‘Kleinste Quadrate’, Thema und Variationen

Angewandte Optimierungsaufgaben in der Statistik

14. August 2005

Zusammenfassung

Zufällige Messfehler können im Prinzip jede Messung verfälschen. Daher enthalten wiederholte Messungen für eine einzige unbekannte Grösse in der Regel Widersprüche. Diese lassen sich mit verschiedenen Massen erfassen. Jedes derartige Mass (technisch eine Metrik) für Widersprüche führt zu einer eigenen ‘optimalen’ Schätzmethode, mit welcher sich der Datensatz so bereinigen lässt, dass der Widerspruch – gemessen in der jeweiligen Metrik – minimiert wird. Gängige Mittelbildungen lassen sich dann als Minimierungsaufgaben interpretieren. Das arithmetische Mittel minimiert die mittlere quadratische Abweichung zwischen der Schätzung und den Messdaten. Weitere ‘Mittelbildungen’ werden in dieser Sichtweise untersucht. Der Text ist begleitet von Lernaufgaben, in denen verschiedene Punktschätzungen als Minimierungsprobleme erscheinen. Etwas Differentialrechnung wird vorausgesetzt. Sie liefert jedoch nur eines von mehreren Mitteln um Minima zu finden.

1 Einleitung

Euklid hat bewiesen, dass die Winkelsumme in einem ebenen Dreieck 180° beträgt. Wer jedoch diese Aussage durch genaue Messungen prüft, findet, dass sie in Wirklichkeit kaum je exakt stimmt. Ist die Geometrie Euklids zu einfach angelegt, wenn es um sehr genaue Messungen geht? Handelt es sich um eine Idealisierung, die in der Wirklichkeit nicht besteht? Oder sind etwa alle unsere Messungen unzuverlässig?

Es ist offenbar bequem, die Geometrie in einer einfachen und idealisierten Form zu behalten und mit der Annahme zu leben, dass jede Messung durch Fehler beeinträchtigt sein kann. Folglich müssen wir uns darum kümmern, Widersprüche zwischen Theorien und Messergebnissen zu beseitigen. Für alle empirischen Wissenschaften ist diese Vereinfachung durch gegenseitige Anpassung der Daten und der theoretischen Voraussagen von fundamentaler Bedeutung. Weil die Gefahr einer willkürlichen Datenmanipulation besteht, sind ‘objektive’ Auswerteverfahren gesucht. Beliebt sind Algorithmen, welche die Datenreduktion ohne Einflussnahme durch den Auswerter besorgen. Legendre und Gauss haben als erste die Methode der kleinsten Quadrate zu diesem Zweck entwickelt und benutzt. Im einfachsten Fall führt dieses Verfahren auf das arithmetische Mittel. Gauss hat abgeklärt, unter welchen Annahmen das arithmetische Mittel eine ‘optimale’ Datenreduktion darstellt.

Es wird im folgenden ein Projekt zum Thema ‘optimale’ Schätzung und ‘objektive’ Datenauswertung im Rahmen des Mathematikunterrichtes der Mittelschule umrissen. Skelettartig sind jene Fragen im Text eingebaut, die den Schülern zur Bearbeitung vorgelegt werden und die den Gang des Projektes bestimmen. Diese Art des Unterrichtes strebt mehrere ineinander verschachtelte Ziele an: Beim Lösen der Aufgaben werden die Schüler vorerst Kenntnisse aus der Analysis vertiefen. Sodann führen die betrachteten Beispiele zu Beobachtungen und Vermutungen; die Schüler werden entdecken und beweisen, dass verschiedene weitere, in der beschreibenden Statistik geläufige Verfahren zur Datenreduktion aus einem einheitlichen *Extremalprinzip* herleitbar sind. Die einzelnen Verfahren unterscheiden sich in dieser Sichtweise bloss in der Art, wie der Widerspruch zwischen der Datenmenge und der idealen Voraussage gemessen wird. Wir beschränken uns in einem ersten Schritt auf Punktschätzungen auf der Zahlgeraden, da hier die Mittelschulanalysis ausreicht. Wir zeigen aber anschliessend auch das Bedürfnis für eine Weiterentwicklung dieser Methoden. In einem zweiten Schritt werden Punktschätzungen in der Ebene oder allgemeiner in mehrdimensionalen Räumen betrachtet. Hier sind wir auf geometrische berlegungen angewiesen, da wir nicht annehmen, dass partielle Ableitungen eingesetzt werden. In der Folge müssen wir uns mit einigen Beispielen und Ausblicken begnügen und mit der Hoffnung, dass die Aussicht auf Neuland einige Schüler zu eigenen weitergehenden Untersuchungen und Entdeckungen anrege. Dieser Bericht enthält darum wohl mehr Aufgaben, als für den Klassenunterricht mit Rücksicht auf die verfügbare Zeit nötig sind. Beim Einsatz im Unterricht kann und soll von den zahlreichen Möglichkeiten bezüglich der Auswahl einzelner Themen oder der Arbeitsformen Gebrauch gemacht werden. Gute Erfahrungen liegen vor mit der Bearbeitung einzelner Verfahren nach der Puzzlemethode, in einer Werkstatt oder als Facharbeiten.

Zu den inhaltlichen und methodischen Voraussetzungen gehören im ersten Teil mindestens die Grunderfahrungen im Umgang mit Extremalproblemen im Rahmen der Analysis. Im zweiten Teil wird angenommen, dass die geometrische Lösung von Extremalproblemen im Zusammenhang mit Normalenproblemen bekannt sei. Ferner tritt das Skalarprodukt auf. ber weitergehende spezielle Anforderungen bei einzelnen Aufgaben sollen die grob skizzierten Lösungen und die dort eingeschobenen Literaturhinweise Auskunft geben. Dieser Bericht enthält keine Musterlösungen, die sich für Korrekturen beim Selbststudium eignen. Dazu sind einige Aufgaben zu anspruchsvoll. Es wird erwartet, dass der Schüler üblicherweise durch einen Fachmann beraten oder betreut werden muss. Zudem werden sich gemäss dem jeweiligen Unterrichtsstil und den persönlichen Eigenheiten der Lehrer und Schüler wohl erhebliche Varianten beim Lösen der Aufgaben ergeben.

Kenner werden bemerken, dass wir mit den allgemeinen Fehlermassen implizit Beispiele zum Begriff des allgemeinen metrischen Raumes behandeln. Natürlich braucht es dazu keine Theorie. Wir versuchen, Begriffsbildungen, welche für die Mathematik des zwanzigsten Jahrhunderts etwa in der Funktionalanalysis wichtig und nützlich geworden sind, im Rahmen der Schulmathematik anzudeuten und mit der elementaren beschreibenden Statistik in Beziehung zu setzen. In einem ähnlichen Sinn sollte die Verwendung eines Extremalprinzips bewertet werden. Die Mittelschulmathematik pflegt zwar die Behandlung von Extremalaufgaben aller Art, sie zeigt aber kaum die Perspektive bis hin zur Variationsrechnung und den Extremalprinzipien, die zu den Grundlagen physikalischer Theorien gehören. Extremalprinzipien waren in der Physik so erfolgreich, dass sie als Denkmuster in anderen Gebieten übernommen wurden, etwa von Darwin. In seiner Evolutionstheorie sorgt die Selektion für das Überleben der geeignetsten Arten. Anderseits werden die Extremalprinzipien in der Datenauswertung benutzt, um bei gegebenen Daten jene Schätzungen auszulesen, die den Widerspruch zum Datensatz minimieren, also für das Überleben der zugrundegelegten Theorie am geeignetsten sind.

2 Das Thema, Methode der kleinsten Quadrate

Angenommen, eine unbekannte Grösse x soll gemessen werden. Es wird also ein Experiment angestellt, das einen Messwert m_1 liefert. Wenn wir der Messung voll vertrauen, werden wir behaupten $x = m_1$. Andere Messungen werden folgen und Ergebnisse $m_2, \ldots, m_n$ liefern. Stimmen alle Messungen überein, so wird unser Vertrauen gestärkt und wir bleiben bei der Behauptung $x = m_1$. Die Erfahrung zeigt aber, dass die angenommene bereinstimmung gerade in den interessantesten Fällen nicht eintritt. Statt dessen zeigen sich *Widersprüche* in den Ergebnissen der wiederholt und unabhängig unter 'gleichen' Bedingungen ausgeführten Messungen.

Legendre und Gauss haben vorgeschlagen, den Widerspruch in den Daten $m_1, \ldots, m_n$ nach folgendem *Extremalprinzip* zu beseitigen:
Wir betrachten eine beliebige reelle Zahl t als Schätzung für die unbekannte Konstante x. Als *'optimale' Schätzung* einigen wir uns auf jene Zahl $\hat{t}$, für welche die Summe

$$\sum_{i=1}^{n} (m_i - t)^2$$

den kleinsten Wert annimmt.

Aufgaben

2.1 Welches ist der Wert $\hat{t}$, den die Methode der kleinsten Quadrate zu den Daten $m_1, \ldots, m_n$ liefert ?

2.2 Angenommen, wir haben drei Messdaten m_1, m_2, m_3 für die unbekannte Grösse x. Deuten Sie den Ausdruck $\sum_{i=1}^{3} (m_i - t)^2$ im Sinne der Raumgeometrie. Welche Bedeutung hat die optimale Schätzung $\hat{t}$ in diesem Rahmen ?

2.3 Ein Versuch mit einer Alternative: 'Methode der kleinsten Biquadrate'.
Angenommen, die Messdaten seien $m_1 = 0, m_2 = 1, m_3 = 4$. Für welche Zahlen $\hat{t}$ wird der Ausdruck

$$\sum_{i=1}^{3} (m_i - t)^4$$

minimal ? Ist die Antwort eindeutig ? Gibt es Gründe, die Methode der kleinsten Quadrate der hier besprochenen Variante vorzuziehen ?

2.4 Welche Gründe könnten Legendre und Gauss dazu geführt haben, die Abweichung der Schätzung t von einem Einzelergebnis m_i durch $(m_i - t)^2$ zu beschreiben ?

Lösungen

2.1 arithmetisches Mittel der Daten

2.2 Die Summe der quadratischen Abweichungen entspricht dem Quadrat des Abstandes der Punkte $\mathbf{T} : (t, t, t)$ und $\mathbf{M} : (m_1, m_2, m_3)$, vorausgesetzt, dass sich die Koordinaten auf ein orthonormierte Koordinatensystem beziehen und die Abstände im Sinne der gewöhnlichen Geometrie gemessen werden.
Lösung ist der Lotfusspunkt zu $\mathbf{M}$ auf der Geraden, welche $\mathbf{T}$ durchläuft, wenn t variiert.

2.3 Numerische Bestimmung: $\hat{t}_4 \approx 1.9(63)$. Der Rechenaufwand ist erheblich grösser als beim arithmetischen Mittel. Das Minimum ist eindeutig, da δ^4 eine konvexe Funktion ist.

2.4 (a) Das Mass für den Fehler soll nur von $\mid m_i - t \mid$ abhängen und unabhängig sein von der Reihenfolge der Messungen.

(b) Verschiedene Fehler sollen sich im Fehlermass nicht gegenseitig aufheben. Dies wird zum Beispiel erreicht, indem die einzelnen Summanden nie negative Werte annehmen.

(c) Die gewählte Variante ist die 'einfachste', welche die beiden obigen Forderungen erfüllt. Sie lässt sich geometrisch rechtfertigen. Gauss hat die Methode in der Ausgleichsrechnung ausgebaut, um redundante Messdaten mittels Linearkombinationen von vorgebenen Funktionen 'optimal' an eine Theorie anzupassen. (Für uns war die theoretische Annahme, dass x eine Konstante sei.) Bei der Auswertung sind typischerweise nur lineare Gleichungen zu lösen.

(d) Das Ergebnis der Methode rechtfertigt die verbreitete Anwendung des arithmetischen Mittels zur Datenreduktion durch ein Extremalprinzip. (Weitere Rechtfertigungen hat Gauss gegeben unter der Annahme, dass Messfehler annähernd normal verteilt seien.)

3 Variationen zum Thema

Nach den bisherigen Überlegungen ist die Stärke der Methode der kleinsten Quadrate in der Einfachheit des Verfahrens zu finden. Die Antwort auf die Frage nach der 'besten Schätzung' kann nach den Konventionen dieser Methode elementar und ohne Kenntnis der Infinitesimalrechnung gefunden werden. Der Algorithmus ist einfach und effizient. Ein weiterer und noch wichtigerer Vorteil hängt mit dieser 'Einfachheit' zusammen: Die Methode ist so flexibel, dass sie auf weit kompliziertere Fälle übertragen werden kann: Anpassung von Funktionen eines vorgegebenen Typs an Messdaten, Anpassung ganzer Theorien an Messdaten. Sie führt zu einem Algorithmus, in dem bloss lineare Gleichungssysteme zu behandeln sind. Dies ist für die Anwendungen von grosser Bedeutung.

Diese Eigenschaften sind zwar gewichtig, sie verbieten aber keineswegs Versuche mit Variationen der Grundidee. Deshalb werden wir hier ein allgemeines *Extremalprinzip* für die Datenreduktion formulieren, in welchem ein noch zu bestimmendes *Fehlermass* auftritt. Beispiele zu Fehlermassen werden im folgenden angegeben und in den zugehörigen Aufgaben untersucht. Aus der Geometrie und der Analysis sind Fälle bekannt, in denen die vertraute Abstandsdefinition aus der Euklidischen Geometrie mit Erfolg durch andere Varianten ersetzt wurden. Wir halten uns an diese Vorbilder und experimentieren mit verschiedenen Fehlermassen. Dabei werden sich verschiedene 'Mittelbildungen' aus der beschreibenden Statistik als 'optimale' Schätzung erweisen, wenn nur das 'richtige' Fehlermass verwendet wird.

Ein allgemeines Extremalprinzip zur Reduktion von Messdaten

Um eine unbekannte Zahl x zu bestimmen, wurden n Messungen ausgeführt. Die Ergebnisse bezeichnen wir mit $\mathbf{M} = \{m_1, \ldots, m_n\}$. Wählen wir eine beliebige reelle Zahl t als Schätzung für x, so ergibt sich ein Widerspruch zu den experimentellen Daten $\mathbf{M}$, es sei denn, alle Messungen führen zum gleichen Ergebnis und wir wählen dieses als Schätzung für x. Wir suchen nun zuerst eine Funktion

$$\begin{array}{rccl} \delta_{\mathbf{M}}: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & t & \mapsto & \delta_{\mathbf{M}}(t) \end{array}$$

aus, welche diesen Widerspruch quantitativ zu erfassen gestattet. Eine solche Funktion werden wir *Fehlermass* nennen. Wir setzen voraus, dass jedes Fehlermass ein globales Minimum besitzt. Eine formale Charakterisierung der Fehlermasse werden wir erst nach der Untersuchung von Beispielen wagen (Aufgabe 3.9).

Als Verallgemeinerung der Summe der quadratischen Abweichungen betrachten wir zunächst Summen der Art

$$\sum_{i=1}^{n} f(\mid m_i - t \mid),$$

wobei die Funktion f noch genauer festgelegt wird.

Definitionen

- Es sei $r \geq 1$: Das Fehlermass $\delta^r_{\mathbf{M}}$ ist eine Funktion

  $$\begin{array}{rccl} \delta^r_{\mathbf{M}}: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & t & \mapsto & \sum_{i=1}^{n} \mid m_i - t \mid^r \end{array}$$

  Bemerkung: Der Fall $\delta^2_{\mathbf{M}}$ führt zur Methode der kleinsten Quadrate.

- Das Fehlermass $\delta^\infty_{\mathbf{M}}$ ist definiert als Funktion

  $$\delta^\infty_{\mathbf{M}} : t \mapsto \max{}_{1 \leq i \leq n}\{\mid m_i - t \mid\}.$$

- Das Fehlermass $\delta^*_{\mathbf{M}}(t)$ zählt für jeden Wert m_i der Datenmenge einen Strafpunkt, falls m_i nicht mit t übereinstimmt (Hamming-Distanz). Für eine formale Definition lässt sich

  $$\delta^*_{\mathbf{M}} : t \mapsto \sum_{i=1}^{n} \operatorname{sgn}(|m_i - t|)$$

  verwenden.

Punktschätzung mit Extremalprinzip

Frage: Um eine unbekannte Zahl x zu bestimmen, wurden die Zahlen $\{m_1, \ldots, m_n\}$ gemessen. Welche Zahl $\hat{x}$ schätzt x mit Berücksichtigung der Menge $\mathbf{M}$ der Messergebnisse am besten? (Punktschätzung)
Antwort: Eine 'optimale' Schätzung hängt noch von der Wahl einer Fehlerfunktion $\delta_{\mathbf{M}}$ ab. Die globalen Minimalstellen von $\delta_{\mathbf{M}}$ sind die gesuchten Punktschätzungen, welche $\mathbf{M}$ *und* $\delta_{\mathbf{M}}$ Rechnung tragen.

Zur Notation: Falls die zugrundegelegte Datenmenge im Zusammenhang klar ist, werden wir die einfachere Notation δ^r statt $\delta^r_{\mathbf{M}}$ verwenden. Da wir den Einfluss verschiedener Fehlermasse auf die 'optimale' Schätzung bei gleichen Datenmengen $\mathbf{M}$ untersuchen wollen, ist es nützlich, die mit δ^r bestimmte 'optimale' Schätzung von nun an mit $\hat{t}_r$ zu bezeichnen.

Aufgaben

3.1 Es sei $m_1 = 0$ der einzige Messwert. Skizzieren Sie die Graphen der folgenden Fehlermasse:
$\delta^1, \delta^2, \delta^4, \delta^\infty, \delta^*$.

3.2 Es sei $\mathbf{M} = \{0, 1\}$ die Datenmenge. Skizzieren Sie die Graphen der Fehlermasse δ^1, δ^2, δ^4, δ^∞, δ^*. Was fällt auf ? Welches sind in den einzelnen Fällen die 'optimalen' Schätzungen $\hat{t}_r$?

3.3 Es sei $\mathbf{M} = \{0, 1, 4\}$ die Datenmenge. Skizzieren Sie die Graphen der Fehlermasse δ^1, δ^2, δ^4, δ^∞, δ^*. Was fällt auf ? Welches sind in den einzelnen Fällen die 'optimalen' Schätzungen ?

3.4 Es sei $\mathbf{M} = \{0, 1, 1, 4\}$ die Datenmenge. Skizzieren Sie die Graphen der Fehlermasse δ^1, δ^2, δ^4, δ^∞, δ^*. Was fällt auf ? Welches sind in den einzelnen Fällen die 'optimalen' Schätzungen $\hat{t}_r$?

3.5 Es sei $\mathbf{M} = \{m_1, \ldots, m_n\}$ eine allgemeine, nichtleere Datenmenge. Begründen oder widerlegen Sie:

- Für jede gerade natürliche Zahl $k \geq 2$ hat die Funktion δ^k genau ein Minimum.
- Für jedes reelle $r > 1$ ist die Funktion δ^r mindestens einmal stetig differenzierbar.
- Für jede reelle Zahl $r \geq 2$ besitzt die Funktion δ^r genau ein Minimum.

3.6 Es sei $\mathbf{M} = \{m_1, \ldots, m_n\}$ eine allgemeine, nichtleere Datenmenge. Geben Sie für δ^1, δ^2, δ^4, δ^∞ ,δ^* je einen Algorithmus an, der die 'optimale' Schätzung $\hat{t}_r$ liefert. Was ist bemerkenswert an den Ergebnissen ?

3.7 Nennen Sie Beispiele von Datenmengen und allgemeine beschreibende Eigenschaften der Messdaten, die dazu führen, dass alle Werte $\hat{t}_r$ der 'optimalen' Schätzungen übereinstimmen.

3.8 Die Datenmenge $\mathbf{M}$ bestehe aus lauter positiven Zahlen $m_1, \ldots, m_n$. Für welches Fehlermass γ, das für positive Zahlen t definiert sein soll, ist das *geometrische Mittel* eine 'optimale' Schätzung ?

3.9 Welche Eigenschaften soll eine Funktion δ besitzen, damit sie als Fehlermass bezeichnet werden kann ?

Lösungen

3.2 Für alle r ist $\hat{t}_r = 0.5$.

3.3 $\hat{t}_1 = 1$ $\quad \hat{t}_2 = 5/3$ $\quad \hat{t}_4 \approx 1.9(63)$ $\quad \hat{t}_\infty = 2$ $\quad \hat{t}_* = \mathbf{M}$

3.4 $\hat{t}_1 = 1$ $\quad \hat{t}_2 = 3/2$ $\quad \hat{t}_4 \approx 1.9(32)$ $\quad \hat{t}_\infty = 2$ $\quad \hat{t}_* = 1$

3.5
- Die zweite Ableitung der Fehlerfunktion ist stetig und überall positiv.
- Es genügt, von der Funktion $x \mapsto \mid x \mid^r$ zu zeigen, dass sie in 0 stetig differenzierbar ist.

- Die zweite Ableitung der Fehlerfunktion ist stetig und überall positiv.
  Bemerkung: Wesentlich an dieser Begründung ist, dass δ^r konvex ist. Dies gilt sogar für $r = 1$, allerdings ist dann nur noch der Minimalwert, nicht jedoch die Minimalstelle stets eindeutig.

3.6 (a) $\hat{t}_1$ ist der Median der Datenmenge **M**.

(b) $\hat{t}_2$ ist das arithmetische Mittel der Datenmenge.

(c) $\hat{t}_4$ kann als einzige Nullstelle der Ableitung der Fehlerfunktion (Polynomfunktion vom Grade 3) numerisch bestimmt werden. Als Schätzwert eignet sich das arithmetische Mittel zwischen dem kleinsten und dem grössten Messwert.

(d) $\hat{t}_\infty$ ist das arithmetische Mittel zwischen dem kleinsten und dem grössten Messwert. Alle andern Messwerte haben keinen Einfluss auf $\hat{t}_\infty$.

(e) $\hat{t}_*$ ist der häufigste Messwert der Datenmenge ('Modus').

3.7 Symmetrische Anordnung der Messwerte bezüglich dem häufigsten Messwert. (Spezialfall: alle Messungen ergeben denselben Wert).

3.8 Setzt man für die erste Ableitung

$$\gamma'(t) = \sum_{i=1}^{n} \ln \frac{t}{m_i} \ ,$$

so wird die zweite Ableitung γ'' überall positiv und die einzige Nullstelle $\hat{t}_\gamma$ der ersten Ableitung γ' erfüllt

$$\ln(\hat{t}_\gamma) = \frac{1}{n} \sum_{i=1}^{n} \ln(m_i) \ ,$$

also ist $\hat{t}_\gamma$ das geometrische Mittel der Messwerte. Durch Integration wird γ aus γ' erhalten, die Wahl der Integrationskonstante hat auf die Lage von $\hat{t}_\gamma$ keinen Einfluss. Man wird aus Gründen der Anschaulichkeit die Integrationskonstante so wählen, dass $\gamma(t) \geq 0$ für t zwischen dem kleinsten und dem grössten Messwert erfüllt ist.

$$\gamma(t) = C - nt + \sum_{i=1}^{n} t \cdot \ln(\frac{t}{m_i})$$

3.9
- $\delta(t) \geq 0$ für alle t.
- δ hat ein globales Minimum.
- ist $\mathbf{M} = \{m, m, \ldots, m\}$, so ist $\hat{t} = m$ eindeutig.
- δ ist konvex.

4 Punktschätzungen in der Ebene

In diesem Abschnitt überschreiten wir eine Grenze bezüglich der Probleme, die der traditionellen Mittelschulanalysis zugerechnet werden. Wir betrachten Punkte in der Ebene als Daten. Folglich hängen die Fehlerfunktionen von zwei Variablen ab, es treten ganz natürlich Extremalprobleme für Funktionen mehrerer Variablen auf. Obwohl unsere Fragen eine starke Motiovation zur Erweiterung der Analysis in dieser Richtung sind, bleiben wir bei den Methoden der Schulmathematik. Die betreffenden Extremalaufgaben werden wir in den vorgelegten Spezialfällen geometrisch behandeln. Von Fall zu Fall benötigen wir eine andere Methode. Falls partielle Ableitungen verfügbar sind, lassen sich einige der Aufgaben nach einem einheitlichen Schema lösen. Die Beispiele zeigen, wie ausgezeichnete Punkte in Dreiecken mittels Extremaleigenschaften charakterisiert werden können.

Aufgaben

Angenommen, ich will meinen Standort in einem ebenen Gelände anhand einer Landkarte bestimmen. Ohne einen guten Kartenpunkt in der Nähe kann das ein heikles Problem sein. Ich versuche es und finde einen Punkt P_1, der mir plausibel erscheint. Zur Kontrolle wiederhole ich die Ortsbestimmung nochmals. Das Ergebnis ist ein Punkt P_2, der nicht mit dem ersten übereinstimmt. Ein dritter Versuch ist nötig. Er führt zu einem Punkt P_3. Ich entschliesse mich, den Schwerpunkt im Dreieck $P_1P_2P_3$ als meinen Standort zu postulieren. Lässt sich diese Wahl rechtfertigen ?

4.1 Welche Fehlermasse für Punkte in der Ebene verallgemeinern die Definitionen von $\delta^r_{\mathbf{M}}$ aus dem letzten Abschnitt ?

4.2 Gegeben sind drei Punkte P_1, P_2, P_3 der Ebene. Für welchen Punkt $\hat{P}$ der Ebene ist:

- (a) die Summe der Abstände zu P_1, P_2, P_3 minimal ?
- (b) die Summe der Quadrate der Abstände zu P_1, P_2, P_3 minimal ?
- (c) die Summe der Abstände zu den Seiten des Dreiecks $P_1P_2P_3$ minimal ?
- (d) die Summe Quadrate der Abstände zu den Seiten des Dreiecks $P_1P_2P_3$ minimal ?
- (e) das Maximum der Abstände zu P_1, P_2, P_3 minimal ?
- (f) das Maximum der Abstände zu den Seiten im Dreieck $P_1P_2P_3$ minimal ?

4.3 Welche der in Aufgabe 4.2 gefundenen Lösungen lassen sich übertragen auf den Fall, dass

- (a) in der Ebene n Punkte gegeben sind ?
- (b) vier Punkte im Raum bestimmt wurden ?
- (c) im Raum n Punkte gegeben sind ?

4.4 Werden die 'optimalen' Schätzungen von 4.2 auf die Daten der Ortsbestimmung angewendet, so ist denkbar, dass das Optimierungskriterium dem Verfahren zur Standortbestimmung angepasst werden kann.
Welches Fehlermass aus Aufgabe 4.2 könnte den folgenden Verfahren besonders angemessen sein ?

- (a) reines Kartenlesen

(b) Rückwärtseinschneiden
Mit einer Bussole wurde vom unbekannten Standort aus die Richtung auf je drei gute Kartenpunkte gemessen. Der Standort wäre dann im gemeinsamen Schnittpunkt der drei Standlinien vorgegebener Richtung durch die drei Kartenpunkte zu finden, wenn die Winkelmessung fehlerfrei erfolgte.

(c) Bogenschnittmethode
Vom unbekannten Standort aus werden die Entfernungen zu drei markanten Kartenpunkten gemessen. Der Standort würde sich bei ideal genauer Messung im gemeinsamen Schnittpunkt der drei Distanzkreise um die Kartenpunkte befinden.

Lösungen

4.1 (a) allgemeine Konstruktion:

$$\Delta^r_{\mathbf{M}}(x,y) = \delta^r_{\pi_1\mathbf{M}}(x) + \delta^r_{\pi_2\mathbf{M}}(y) \ ,$$

wobei π_i die Projektion auf die $i-$te Koordinate bezeichnet. Mit dieser Konstruktion lassen sich alle Ergebnisse aus Abschnitt 3 leicht auf höhere Dimensionen übertragen, denn es gilt:
Die Koordinaten des 'optimalen Punktes' $\hat{P}$ bezüglich $\Delta^r_{\mathbf{M}}$ lassen sich einzeln mit $\delta^r_{\pi_i\mathbf{M}}$ bestimmen. Damit wird $(\hat{x}_r \mid \hat{y}_r)$ die optimale Schätzung für die Koordinaten des gesuchten Punktes.

Vorteil dieser Definition ist die leichte Übertragbarkeit der früheren Ergebnisse und Erfahrungen auf höhere Dimensionen. Dem steht ein allenfalls gewichtiger Nachteil gegenüber. Zu gegebenen Daten hängt die Schätzung $\hat{P}$ von der allenfalls noch willkürlich wählbaren Richtung der Koordinatenachsen ab. Dies ist jeweils dann möglich, wenn kein 'natürliches' Koordinatensystem vorgegeben ist, beispielsweise bei geometrischen Daten wie Punkten in der Ebene im Unterschied zu Zahlenpaaren der Form (t, x), wo x eine Ortskoordinate und t eine Zeit bezeichnen.

(b) Die Summe der Abstände von P zu jedem Datenpunkt.

(c) Die Summe der Quadrate der Abstände von P zu jedem Datenpunkt. Dies entspricht gerade dem Fehlermass Δ^2 aus der allgemeinen Konstruktion.

(d) Die Summe der $r-$ten Potenzen der Abstände von P zu jedem Datenpunkt.

(e) Das Maximum der Abstände von P zu jedem Datenpunkt.

(f) Die Anzahl der Datenpunkte, die nicht mit P übereinstimmen.

(g) Weitere Varianten: Statt Abstände zu den Datenpunkten können die Flächeninhalte der Dreiecke der Form P_iP_kP verwendet werden oder die Abstände zu Verbindungsgeraden verschiedener Datenpunkte. Dieses Thema wird hier nicht weiter verfolgt, obwohl es sich für Schüleraufgaben und eigene Erkundungen im mathematischen Neuland eignet.

4.2 (a) Der 'Steinerpunkt' Σ des Dreiecks $P_1P_2P_3$, das ist der Punkt, von dem aus je zwei der drei Datenpunkte unter einem 120°-Winkel gesehen werden, beziehungsweise die Ecke des Dreiecks mit dem grössten Winkel, falls dieser mindestens 120° misst.[vgl. Courant, Robbins, Was ist Mathematik? Kap. VII, §5].

(b) Der Schwerpunkt des Dreiecks.

(c) Es sei x_i der Abstand des Punktes P von der Seite p_i im Dreiseit $p_1p_2p_3$. Liegt P im Innern dieses Dreiseits, so gilt für dessen Flächeninhalt F

$$F = \frac{1}{2}\sum_{i=1}^{3} x_i \mid p_i \mid ,$$

wobei $\mid p_i \mid$ die Länge der Seite p_i bezeichnet. Ferner ist $x_i \geq 0$. Es entsteht ein *lineares Optimierungsproblem* mit der Zielfunktion $\sum_{i=1}^{3} x_i$, die zu minimieren ist. Lösungspunkt ist jene Ecke des Dreiecks $P_1P_2P_3$, die der längsten Seite gegenüber liegt.

(d) Mit den gleichen Bezeichnungen wie in der vorangehenden Lösung ist $\sum_{i=1}^{3} x_i^2$ zu minimieren unter der Nebenbedingung

$$F = \frac{1}{2}\sum_{i=1}^{3} x_i \mid p_i \mid$$

und den Einschränkungen $x_i \geq 0$. Die Nebenbedingung und die Einschränkungen beschreiben eine Dreiecksfläche $\mathcal{D}$ im ersten Oktanten. Lösung ist der Punkt $H : (x_1, x_2, x_3)$ in $\mathcal{D}$, der am nächsten bei $\mathbf{0}$ liegt. Gleichwertige Beschreibungen von H sind:

- der Lotfusspunkt zu $\mathbf{0}$ im Dreieck $\mathcal{D}$.
- der Höhenschnittpunkt im Dreieck $\mathcal{D}$.

(e) Der Umkreismittelpunkt.

(f) Der Inkreismittelpunkt.

4.3 (a) Schwerpunkt, Methode (4.2,b)

(b) Methode (4.2,b,c,d,e,f)

(c) Schwerpunkt, Methode (4.2,b)

4.4 (a) Methode (4.2,a,b oder e)

(b) Methode (4.2,c,d)

(c) Methode (4.2,e evtl. a,b)

5 Diskussions- und Aufsatzthemen

- Warum stimmen die Ergebnisse verschiedener 'optimaler' Schätzungen bei gleichen Daten nicht notwendigerweise überein?
- Warum stimmen die Ergebnisse verschiedener 'objektiver' Schätzungen bei gleichen Daten nicht notwendigerweise überein?
- Gibt es objektive Kriterien für die Wahl von Fehlermassen?