt sind.

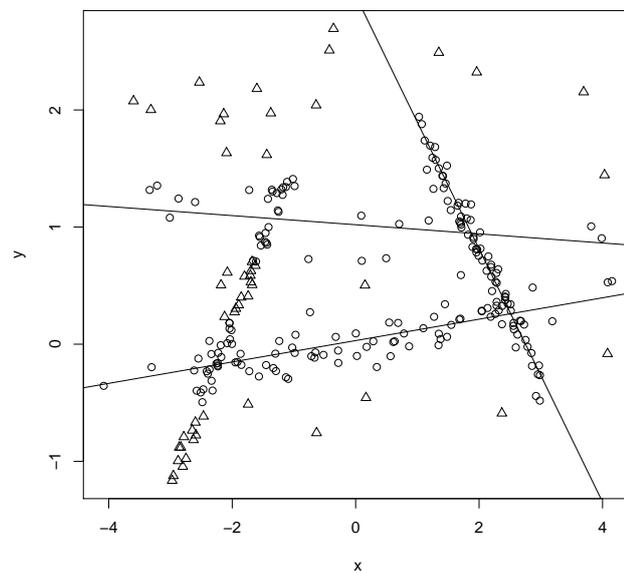
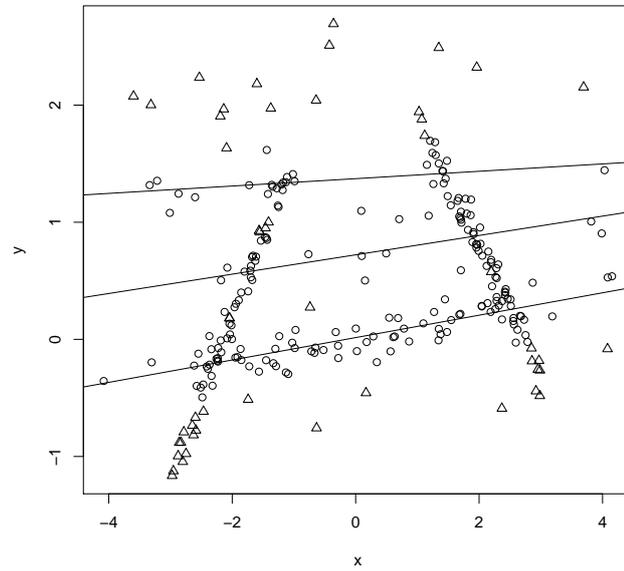


Abbildung 5.2: FastTLE-Schätzung 1 (oben) und 2 (unten) für das Modell aus Abbildung 5.1

5.3 Der FDS-Algorithmus

Der Vollständigkeit halber wird hier noch einmal der Algorithmus beschrieben, der das (K -TLAD)-Problem mit Hilfe der endlichen Kandidatenmenge aus Satz 4.7 löst. Er soll ab jetzt FDS-Algorithmus (finite dominating set = endliche Kandidatenmenge) heißen, denn in Abschnitt 5.4 soll er mit dem FastTLE-Algorithmus in einer Simulation verglichen werden.

Da der FDS auf Satz 4.7 beruht, müssen im Wesentlichen die $2K$ -elementigen Teilmengen der n Beobachtungen abgezählt werden. Dies geschieht durch eine Art kanonische Ordnung der Teilmengen, die man durch die Bijektion

$$\begin{aligned} & \{(i_1, \dots, i_L) : i_1, \dots, i_L \in \{1, \dots, n\}, i_1 < \dots < i_L\} \\ & \qquad \qquad \qquad \rightarrow \\ & \{I \subset \{1, \dots, n\} : I = \{i_1, \dots, i_L\}\} \end{aligned}$$

erhält. Dann kann man sagen, dass eine Teilmenge $I = \{i_1, \dots, i_L\}$ kleiner ist als eine andere $I' = \{i'_1, \dots, i'_L\}$, wenn für $\tilde{k} = \arg \min_k i_k \neq i'_k$ gilt, dass $i_{\tilde{k}} < i'_{\tilde{k}}$. Gilt $i_{\tilde{k}} > i'_{\tilde{k}}$, ist I größer als I' , und wenn \tilde{k} nicht existiert ist $I = I'$.

Nun lässt sich der FDS-Algorithmus folgendermaßen formulieren:

FDS-Algorithmus

Input: $z_1 = (x_1, y_1), \dots, z_n = (x_n, y_n) \in \mathbb{R}^2$
 $m \in \mathbb{N} : m \leq n$
 $K \in \mathbb{N} : 2K \leq m$

Initialisierung: Setze $I = \{i_1, \dots, i_{2K}\} := \{1, \dots, 2K\}$, $LAD^* := \infty$.

Schritt 1: Für $k = 1, \dots, K$ berechne

$$\begin{aligned} \beta_{1k} &:= \frac{y_{i_{2k-1}} - y_{i_{2k}}}{x_{i_{2k-1}} - x_{i_{2k}}}, \\ \beta_{0k} &:= y_{i_{2k}} - \beta_{1k} x_{i_{2k}}. \end{aligned}$$

Schritt 2: Berechne für alle $k = 1, \dots, K$ und alle $i = 1, \dots, n$

$$d_{ik} := d_v(\ell(\beta_{0k}, \beta_{1k}), z_i) = |y_i - \beta_{1k} x_i - \beta_{0k}|.$$

Schritt 3: Bestimme für alle $i = 1, \dots, n$

$$\tilde{k}_i := \arg \min_{k=1, \dots, K} d_{ik}$$

Schritt 4: Sortiere durch eine Permutation (\cdot) die Ausdrücke $d_{i\tilde{k}_i}$ absteigend, d. h.

$$d_{(1)\tilde{k}_{(1)}} \leq \dots \leq d_{(n)\tilde{k}_{(n)}}.$$

Schritt 5: Bilde $LAD := \sum_{i=1}^m d_{(i)\tilde{k}_{(i)}}$.

Falls $LAD < LAD^*$:

$$LAD^* := LAD,$$

$$\beta^* := (\beta_{01}, \beta_{11}, \dots, \beta_{0K}, \beta_{1K})$$

$$c := (\tilde{k}_1, \dots, \tilde{k}_n)$$

Schritt 6: Falls $I \neq \{n - 2K + 1, \dots, n\}$, gehe zur nächsten Indexmenge und zu Schritt 1.

Output: Parameter der für (K -TLAD) optimalen Geraden $\ell(\beta_{0k}^*, \beta_{1k}^*)$

Optimaler Zielfunktionswert LAD^*

Optimale Klassenzugehörigkeiten c_1, \dots, c_n

Hieran kann man noch einmal sehr gut die Zeitkomplexität des Algorithmus ablesen. Dazu nehmen wir an, dass $K \leq \log n$.

Schritt 1: $O(K) = O(\log n)$

Schritt 2: $O(nK) = O(n \log n)$

Schritt 3: $O(nK) = O(n \log n)$

Schritt 4: $O(n \log n)$ nach Hoare [10]

Schritt 5: $O(m) = O(n)$

Schritt 6: konstante Zeit

Da die Schritte insgesamt $\binom{n}{2K} = O(n^{2K})$ mal ausgeführt werden müssen ergibt sich wie schon gesehen die Zeitkomplexität $O(n^{2K+1} \log n)$. Eine Implementierung des FDS-Algorithmus findet sich in der Datei `Simulation\fds.r` auf dem beiliegenden Datenträger.

5.4 Vergleich der Algorithmen

Um den FastTLE- und den FDS-Algorithmus zu vergleichen, wurde eine Simulation durchgeführt. Dazu wurden zufällige Regressionmodelle mit latenten Klassen und in

verschiedenen Größen erzeugt und anschließend Schätzer mit beiden Algorithmen bestimmt. Die Ergebnisse werden verglichen in Bezug auf ihre Laufzeit und die Qualität des erhaltenen Schätzers. Aufgrund der langen Laufzeit des FDS-Algorithmus für große Klassenzahlen K sind hier nur Modelle mit $K = 1, 2, 3$ zum Einsatz gekommen.

Die zufällige Erzeugung eines Modells verläuft folgendermaßen: Wähle zu gegebenen K , n und m zufällige Klassengrößen n_1, \dots, n_K und in jeder Klasse k die Parameter $\beta_{0k}, \beta_{1k} \in [-5, 5]$ und $\sigma_k \in [0.1, 1]$ zufällig gemäß einer Gleichverteilung. Wähle die Messstellen x_{ik} , $i = 1, \dots, n_k$ gleichverteilt aus $[a_k, a_k + b_k]$, wobei a_k gleichverteilt aus $[-5, 5]$ und b_k aus $[3, 5]$ gewählt wird. Erzeuge dann die endgültigen Beobachtungen durch

$$Y_{ik} = \beta_{0k} + \beta_{1k}x_{ik} + \epsilon_{ik} \quad \forall i = 1, \dots, n_k \text{ und } \forall k = 1, \dots, K,$$

mit $\epsilon_{1k}, \dots, \epsilon_{n_k k} \sim \mathcal{N}(0, \sigma_k)$ für alle $k = 1, \dots, K$. Alle Fehler sollen außerdem unabhängig sein. Füge schließlich noch $n - m$ Ausreißer hinzu.

Zur Simulation wird hier immer $m = 0.8n$ gewählt, um die Anzahl der verschiedenen Modelle klein zu halten. Ein Modell wird also nach Vorgabe der freien Parameter K und n zufällig erzeugt. Es werden L Instanzen des Modells erzeugt und dann die Schätzungen mit beiden Algorithmen durchgeführt. Um die Güte der Approximation unter Annahme normalverteilter Fehler zu vergleichen, wird der Mittelwert der getrimmten Summe der quadrierten Residuen (trimmed sum of squared residuals, TSSR) und der entsprechenden Loglikelihoodfunktion über die jeweils L Schätzungen gebildet. Analog kann die Güte der Approximation unter Annahme Laplace-verteilter Fehler durch Mittelung über die getrimmte Summe der absoluten Abweichungen (trimmed sum of absolute deviations, TSAD) und wiederum die entsprechende Loglikelihoodfunktion ermittelt werden. Außerdem wird zu jedem Schätzer eine Art relativer Abstand der geschätzten Parameter zu den wahren Parametern angegeben. Für eine Gerade, also $K = 1$ ist der Abstand einfach

$$\frac{\sqrt{(\hat{\beta}_0 - \beta_0)^2 + (\hat{\beta}_1 - \beta_1)^2}}{\sqrt{\beta_0^2 + \beta_1^2}},$$

wobei β_0 und β_1 die wahren Parameter sind und $\hat{\beta}_0$ und $\hat{\beta}_1$ die geschätzten. Bei mehreren Geraden, also $K \geq 2$ kann der Abstand nicht einfach durch

$$\frac{\sqrt{\sum_{k=1}^K (\hat{\beta}_{0k} - \beta_{0k})^2 + (\hat{\beta}_{1k} - \beta_{1k})^2}}{\sqrt{\sum_{k=1}^K \beta_{0k}^2 + \beta_{1k}^2}} \quad (5.6)$$

berechnet werden, weil das statistische Modell nicht identifizierbar ist. Das bedeutet, dass die Schätzung $\hat{\beta}_{0k}, \hat{\beta}_{1k}$ nicht unbedingt β_{0k}, β_{1k} schätzt. Das sieht man am besten an einem Bild wie in Abbildung 5.3.

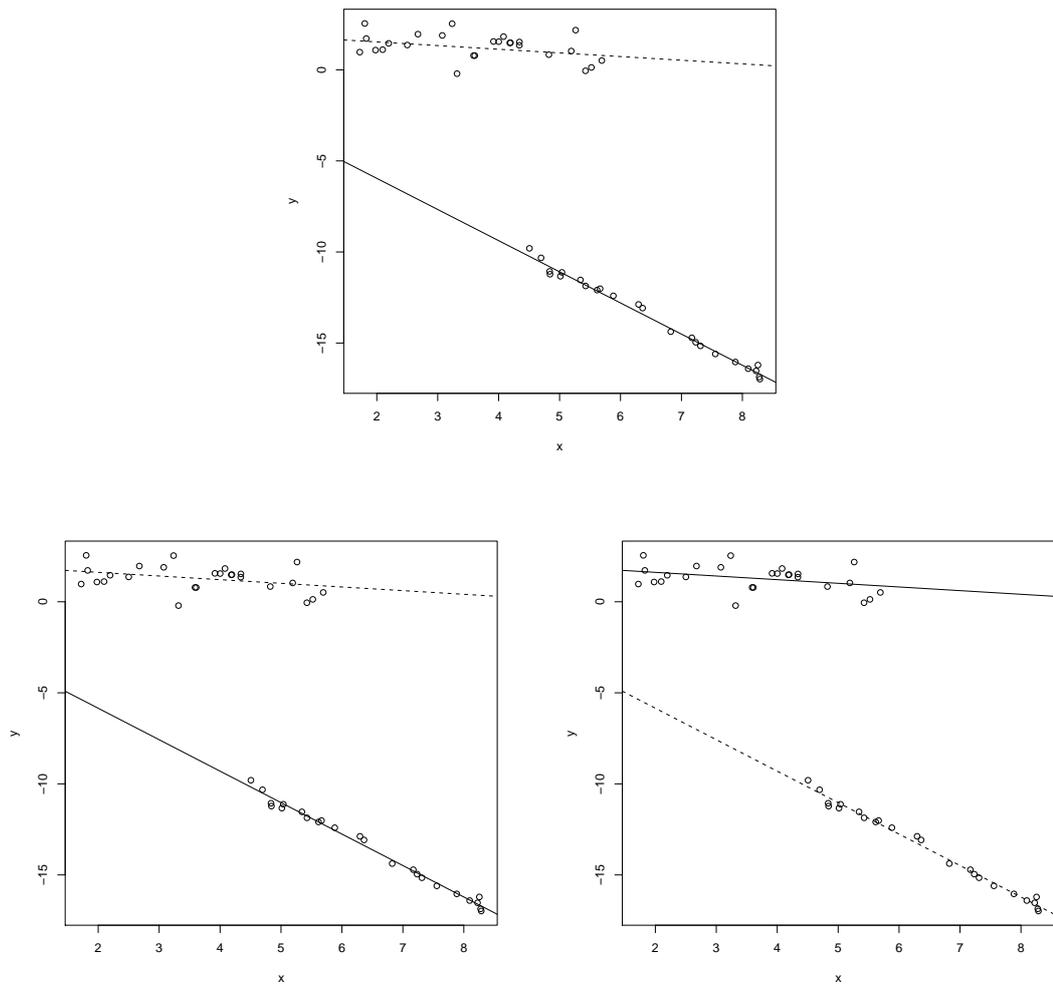


Abbildung 5.3: oben: ein Modell mit zwei latenten Klassen; links: Schätzer (1) für das Modell; rechts: Schätzer (2) für das Modell

Hier sind ein Modell mit zwei Klassen ohne Ausreißer und zwei verschiedenen Schätzungen für dieses gezeigt. Die Parameter der jeweils zwei Geraden sind (auf zwei Nachkommastellen gerundet)

$$\begin{array}{cccc}
 \beta_{01} = -2.54 & \beta_{11} = -1.71 & \beta_{02} = 1.94 & \beta_{12} = -0.20 \\
 \hat{\beta}_{01}^{(1)} = -2.39 & \hat{\beta}_{11}^{(1)} = -1.73 & \hat{\beta}_{02}^{(1)} = 2.02 & \hat{\beta}_{12}^{(1)} = -0.20 \\
 \hat{\beta}_{01}^{(2)} = 2.02 & \hat{\beta}_{11}^{(2)} = -0.20 & \hat{\beta}_{02}^{(2)} = -2.39 & \hat{\beta}_{12}^{(2)} = -1.73.
 \end{array}$$

Daraus ergeben sich durch Einsetzen in (5.6) die Werte 0.05 für Schätzer (1) und 1.83 für Schätzer (2). Dabei sind beide Schätzer gleich gut und im Wesentlichen identisch. Nur die Cluster sind einmal wie im Modell durchnummeriert und einmal anders.

Um doch so etwas wie einen relativen Abstand angeben zu können, wird daher der Ausdruck

$$\min_{\tau} \frac{\sqrt{\sum_{k=1}^K (\hat{\beta}_{0\tau(k)} - \beta_{0k})^2 + (\hat{\beta}_{1\tau(k)} - \beta_{1k})^2}}{\sqrt{\sum_{k=1}^K \beta_{0k}^2 + \beta_{1k}^2}}$$

angegeben, wobei τ alle Permutationen von $\{1, \dots, K\}$ durchläuft.

Es ist besonders hervorzuheben, dass der FDS-Algorithmus TSAD (global) minimiert, während der FastTLE-Algorithmus die Loglikelihoodfunktion im normalverteilten Fall (lokal) maximiert. Der Einfachheit halber sind in Tabelle 5.2 die Werte der negativen Loglikelihoodfunktion anstelle der normalen Loglikelihoodfunktion angegeben. Somit bedeutet in allen Spalten ein niedrigerer Wert ein besseres Ergebnis. Zusätzlich zu den FDS- und FastTLE-Schätzern (Einträge „FDS“ und „TLE“) sind in den Tabellen auch noch die wahren Parameter des Modells (Einträge „model“) verzeichnet und ein kombinierter Schätzer, bei dem der FastTLE-Algorithmus die Schätzer des FDS-Algorithmus als Startwerte verwendet (Einträge „mix“). In der Spalte „Laplace“ stehen die Werte der negativen Loglikelihoodfunktion bei Annahme Laplace-verteilter Fehler und in der Spalte „normal“ die Werte für normalverteilte Fehler. Alle Werte sind auf drei Nachkommastellen gerundet. Die Simulation wurde mit R programmiert und der entsprechende Quelltext befindet sich im Ordner `Simulation` auf dem beiliegenden Datenträger.

Man sieht, dass der FDS-Algorithmus dem FastTLE-Algorithmus bei der Güte der Approximation deutlich überlegen ist. Für $K = 1$ und kleine n ist er sogar schneller. Bei großen n oder $K \geq 2$ macht sich die Zeitkomplexität des FDS-Algorithmus von $O(n^{2K+1} \log n)$ stark bemerkbar.

Ein interessanter Punkt ist die Verwendung des FDS-Algorithmus zur Berechnung von Startlösungen für den FastTLE-Algorithmus. Geht man von normalverteilten Fehlern

5 Algorithmen

Algorithmus	n	m	K	L	TSAD	Laplace	TSSR	normal	rel.Fehler	Laufzeit
model	25	20	1	100	8.962	14.813	7.806	13.301	0	0
FDS	25	20	1	100	7.792	11.345	6.635	11.145	0.132	0.124
TLE	25	20	1	100	9.347	13.686	10.098	12.276	0.210	6.683
mix	25	20	1	100	8.045	11.947	6.220	10.587	0.130	6.763
model	50	40	1	100	18.358	32.543	15.439	29.624	0	0
FDS	50	40	1	100	16.776	28.089	13.398	26.579	0.086	0.502
TLE	50	40	1	100	21.410	34.382	25.965	31.176	0.268	7.325
mix	50	40	1	100	17.071	28.750	12.841	25.830	0.075	7.335
model	100	80	1	100	34.729	61.12	28.881	55.43	0	0
FDS	100	80	1	100	31.887	52.138	24.062	48.258	0.101	2.118
TLE	100	80	1	100	36.004	57.901	34.228	52.416	0.151	9.065
mix	100	80	1	100	32.175	52.873	23.737	47.672	0.085	9.067
model	250	200	1	100	91.187	152.927	80.848	138.369	0	0
FDS	250	200	1	100	84.349	134.787	66.479	122.130	0.075	14.902
TLE	250	200	1	100	85.972	137.937	68.176	123.836	0.091	14.509
mix	250	200	1	100	84.643	135.456	66.053	121.508	0.071	14.540
model	500	400	1	100	187.433	323.152	165.905	294.743	0	0
FDS	500	400	1	100	177.682	293.889	143.149	268.504	0.036	71.945
TLE	500	400	1	100	179.204	296.365	145.250	269.688	0.058	23.193
mix	500	400	1	100	177.951	294.526	142.798	267.970	0.048	23.083
model	50	40	2	100	17.051	48.767	14.555	46.013	0	0
FDS	50	40	2	100	13.262	40.432	9.525	41.076	0.418	105.375
TLE	50	40	2	100	17.163	46.907	16.689	45.035	0.651	8.290
mix	50	40	2	100	14.369	41.934	10.123	40.060	0.461	8.413
model	100	80	2	24	32.187	101.161	24.929	95.475	0	0
FDS	100	80	2	24	28.334	89.962	19.546	87.576	0.107	1883.403
TLE	100	80	2	24	33.926	101.164	26.306	96.487	0.416	10.890
mix	100	80	2	24	29.134	91.190	19.603	86.863	0.100	11.169
model	35	28	3	24	4.260	22.813	1.094	20.765	0	0
FDS	35	28	3	24	3.073	12.727	0.746	14.368	0.088	769.741
TLE	35	28	3	24	7.640	30.692	4.562	28.661	1.111	8.751
mix	35	28	3	24	3.290	14.344	0.674	12.516	0.117	9.068
model	40	32	3	24	4.889	25.604	1.279	23.153	0	0
FDS	40	32	3	24	3.772	16.185	0.924	17.366	0.090	1840.734
TLE	40	32	3	24	7.169	29.501	3.878	27.932	0.739	9.266
mix	40	32	3	24	4.046	17.786	0.893	15.915	0.141	9.667

Tabelle 5.2: Durchschnittliche Ergebnisse der Schätzverfahren FDS, FastTLE und FastTLE mit Startwerten aus FDS

aus, können die TSSR-Werte des FastTLE durch die Startwerte deutlich verbessert werden. Da FastTLE eine recht kurze Laufzeit hat, lohnt es sich eigentlich immer, ihn nach dem FDS-Algorithmus noch einmal zusätzlich zu verwenden, um möglicherweise eine Verbesserung der Schätzung zu erzielen.

Insgesamt ist der FDS-Algorithmus dem FastTLE-Algorithmus also vorzuziehen, wenn die Datenmenge klein ist oder wenn unbedingt sichergestellt werden muss, dass man einen global optimalen Schätzer erhält. Wenn eine lange Rechenzeit akzeptabel ist, kann er auch für größere Datensätze verwendet werden. Das bietet sich besonders für einmalige Analysen an, wenn zum Beispiel die Datenerhebung abgeschlossen ist und nicht damit zu rechnen ist, dass innerhalb kurzer Zeit neue Daten erhoben werden müssen, da die alten nicht mehr gültig sind. Ein gutes Beispiel ist sicher die Analyse der Wirksamkeit eines neuen Medikaments, denn es muss nur einmal getestet werden und nicht immer wieder neu. Außerdem sollte hier mit maximaler Genauigkeit gearbeitet werden, da falsche Schlüsse aus den Daten, wie sie durch falsches Erkennen ihrer Struktur gezogen werden könnten (siehe Abschnitt 5.2), schwere Folgen nach sich ziehen können.

6 Ausblick

Um die LAD-Methode in der Praxis nutzen zu können, wären einige weitere Ergebnisse sehr hilfreich. Am naheliegendsten ist wahrscheinlich ein Algorithmus mit geringerer Laufzeit. Zwar verhindert die NP-Vollständigkeit aus Satz 4.4 eine effiziente Lösung – unter $\mathcal{P} \neq \mathcal{NP}$ – allerdings gibt es einige Möglichkeiten, das Lösen des Problems zu beschleunigen. Im Fall $K = 1$ ohne Trimmen, in dem durch die endliche Kandidatenmenge eine Rechenzeit von $O(n^3 \log n)$ erzielt wird, könnte sicherlich unter Ausnutzung der Konvexität (Lemma 3.12) und Anwendung von Verfahren, die auf der Methode des steilsten Abstiegs basieren, eine geringere Komplexität erreicht werden. Auch im Fall $K = 1$ mit Trimmen ist eine Verbesserung der Zeitkomplexität möglich. Durch ein sweep-line Verfahren kann sie eventuell auf $O(n^2 \log n)$ reduziert werden. Schwierigkeiten bei der Verbesserung treten naturgemäß bei der Platzierung mehrerer Geraden auf. Hier wäre es interessant, zu untersuchen, ob sich das Konzept der sweep-line Verfahren auf die Problemstellung für $K > 1$ übertragen lässt, um so wenigstens die Exponenten der Zeitkomplexität $O(n^{2K+1} \log n)$ ein wenig zu reduzieren. Zu klären wäre auch noch die Frage, ob die Platzierung mehrerer Geraden wie in dieser Arbeit einem TLE nach Neykov u. a. [24] entspricht oder nicht.

Eine weitere naheliegende Fragestellung ist die nach höheren Dimensionen. In dieser Arbeit wurde nur die einfache Regression in zwei Dimensionen behandelt, allerdings ist in der Praxis häufig mehr als eine erklärende Variable von Nöten und eine Beobachtung ist bei dieser multiplen Regression nicht mehr durch $z_i = (x_i, y_i) \in \mathbb{R}^2$ gegeben, sondern durch $z_i = (x_{i,1}, \dots, x_{i,p-1}, y_i) \in \mathbb{R}^p$. Auch für den p -dimensionalen Fall lässt sich eine duale Transformation analog zu Definition 3.6 angeben, die jetzt Punkten Hyperebenen zuordnet und umgekehrt. Es bezeichne

$$\ell(\beta_0, \beta_1, \dots, \beta_{p-1}) = \{(s_1, \dots, s_p) \in \mathbb{R}^p : s_p = \beta_0 + \beta_1 s_1 + \dots + \beta_{p-1} s_{p-1}\}$$

eine affine Hyperebene im \mathbb{R}^p , die nicht parallel zur p -ten Koordinatenachse ist, denn sie schneidet diese im Punkt $(0, \dots, 0, \beta_0)$, so dass β_0 wieder die Rolle des Achsenabschnitts aus dem zweidimensionalen Fall übernimmt. Weiter sei $z = (x_1, \dots, x_{p-1}, y)$ ein Punkt im \mathbb{R}^p , dann ist die duale Transformation gegeben durch

$$\begin{aligned} z^* &= \ell(-y, x_1, \dots, x_{p-1}) && \text{für } z \in \mathbb{R}^p \\ l^* &= (\beta_1, \dots, \beta_{p-1}, -\beta_0) && \text{für } l = \ell(\beta_0, \beta_1, \dots, \beta_{p-1}). \end{aligned}$$

Der vertikale Abstand wird dann zum Abstand in der p -ten Koordinate:

$$d_v(z, l) = |y - \beta_1 x_1 - \dots - \beta_{p-1} x_{p-1} - \beta_0|.$$

Dieser ist wieder unter der dualen Transformation erhalten:

$$d_v(l^*, z^*) = |-\beta_0 - \beta_1 x_1 - \dots - \beta_{p-1} x_{p-1} + y| = d_v(z, l).$$

Auch sonst hat man Eigenschaften der dualen Transformation wie in Lemma 3.8. Damit lassen sich analoge Ergebnisse zur Hyperebenenplatzierung wie zur Geradenplatzierung erzielen.

weak incidence property Eine endliche Kandidatenmenge für die Platzierung einer Hyperebene erhält man wie in Lemma 3.15 über die Aussage, dass es eine optimale Hyperebene gibt, die durch p affin unabhängige der gegebenen Punkte verläuft (vgl. Schöbel [30]). Dies könnte man vielleicht wieder auf die Platzierung mehrerer Hyperbenen verallgemeinern wie in Satz 4.7. Die Zeitkomplexität beim Durchprobieren der Kandidatenmenge erhöht sich durch die höhere Dimension allerdings auf $O(n^{pK+1} \log n)$.

pseudo-halving property Wie auf Seite 49 beschrieben, gilt die pseudo-halving property in beliebigen Dimensionen (vgl. Schöbel [30]).

Starker Punkt Ein starker Punkt wie in Satz 3.13 sollte sich auch für p Dimensionen herleiten lassen.

Außerdem sollte man auch die weak incidence property auf die Platzierung mehrerer Geraden und auch Hyperebenen durch norminduzierte und gerichtete Abstände verallgemeinern können.

Da a priori ja die Klassen der Beobachtungen unbekannt sind, ist insbesondere auch deren Anzahl K meist nicht bekannt. In dieser Arbeit wurde K als Parameter und nicht als Variable des Problems aufgefasst und musste somit vor dem Lösen angegeben werden. Gleiches gilt für die Anzahl m der Punkte, die zum Modell gehören. Wünschenswert wären Verfahren, die diese beiden Größen auch als Variablen auffassen und optimal bestimmen. Das birgt natürlich gewisse Schwierigkeiten, denn für $m = 0$ würden keine Punkte zur Zielfunktion beitragen, der optimale Zielfunktionswert wäre Null. Ebenso wird das Problem durch $K \geq \lceil m/2 \rceil$ trivial, da man dann immer je zwei Punkte mit einer eigenen Gerade ohne Fehler annähern könnte. Diese Fälle sind theoretisch so uninteressant wie praktisch irrelevant. Daher müssen geeignete Einschränkungen an K und m gestellt werden. Ein gutes K könnte durch die Verwendung eines Informationskriteriums zur Modellauswahl bestimmt werden. Die beiden bekanntesten sind das

Akaike information criterion [1] $AIC = \log L_{z_1, \dots, z_n}(\hat{\theta}_j) - P_j$ und das

Bayesian information criterion [32] $BIC = \log L_{z_1, \dots, z_n}(\hat{\theta}_j) - \frac{P_j}{2} \log n$,

wobei j verschiedene zur Auswahl stehende Modell indiziert, P_j die Zahl der freien Parameter eines Modells ist und $L_{z_1, \dots, z_n}(\hat{\theta}_j)$ der größte Wert der Likelihoodfunktion, der mit Modell j zu erreichen ist. Die Auswahl geschieht durch die Maximierung eines der Kriterien. Anschaulich bedeutet das eine Belohnung für große Likelihoodwerte, also für Modelle, die kleine Fehler machen, und gleichzeitig eine Strafe für Modelle mit vielen freien Parametern. In unserem Fall ist das eine Strafe für Modelle, die viele Geraden verwenden.

Der FDS-Algorithmus könnte auch angepasst werden, um Geraden zu platzieren, von denen einige Parameter bekannt sind. Zum Beispiel ist bei Ursprungsgeraden der Achsenabschnitt bekannt. In dem Fall wäre die Zielfunktion für eine Gerade im Dualraum auf eine Achsenparallele eingeschränkt und man kann auf dieser die Linearitätsbereiche als Geradenstücke indentifizieren (siehe Lemma 3.11). Dann würde möglicherweise ein Punkt aus den gegebenen genügen, um die Geraden endgültig festzulegen. Falls das auch auf die Platzierung mehrerer Geraden übertragen werden könnte, würde man damit die Komplexität auf $O(n^{2K-P+1} \log n)$ reduzieren, wenn P die Anzahl der bekannten Parameter ist (außer im trivialen Fall $P = 2K$).

Insgesamt ist die LAD-Methode durchaus konkurrenzfähig zur weiter verbreiteten LS-Methode, besonders, wenn von vornherein klar ist, dass ein robuster Schätzer benötigt wird. Insbesondere die Effizienz der beiden Methoden bietet vom theoretischen Standpunkt aus betrachtet kein ausreichendes Argument mehr für die LS-Methode. Die Platzierung einer Geraden ist aufgrund schneller Computer mit beiden Verfahren schnell durchzuführen und die Platzierung mehrerer Geraden ist in beiden Fällen NP-vollständig.

Abbildungsverzeichnis

1.1	Modellierungsfehler bei Platzierung einer Geraden	7
2.1	Dichte der Laplace-Verteilung	15
2.2	Laplace-Verteilung streut stärker als die Normalverteilung	17
2.3	LAD ist robuster als LSE	18
3.1	Nicht-Eindeutigkeit von (TLAD)	34
3.2	Beispiel für eine duale Parkettierung	38
3.3	Beispiel für eine duale Zelle	39
3.4	Die Zielfunktion von (TLAD) ist im Allgemeinen nicht quasikonvex.	43
3.5	Erläuterung zu Schritt 2 des Beweises von Satz 3.13	45
3.6	Erläuterung zu Schritt 3 des Beweises von Satz 3.13	46
5.1	Lineares Modell mit drei latenten Klassen, 250 Punkte	70
5.2	FastTLE-Schätzung 1 (oben) und 2 (unten) für das Modell aus Abbildung 5.1	72
5.3	Nicht-Identifizierbarkeit der latenten Klassen	76

Literaturverzeichnis

- [1] AKAIKE, H.: A New Look at the Statistical Identification Model. In: *IEEE Transactions on Automatic Control* 19 (1974), S. 716–723
- [2] BASSET, G. ; KOENKER, R.: Asymptotic Theory of Least Absolute Error Regression. In: *Journal of the American Statistical Association* 73 (1978), September, Nr. 363, S. 618–622
- [3] BERTSIMAS, D. ; SHIODA, R.: Classification and Regression via Integer Optimization. In: *Operations Research* 55 (2007), März–April, Nr. 2, S. 252–271
- [4] COOPER, L. ; STEINBERG, D. I.: *Introduction to Methods of Optimization*. 1. Auflage. Philadelphia : W. B. Saunders Company, 1970
- [5] DANTZIG, G. B.: *Linear Programming and Extensions*. Princeton : Princeton University Press, 1963
- [6] DEHLING, H. ; HAUPT, B.: *Einführung in die Wahrscheinlichkeitstheorie und Statistik*. 1. Auflage. Berlin : Springer-Verlag, 2003
- [7] DEMPSTER, A. P. ; LAIRD, N. M. ; RUBIN, D. B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. In: *Journal of the Royal Statistical Society. Series B (Methodological)* 39 (1977), Nr. 1, S. 1–38
- [8] FIEBIG, U.-R.: *Angewandte Statistik*. 2008. – Vorlesung an der Georg-August-Universität Göttingen
- [9] FORSTER, Otto: *Analysis 2*. 6. Auflage. Wiesbaden : Vieweg-Verlag, April 2005
- [10] HOARE, C. A. R.: Quicksort. In: *The Computer Journal* 5 (1962), Nr. 1, S. 10–16
- [11] HUBER, P. J.: Robust Regression: Asymptotics, Conjectures and Monte Carlo. In: *The Annals of Statistics* 1 (1973), Nr. 5, S. 799–821

- [12] HUBER, P. J. ; RONCHETTI, E. M.: *Robust Statistics*. 2. Auflage. John Wiley & Sons, 2009
- [13] KARMARKAR, N.: A New Polynomial-time Algorithm for Linear Programming. In: *Combinatorica* 4 (1984), Dezember, Nr. 4, S. 373–395
- [14] KARP, R. M.: *Reducibility Among Combinatorial Problems 1958–2008*. S. 219–242. In: JÜNGER, M. (Hrsg.) ; LIEBLING, T. (Hrsg.) ; NADDEF, D. (Hrsg.) ; NEMHAUSER, G. (Hrsg.) ; PULLEYBLANK, W. (Hrsg.) ; REINELT, G. (Hrsg.) ; RINALDI, G. (Hrsg.) ; WOLSEY, L. (Hrsg.): *50 Years of Integer Programming*. Berlin : Springer-Verlag, 2010
- [15] KHACHIYAN, L. G.: A Polynomial Algorithm in Linear Programming. In: *Soviet Mathematics Doklady* 20 (1979), Nr. 1, S. 191–194
- [16] KLEE, V. ; MINTY, G. L.: *How Good is the Simplex Algorithm?* S. 159–179. In: SHISHA, O. (Hrsg.): *Inequalities III*. New York : Academic Press, 1972
- [17] LAPLACE, P.-S.: *Théorie Analytique des Probabilités*. 3. Auflage. 1818
- [18] LEGENDRE, Adrien-Marie: *Nouvelles methodes pour la détermination des orbites des comètes*. Paris, 1805. – Quelle: Service Commun de la Documentation University of Strasbourg - Digital old books, URL <http://num-scd-ulp.u-strasbg.fr:8080/327/>
- [19] LEISCH, F.: FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R. In: *Journal of Statistical Software* 11 (2004), Oktober, Nr. 8, S. 1–18
- [20] LOZANO, A. J. ; PLASTRIA, F.: The Ordered Median Euclidean Straight-line Location Problem. In: *Studies in Locational Analysis* (2009), September, Nr. 17, S. 29–43
- [21] MEGIDDO, N.: Towards a Genuinely Polynomial Algorithm for Linear Programming. In: *SIAM Journal on Computing* 12 (1983), Mai, Nr. 2, S. 347–353
- [22] MEGIDDO, N.: Linear Programming in Linear Time When the Dimension Is Fixed. In: *Journal of the ACM* 31 (1984), Januar, Nr. 1, S. 114–127
- [23] MEGIDDO, N. ; TAMIR, A.: On the Complexity of Locating Linear Facilities in the Plane. In: *Operations Research Letters* 1 (1982), November, Nr. 5, S. 194–197

- [24] NEYKOV, N. ; FILZMOSE, P. ; DIMOVA, R. ; NEYTCHEV, P.: Robust Fitting of Mixtures Using the Trimmed Likelihood Estimator. In: *Computational Statistics & Data Analysis* 52 (2007), September, Nr. 1, S. 299–308
- [25] NICKEL, S. ; PUERTO, J.: *Location Theory. A Unified Approach*. 1. Auflage. Berlin : Springer-Verlag, Mai 2005
- [26] POLLARD, D.: Asymptotics for Least Absolute Deviation Regression Estimators. In: *Econometric Theory* 7 (1991), Juni, Nr. 2, S. 186–199
- [27] R DEVELOPMENT CORE TEAM: *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (Veranst.), 2009. – URL <http://www.R-project.org>. – ISBN 3-900051-07-0
- [28] ROUSSEEUW, P. J. ; LEROY, A. M.: *Robust Regression and Outlier Detection*. John Wiley & Sons, 2003
- [29] SCHABACK, R. ; WENDLAND, H.: *Numerische Mathematik*. 5. Auflage. Berlin : Springer-Verlag, 2005
- [30] SCHÖBEL, A.: *Locating Lines and Hyperplanes: Theory and Algorithms*. 1. Auflage. Dordrecht, Niederlande : Kluwer Academic Publishers, 1999
- [31] SCHÖBEL, A.: *Robuste Optimierung*. 2009/2010. – Vorlesung an der Georg-August-Universität Göttingen
- [32] SCHWARZ, Gideon: Estimating the Dimension of a Model. In: *The Annals of Statistics* 6 (1978), Nr. 2, S. 461–464
- [33] STIGLER, S. M.: Simon Newcomb, Percy Daniell, and the History of Robust Estimation 1885–1920. In: *Journal of the American Statistical Association* 68 (1973), Dezember, Nr. 344, S. 872–879

Danksagung

Ich bedanke mich für die Hilfe und vielen Anregungen bei der Erstellung dieser Arbeit bei der Arbeitsgruppe Optimierung des Instituts für Numerische und Angewandte Mathematik der Universität Göttingen, insbesondere bei meiner Betreuerin Prof. Anita Schöbel und Mark-Christoph Körner.

Außerdem danke ich meinem Zweitkorrektor Prof. Martin Schlather vom Institut für Mathematische Stochastik der Universität Göttingen, der sich schon während der Erstellung der Arbeit Zeit zur Beantwortung von Fragen genommen hat.

Schließlich geht mein besonderer Dank an Neyko Neykov vom Nationalen Institut für Meteorologie und Hydrologie Bulgarien, der mir den Quelltext des FastTLE-Algorithmus [24] zur Verfügung gestellt hat.