

Die Methode der maximalen Untergruppe (MSS) und ihre Anwendung in der Kongruenzuntersuchung geodätischer Netze

Frank Neitzel

Zusammenfassung

Es wird eine neue Methode der geodätischen Datenanalyse namens MSS - Die Methode der maximalen Untergruppe vorgestellt. Basierend auf einer kombinatorischen Suche, ist diese Methode grundsätzlich in der Lage, die maximale Untergruppe konsistenter Daten aus den Eingangsdaten zu identifizieren, die in einer Ausgleichung nach kleinsten Quadraten zu akzeptablen Residuen führt. Eine spezielle Strategie für den Einsatz der MSS-Methode in der Kongruenzuntersuchung geodätischer Netze wird ausgearbeitet.

Summary

A new method for geodetic data analysis named MSS (the maximum subsample method) is presented. Based on a combinatorial search this method is in principle always able to identify the maximum consistent subsample of the input data, which results in acceptable residuals after a least squares adjustment. A special strategy for the application of the MSS-method in the investigation of congruence of geodetic networks is elaborated.

1 Einführung

Bei vielen Aufgabenstellungen der Parameterschätzung sind in den Eingangsdaten auch solche enthalten, die mit groben Fehlern behaftet sind oder für die das angesetzte funktionale Modell unzureichend ist. Die grundsätzliche Aufgabe, die es in diesem Fall zu lösen gilt, besteht darin, diese Daten zu identifizieren und aus der Parameterschätzung auszuschließen. Die verbleibenden Daten, die dann in einer abschließenden Ausgleichung nach kleinsten Quadraten zu akzeptablen Residuen führen, werden als „maximale Untergruppe konsistenter Daten“ bezeichnet. Für die Suche nach dieser Gruppe ist in der Literatur eine Vielzahl von Methoden zu finden, die aber nicht grundsätzlich zum Erfolg führen.

In diesem Artikel wird eine neue Methode vorgestellt, die auf Grundlage einer kombinatorischen Suche grundsätzlich in der Lage ist, die maximale Untergruppe konsistenter Daten aus den Eingangsdaten zu identifizieren. Diese wird als „Methode der maximalen Untergruppe (MSS)“ bezeichnet.

Neben einer Darstellung der generellen Methodik wird der Einsatz dieser neuen Methode in der Kongruenzuntersuchung geodätischer Netze aufgezeigt.

2 Bestehende Verfahren zur Identifizierung konsistenter Datengruppen

Ist zu befürchten, dass für einen Teil der Eingangsdaten einer Parameterschätzung das angesetzte Modell nicht gültig ist oder werden grobe Fehler in den Messungen vermutet, ist ein Preprocessing erforderlich. Das Ziel aller Voruntersuchungen, unabhängig von der gewählten Methode, besteht grundsätzlich darin, die Daten zu bereinigen und das Modell gegebenenfalls zu erweitern, so dass aus einer abschließenden Parameterschätzung nach der Methode der kleinsten Quadrate akzeptable Residuen resultieren. Für diese Voruntersuchung ist in der Literatur eine Vielzahl von Strategien zu finden, die im Folgenden kurz aufgezeigt werden.

2.1 Residuenanalyse nach der Ausgleichung nach kleinsten Quadraten

Die Suche nach groben und systematischen Fehlern anhand der Ergebnisse einer Ausgleichung nach kleinsten Quadraten kann in Form einer Residuenanalyse erfolgen. Mit der in Helmert (1924, S. 333) vorgeschlagenen Prüfung der Vorzeichenverteilung der Residuen kann auf das Vorhandensein systematischer Fehler geschlossen werden. Über die Art des systematischen Fehlers lassen sich jedoch in der Regel keine Rückschlüsse ziehen.

Die Elimination von Ausreißern kann mit Hilfe einer sequentiellen Untersuchung der normierten Verbesserungen der Beobachtungen, dem sogenannten „data snooping“, erfolgen. Dies kann gelingen, wenn maximal 3 – 5% der

Daten verfälscht sind, siehe Niemeier (2002, S. 198). Der Erfolg hängt jedoch sehr stark von der Geometrie des Ausgleichsproblems (Verteilung der Redundanzanteile) ab, zudem stößt dieses Verfahren beim Vorliegen mehrerer Ausreißer aufgrund der bekannten „Verschmierungeffekte“ (Einfluss grober Fehler auf andere Beobachtungen), siehe Niemeier (2002, S. 202), schnell an seine Grenzen.

2.2 Parameterschätzung mit alternativen Zielfunktionen

Sind mehrere grobe Fehler in den Daten zu befürchten, wird oftmals ein Preprocessing mit einer alternativen Zielfunktion angeraten. Hierzu zählt die Anwendung der sogenannten resistenten und robusten Verfahren, deren Klassifizierung in Caspary (1996) zu finden ist. Ein typisches resistentes Verfahren aus der Gruppe der L_p -Norm-Schätzer ist die L_1 -Norm-Schätzung mit der Forderung

$$\sum_{i=1}^n |v_i| \rightarrow \min, \quad (2.1)$$

mit deren Hilfe versucht wird, grob fehlerhafte Beobachtungen zu identifizieren.

Aus einer Verallgemeinerung der Maximum-Likelihood-Methode, bei der unter Verwendung einer bekannten Verteilung der Messwerte die unbekannt Parameter derart geschätzt werden, dass die Likelihood-Funktion ihren maximalen Wert annimmt, ergeben sich die von Huber (1981) entwickelten robusten M-Schätzer mit der Forderung

$$\sum_{i=1}^n \rho(v_i) \rightarrow \min, \quad (2.2)$$

wobei mit $\rho(v_i)$ eine geeignet auszuwählende Verlustfunktion bezeichnet wird. Eine Weiterentwicklung stellen die modifizierten M-Schätzer dar, bei denen versucht wird, den Einfluss von Ausreißern in Hebelbeobachtungen durch Einführung zusätzlicher Gewichte zu reduzieren. Ein aktuelles Beispiel hierfür ist der von Wicki (1999, S. 87 ff.) entwickelte BIBER-Schätzer.

Einen sehr interessanten Sonderfall der alternativen Parameterschätzung stellt die LMS-Schätzung von Rousseeuw (1984) dar, bei der aus allen n Beobachtungen jeweils m (Anzahl der Unbekannten) Beobachtungen ausgewählt werden, um die unbekannt Parameter eindeutig (widerspruchsfrei) zu bestimmen. Auf Grundlage dieser Lösung werden dann für alle Beobachtungen die Verbesserungen berechnet. Diese werden quadriert und der Median wird bestimmt. Dieser Vorgang wird für alle möglichen Kombinationen von m Beobachtungen durchgeführt. Als Lösung wird dann diejenige ausgewählt, die die Forderung

$$\text{median}(v_i^2) \rightarrow \min \quad (2.3)$$

erfüllt. Der interessante Aspekt an der LMS-Schätzung liegt darin, dass eine kombinatorische Suche durchgeführt wird und somit Ausreißer in den Hebelpunkten keinen Einfluss auf das Schätzergebnis ausüben können.

Eine weitere alternative Zielfunktion, die bei unvollständigen geodätischen Modellen angewendet werden kann, um fehlende Modellanteile zu identifizieren, ist der lineare Korrelationskoeffizient (quadriert) zwischen den Daten l und dem funktionalen Modell $f(u)$, der durch eine Ausgleichung nach maximaler Korrelation, siehe Neitzel (2004, S. 65 ff.) und Petrovic (1991, S. 24 ff.), deren Begründung rein geometrischer Natur ist, mit

$$r^2(l, f(u)) \rightarrow \max \quad (2.4)$$

seinen maximal möglichen Wert annimmt.

Grundsätzlich soll, unabhängig von der gewählten Methode, als Ergebnis einer Voruntersuchung feststehen, welche *maximale Untergruppe* der Eingangsdaten (eventuell nach einer Vervollständigung des mathematischen Modells) einer abschließenden Ausgleichung nach kleinsten Quadraten unterzogen werden kann.

2.3 Zuverlässigkeit der Ausreißersuche mit alternativen Schätzverfahren

Als Maß für die globale Zuverlässigkeit eines Schätzverfahrens wird der Bruchpunkt angegeben. Damit wird der Prozentsatz von fehlerhaften Daten beliebiger Größe bezeichnet, der erlaubt ist, bevor das Schätzverfahren falsche Werte liefert beziehungsweise „zusammenbricht“, siehe Niemeier (2002, S. 198). Der Bruchpunkt für die L_1 - und die LMS-Schätzung beträgt bis zu 50%. In Hekimoglu und Koch (1999) wird ausgeführt, dass dieser Wert die Fähigkeit eines Schätzverfahrens charakterisiert, mit sehr großen Ausreißern zurechtzukommen. Da diese aber in der geodätischen Praxis eher selten auftreten, wurde mit der „minimalen mittleren Erfolgsrate“ ein

Zuverlässigkeitsmaß entwickelt, für den Fall, dass mehrere kleine Ausreißer im Datenmaterial vorhanden sind. Untersuchungen der Schätzverfahren nach Huber, eines modifizierten M-Schätzers, der LMS-Methode, des Andrews-Schätzers und der L_1 -Schätzung an Beispielen der einfachen und multiplen linearen Regression haben in Hekimoglu und Koch (1999) zu folgendem Ergebnis geführt:

- Alternative Schätzverfahren können Ausreißer produzieren, obwohl im Datenmaterial keine enthalten waren.
- Die Zuverlässigkeit ist abhängig von der Anzahl der Unbekannten, der Anzahl und Größenordnung der Ausreißer und von der Art der Ausreißer und der Position (Geometrie) der Beobachtungen.
- Die Zuverlässigkeit robuster Methoden nimmt rapide ab, wenn die Anzahl der Unbekannten zunimmt.

Zu vergleichbaren Ergebnissen kommen auch Berber und Hekimoglu (2001) bei der Untersuchung resistenter und robuster Schätzverfahren in einem geodätischen Netz und Hekimoglu (2001) bei einer speziellen Untersuchung der LMS-Schätzung an Beispielen der linearen Regression.

Abschließend bleibt somit festzustellen, dass es unter den resistenten oder robusten Schätzverfahren keine „Universalmethode“ für die Datenanalyse gibt und dass der Erfolg eines gewählten Verfahrens immer sehr stark von der Charakteristik des jeweiligen Ausgleichungsproblems abhängt.

2.4 Das spezielle Problem der Kongruenzuntersuchung

Die generelle Aufgabe, aus den Eingangsdaten die maximale Untergruppe konsistenter Daten zu finden, stellt sich sehr anschaulich bei der Kongruenzuntersuchung geodätischer Netze. Hierbei besteht die Zielsetzung, aus der Gesamtheit aller Netzpunkte (beziehungsweise aller Stützpunkte) die maximale Untergruppe zu lokalisieren, für die die Hypothese der Kongruenz angenommen werden kann.

Seit etwa 1970 wurden intensive Forschungen zu dieser Thematik durchgeführt und schließlich hat sich ein Verfahren herauskristallisiert, das in der aktuellen Literatur zum Beispiel in Welsch et al. (2000, S. 369 ff.) und Niemeier (2002, S. 365 ff.) dargestellt ist. Grundsätzlich gliedert sich die Untersuchung in zwei Teilschritte:

1. Globaler Kongruenztest: Mit Hilfe eines statistischen Tests wird zunächst untersucht, ob zwischen den zu untersuchenden Epochen überhaupt signifikante Punktverschiebungen vorhanden sind.
2. Lokalisierung: Durch eine sukzessive Untersuchung einzelner Punkte nach der Ausgleichung nach kleinsten Quadraten sollen signifikante Punktverschiebungen erkannt werden.

Diese Vorgehensweise wird im folgenden Abschnitt anhand eines numerischen Beispiels veranschaulicht.

2.5 Numerisches Beispiel

Die Untersuchung dieses Streckennetzes, das in Anlehnung an ein numerisches Beispiel in Reinking (1994, S. 53 ff.) konzipiert wurde, erfolgt mit Hilfe datumsinvarianter Elemente nach der in Welsch et al. (2000, S. 381 ff. und S. 393) dargestellten Vorgehensweise. Das in Abb. 2.1 dargestellte Netz besteht aus zehn Punkten, die Punkte 7, 8 und 9 bilden eine kongruente Punktgruppe, Maßstabsunterschiede zwischen den Epochen sind nicht vorhanden. Eine Unterteilung in Stütz- und Objektpunkte ist nicht gegeben, so dass es sich um ein Relativmodell handelt. Die Näherungskoodinaten für die Netzausgleichung und die Streckenmessungen in Epoche 1 und 2 können Neitzel (2004, S. 100 ff.) entnommen werden.

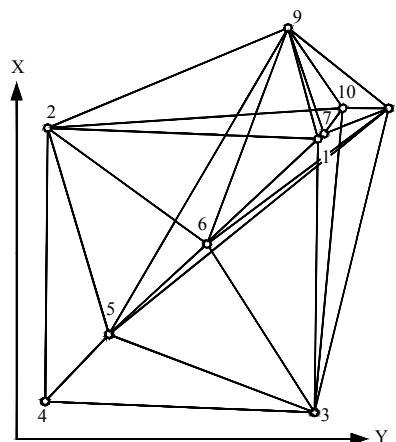


Abb. 2.1: Streckennetz, nicht alle Strecken dargestellt

Aus einer Ausgleichung der Einzelepochen jeweils als freies Netz mit Gesamtpurminimierung, wobei die theoretische Standardabweichung jeweils mit $\sigma_0 = 1$ cm angenommen wird, erhält man die ausgeglichenen Koordinaten in Tab. 2.1 mit den Kofaktorenmatrizen $\mathbf{Q}_{x_1x_1}$ und $\mathbf{Q}_{x_2x_2}$.

Tab. 2.1: Koordinaten aus der freien Netzausgleichung

Pkt. Nr.	Epoche I		Epoche II	
	\hat{Y} [m]	\hat{X} [m]	\hat{y} [m]	\hat{x} [m]
1	220.003	219.991	222.006	217.502
2	20.005	220.006	22.500	222.509
3	219.996	19.993	217.505	17.500
4	19.999	20.000	15.999	25.500
5	70.003	70.000	68.003	73.002
6	140.005	139.997	139.998	140.495
7	224.997	220.007	225.002	219.996
8	275.003	240.004	275.004	239.996
9	199.988	300.002	199.992	299.998
10	240.001	240.001	241.990	237.501

Die gewichtete Quadratsumme der Verbesserungen und die empirische Einheitsvarianz für Epoche 1 ergibt sich zu

$$\Omega_1 = \mathbf{v}_1^T \mathbf{P}_1 \mathbf{v}_1 = 4.5460 \cdot 10^{-3} \text{ m}^2, \quad (2.5)$$

$$s_{0_1}^2 = \Omega_1 / f_1 = 4.5460 \cdot 10^{-3} \text{ m}^2 / 28 = 1.6236 \cdot 10^{-4} \text{ m}^2 \quad (2.6)$$

und für Epoche 2 erhält man

$$\Omega_2 = \mathbf{v}_2^T \mathbf{P}_2 \mathbf{v}_2 = 2.4644 \cdot 10^{-3} \text{ m}^2, \quad (2.7)$$

$$s_{0_2}^2 = \Omega_2 / f_2 = 2.4644 \cdot 10^{-3} \text{ m}^2 / 28 = 8.8015 \cdot 10^{-5} \text{ m}^2. \quad (2.8)$$

Zur Überprüfung der Nullhypothese $H_0: E\{s_{0_1}^2\} = E\{s_{0_2}^2\}$ wird die Testgröße

$$T_F = \frac{s_{0_1}^2}{s_{0_2}^2} = \frac{1.6236 \cdot 10^{-4} \text{ m}^2}{8.8015 \cdot 10^{-5} \text{ m}^2} = 1.84 \quad (2.9)$$

gebildet und mit dem Quantil der Fisher-Verteilung verglichen. Mit den Freiheitsgraden $f_1 = f_2 = 28$ und einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ ergibt sich für $F_{f_1, f_2, 1-\alpha/2} = 2.13$, somit kann die Nullhypothese nicht verworfen werden und mit

$$s_0^2 = \sum_{i=1}^2 \Omega_i / \sum_{i=1}^2 f_i = 7.0104 \cdot 10^{-3} \text{ m}^2 / 56 = 1.2519 \cdot 10^{-4} \text{ m}^2 \quad (2.10)$$

kann ein gemeinsamer Varianzfaktor eingeführt werden. Der Globaltest kann mit einer Minimalkonfiguration aus h datumsinvarianten Elementen durchgeführt werden, die in diesem Fall aus $h = 17$ Strecken besteht, die das Netz eindeutig beschreiben. Mit

$\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$: ausgeglichene Koordinaten in Epoche 1 und 2 und

$\mathbf{L}_1, \mathbf{L}_2$: Funktionalmatrizen zur Berechnung der Strecken in Epoche 1 und 2

erhält man aus

$$R = \mathbf{dl}^T \mathbf{Q}_{dl}^{-1} \mathbf{dl} = \quad (2.11)$$

$$(\mathbf{L}_2 \hat{\mathbf{x}}_2 - \mathbf{L}_1 \hat{\mathbf{x}}_1)^T (\mathbf{L}_1 \mathbf{Q}_{x_1x_1} \mathbf{L}_1^T + \mathbf{L}_2 \mathbf{Q}_{x_2x_2} \mathbf{L}_2^T)^{-1} (\mathbf{L}_2 \hat{\mathbf{x}}_2 - \mathbf{L}_1 \hat{\mathbf{x}}_1)$$

den Zuschlag zur Verbesserungsquadratsumme $R = 53.3 \text{ m}^2$. Damit wird die Testgröße

$$T_F = \frac{R/h}{s_0^2} = \frac{53.3 \text{ m}^2 / 17}{1.2519 \cdot 10^{-4} \text{ m}^2} = 25043 \quad (2.12)$$

gebildet und mit dem Quantil der Fisher-Verteilung verglichen. Mit den Freiheitsgraden $h = 17, f = 56$ und einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ ergibt sich für $F_{h,f,1-\alpha} = 1.81$, somit liegen signifikante Deformationen vor.

Zum Auffinden der kongruenten Punktgruppe wird nun die Strategie der Einzelpunktanalyse angewendet. Dazu wird jeweils ein Punkt aus dem Netz ausgeschlossen und ein Globaltest mit einer Minimalkonfiguration datumsinvarianter Größen (Strecken) durchgeführt. Damit erhält man für jedes Teilnetz einen Wert R_F (Tab. 2.2), wobei der Punkt, der zu dem geringsten Wert geführt hat, als verschoben gilt.

Tab. 2.2: Werte R_F [m²] nach dem ersten Lokalisierungsschritt

R_{F_1}	R_{F_2}	R_{F_3}	R_{F_4}	R_{F_5}	R_{F_6}	R_{F_7}	R_{F_8}	R_{F_9}	$R_{F_{10}}$
50.2	43.0	45.2	43.2	48.8	52.0	49.1	48.1	40.5	50.0

Aus diesem Ergebnis ist ersichtlich, dass ebenfalls wie in Reinking (1994, S. 53 ff.) gezeigt, der Punkt 9 fälschlicherweise aus der Gruppe der stabilen Punkte zu entfernen ist, da er zu dem niedrigsten Wert für R_F führt. Da die Identifizierung dieses Punktes nur mit datumsinvarianten Größen erfolgt ist, ist der Grund für diese falsche Lokalisierung jedoch nicht in der Datumsübertragung zu suchen, wie in Reinking (1994, S. 53) ausgeführt wird, sondern die Probleme müssen an anderer Stelle liegen.

3 Beurteilung und Motivation neuer Methoden

Kritische Anmerkungen zu dem dargestellten Verfahren sind in Caspary und Borutta (1986) zu finden. Da in allen Schritten die Methode der kleinsten Quadrate angewendet wird, auch wenn das Modell noch sehr grob ist und unentdeckte Einzelpunktbewegungen vorhanden sind, wird auf die Gefahr hingewiesen, dass auf der Grundlage nicht optimaler Schätzergebnisse Entscheidungen für das weitere Vorgehen getroffen werden müssen. Es wird die Schlussfolgerung gezogen, dass der Mangel nur durch den Übergang auf andere Schätzverfahren behoben werden kann.

Die Frage, warum sich der Einsatz alternativer Schätzverfahren in der Kongruenzuntersuchung nicht durchgesetzt hat, ergibt sich aus den Ausführungen in Abschnitt 2.3. Betrachtet man zudem das numerische Beispiel, dann ist sofort einzusehen, dass auch die Anwendung eines resistenten oder robusten Verfahrens nicht zum Erfolg führen kann, da von 10 Punkten lediglich 3 stabil sind. Dies entspricht einem Anteil „kontaminierter Daten“ an den Gesamtdaten von 70%, womit auch jedes resistente oder robuste Verfahren überfordert ist.

In Neitzel (2004, S. 103 ff.) wurde die Kongruenzuntersuchung mit Hilfe der Ausgleichung nach maximaler Korrelation untersucht, wobei der Korrelationskoeffizient als nichtmetrisches Kriterium für den Vergleich der Form von Teilnetzen eingesetzt wurde. Doch auch die sukzessive Lokalisierung mit Hilfe des Korrelationskoeffizienten hat in dem numerischen Beispiel nicht zur korrekten Lösung geführt.

Der Grund für das Scheitern aller Verfahren liegt in der Strategie der *sukzessiven Einzelpunktanalyse*. Während eines Lokalisierungsschrittes wird hierbei jeweils immer nur *ein* Punkt als verschoben verdächtigt, für die *übrigen* Punkte wird implizit angenommen, dass sie sich *nicht* verschoben haben. Somit lassen sich auch die Ergebnisse in Abschnitt 2.5 und die Untersuchungen zur Ausgleichung nach maximaler Korrelation in Neitzel (2004, S. 103 ff.) erklären. Punkt 9 ist tatsächlich derjenige, der die Kongruenz beider Netze am meisten stört, wenn man die übrigen Punkte als nicht verschoben annimmt. Es gibt aber keine Gewähr dafür, dass die tatsächlich kongruente Punktgruppe in der Untergruppe, die nach der sukzessiven Elimination einzelner Punkte verbleibt, enthalten ist. Sind mehrere Punkte verschoben, können diese Punkte derart Einfluss auf die Entscheidungsgröße nehmen, dass es zu Fehlinterpretationen kommt. Vergleichbar ist diese Situation mit der Suche nach Ausreißern mittels „data snooping“. Auch dort wird bei der sukzessiven Elimination angenommen, dass jeweils nur ein Ausreißer vorhanden ist. Liegt eine größere Anzahl vor, kann es auch dort passieren, dass Ausreißer an den falschen Stellen angezeigt werden.

In Bezug auf die Kongruenzuntersuchung besteht der Ausweg darin, *alle möglichen* Untergruppen zu untersuchen, das heißt, hat ein Globaltest mit p Punkten ergeben, dass Deformationen vorhanden sind, dann sind *alle* Untergruppen mit $p-1$ Punkten zu untersuchen. Kann die Hypothese der Kongruenz danach nicht angenommen werden, sind, ausgehend von der Gesamtheit aller p Punkte, *alle* Untergruppen mit $p-2$ Punkten zu untersuchen. Hat dies nicht zum Erfolg geführt, dann sind, wieder ausgehend von der Gesamtheit aller p Punkte, *alle* Untergruppen mit $p-3$ Punkten zu untersuchen und so weiter. Schlussfolgerung:

Die Lokalisierung kongruenter Punktgruppen in mehrfach beobachteten geodätischen Netzen ist grundsätzlich eine Aufgabe der Kombinatorik.

Nun soll eine generelle Methodik entwickelt werden, mit der eine kombinatorische Suche nach der maximalen Untergruppe konsistenter Daten möglich wird. Diese soll dann für den speziellen Fall der Kongruenzuntersuchung ausgearbeitet und anhand des numerischen Beispiels aus Abschnitt 2.5 getestet werden.

4 Die Methode der maximalen Untergruppe (MSS)

Da die Identifizierung der maximalen Untergruppe konsistenter Daten weder mit einer sukzessiven Residuenanalyse einzelner Beobachtungen nach einer Ausgleichung nach kleinsten Quadraten, noch mit der Anwendung alternativer Schätzverfahren in allen Fällen gelöst werden kann, stellt sich die Frage nach einer neuen Herangehensweise. Betrachtet man die alternativen Schätzverfahren etwas näher, fällt das LMS-Verfahren auf, da hierbei die Lösung mit Hilfe der Kombinatorik gesucht wird. Mit der Forderung (2.3) wird jeweils eine *minimale* Untergruppe gesucht, für die sich dann die Verbesserungen der Beobachtungen zu null ergeben. Die Beurteilung der übrigen Daten erfolgt dann in Bezug zu dieser Minimalkonfiguration. Dass diese Vorgehensweise nicht immer zu brauchbaren Ergebnissen führt, ist in Abschnitt 2.3 dargelegt und auch Kampmann (1993) kommt bei der Untersuchung der LMS-Methode zu unplausiblen Ergebnissen.

Die Grundidee, Ausgleichungsprobleme mit Hilfe der Kombinatorik zu lösen, soll im Folgenden beibehalten werden. Es soll eine neue Strategie entwickelt werden, mit der es möglich ist, das eigentliche Ziel, nämlich die Identifizierung der *maximalen* Untergruppe, die in einer Ausgleichung nach kleinsten Quadraten zu akzeptablen Residuen führt, auf direktem Wege anzugehen. Diese neue Methode, bei der mit einer *kombinatorischen Suche* die maximale Untergruppe verträglicher Daten aus der Gesamtheit aller Daten gesucht wird, wird mit dem Namen „MSS (Maximum Subsample) - Die Methode der maximalen Untergruppe“ bezeichnet.

Wann man ein Ergebnis als „verträglich“ annehmen kann, ist von der zugrunde liegenden Aufgabenstellung abhängig. Ein mögliches Beurteilungskriterium sind die normierten Verbesserungen der Beobachtungen oder ein statistischer Test (z.B. Globaltest bei der Kongruenzuntersuchung).

In den folgenden Abschnitten wird nun aufgezeigt, wie diese kombinatorische Lösung allgemein gefunden werden kann.

4.1 Direkte Lösung des MSS-Problems

Gegeben ist ein Ausgleichungsproblem mit n Messungen und m Unbekannte. Das Ziel besteht darin, solange alle möglichen Kombinationen von Beobachten einer Ausgleichung nach kleinsten Quadraten zu unterziehen, bis eine akzeptable Lösung gefunden ist. Alle Kombinationen durchzuprobieren ist natürlich bei größeren Problemen nicht realistisch. In einem Beispiel mit $n = 50$ Beobachtungen, $m = 10$ Unbekannte und einem Anteil kontaminierter Daten von 30% (15 fehlerhafte Beobachtungen) wären bis zu

$$K = \binom{50}{49} + \binom{50}{48} + \binom{50}{47} + \dots + \binom{50}{35} = 3.7157E + 12 \quad (4.1)$$

Ausgleichungen zu berechnen, bis die maximale Untergruppe gefunden wird.

Die Frage, die es nun zu beantworten gilt, ist die nach einer Strategie, die Anzahl der möglichen Kombinationen soweit einzuschränken, dass eine Lösung mit vertretbarem Aufwand erzielt werden kann.

4.2 Zufallsbasierte Auswahl von Kombinationen

Eine allgemeine Einschränkung der Kombinationen kann erreicht werden, in dem nicht grundsätzlich alle Kombinationen untersucht werden, sondern dass per Zufallsgenerator jeweils eine festgelegte Anzahl von Untergruppen untersucht wird. Dabei kann dann allerdings nicht mehr garantiert werden, dass tatsächlich die maximale Untergruppe gefunden wird. Um diesem Umstand entgegenzuwirken, sollten die nicht zu berücksichtigenden Beobachtungen derart aus der Ausgleichung eliminiert werden, dass ihre Gewichte zu null gesetzt werden. Somit nehmen diese Beobachtungen nicht an der Parameterschätzung teil, erhalten aber Verbesserungen, anhand derer man sowohl die Größe des groben Fehlers ablesen kann als auch erkennen kann, ob eine Beobachtung doch zur Gruppe der verträglichen Daten zugeschlagen werden kann.

4.3 Ausschluss unmöglicher Kombinationen

Die erfolgversprechendste Lösung des MSS-Problems besteht darin, dass man im Rahmen einer Voruntersuchung diejenigen Kombinationen ausschließt, die sowieso nicht Bestandteil einer konsistenten Gruppe sein können. Mit den verbleibenden Eingangsdaten kann dann die MSS-Lösung in kurzer Zeit gefunden werden. Ein Beispiel zur Anwendung dieser Strategie bei der Ausgleichung direkter Beobachtungen ist in Neitzel (2004, S. 110 ff.) aufgeführt. Diese Strategie bietet sich auch insbesondere bei der Kongruenzuntersuchung geodätischer Netze an.

5 Einsatz der MSS-Methode in der Kongruenzuntersuchung

Hat ein Globaltest ergeben, dass Punktverschiebungen zwischen den zu untersuchenden Epochen aufgetreten sind, besteht die Aufgabe, die maximale Untergruppe stabiler Punkte zu identifizieren. Wie bereits beschrieben, kann diese maximale Gruppe in jedem Fall durch eine Untersuchung aller Punktkombinationen gefunden werden. Um jedoch nicht alle Kombinationen berechnen zu müssen, sollen nur diejenigen Punkte untersucht werden, bei denen überhaupt die Aussicht besteht, dass sie eine kongruente Gruppe bilden können. Dazu wird im Folgenden eine Strategie für die Voruntersuchung unter Verwendung datumsinvarianter Elemente (Strecken) entwickelt.

5.1 Ausschluss unmöglicher Kombinationen

Auf Grundlage der ausgeglichenen Koordinaten in Epoche 1 und 2 werden mit

$$\hat{\mathbf{l}}_1 = \mathbf{L}_1 \hat{\mathbf{x}}_1, \quad \mathbf{Q}_{1,1} = \mathbf{L}_1 \mathbf{Q}_{x_1 x_1} \mathbf{L}_1^T, \quad (5.1)$$

$$\hat{\mathbf{l}}_2 = \mathbf{L}_2 \hat{\mathbf{x}}_2, \quad \mathbf{Q}_{2,2} = \mathbf{L}_2 \mathbf{Q}_{x_2 x_2} \mathbf{L}_2^T \quad (5.2)$$

jeweils *alle* möglichen Strecken mit ihren Kofaktorenmatrizen berechnet. Daraus lässt sich der Differenzvektor und dessen Kofaktorenmatrix

$$\mathbf{dl} = \hat{\mathbf{l}}_2 - \hat{\mathbf{l}}_1, \quad \mathbf{Q}_{dl} = \mathbf{Q}_{1,1} + \mathbf{Q}_{2,2} \quad (5.3)$$

berechnen, zudem steht der gemeinsame empirische Varianzfaktor s_0^2 zur Verfügung. Um unmögliche Konfigurationen auszuschließen, werden die Differenz dl_{ij} der Strecken von Punkt i nach Punkt j , mit $i = (1, \dots, p - 1), j = (i + 1, \dots, p), p =$ Anzahl der zu untersuchenden Punkte, im Sinne der Informationstheorie als „Signal“ und die empirischen Standardabweichungen $s_{dl_{ij}}$ als „Rauschen“ angesehen. Somit bezeichnen die Quotienten

$$q_{dl_{ij}} = \frac{|dl_{ij}|}{s_{dl_{ij}}} \quad (5.4)$$

das „Signal-Rausch-Verhältnis“ bei dem man davon ausgehen kann, dass es sich bei $q_{dl_{ij}} > 5$ um signifikante Abweichungen handelt. Alle Strecken, die eine derartige Abweichung aufweisen, können somit nicht Bestandteil einer kongruenten Punktgruppe sein. Diese Information gilt es nun derart aufzubereiten, dass man „unmögliche“ Kombinationen von vornherein aus der weiteren Suche ausschließen kann.

5.2 Topologiematrix möglicher kongruenter Punktgruppen

Mit den verbleibenden Strecken kann anhand topologischer Beziehungen eine Lokalisierung möglicher kongruenter Punktgruppen durchgeführt werden. Topologische Beziehungen in einem Netz lassen sich allgemein durch eine Kanten-Knoten-Matrix beschreiben; die Verwendung einer derartigen Matrix für geodätische Fragestellungen ist zum Beispiel in Gründig (1988, S. 13 ff.) beschrieben. Die Knoten des Netzes sind durch Kanten verbunden. Werden die Beobachtungen den Kanten zugeordnet, lassen sich die topologischen Beziehungen in einer Kanten-Knoten-Matrix \mathbf{C} abbilden. Für ein Netz mit fünf Punkten, zwischen denen alle möglichen Strecken gemessen wurden (Abb. 5.1), ergibt sich die Topologiematrix in Tab. 5.1. In \mathbf{C} wird die i -te Beobachtung dem j -ten Punkt zugeordnet. Es gilt folgende Definition:

$C_{ij} = 1$, wenn die i -te Beobachtung vom Punkt j ausgeht,

$C_{ij} = -1$, wenn die i -te Beobachtung im Punkt j endet,

$C_{ij} = 0$, in allen übrigen Fällen.

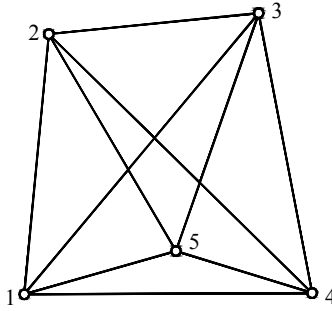


Abb. 5.1: Streckennetz mit fünf Punkten

Tab. 5.1: Kanten-Knoten-Matrix für ein Streckennetz mit fünf Punkten

	Pkt. 1	Pkt. 2	Pkt. 3	Pkt. 4	Pkt. 5
$s_{1,2}$	1	-1	0	0	0
$s_{1,3}$	1	0	-1	0	0
$s_{1,4}$	1	0	0	-1	0
$s_{1,5}$	1	0	0	0	-1
$s_{2,3}$	0	1	-1	0	0
$s_{2,4}$	0	1	0	-1	0
$s_{2,5}$	0	1	0	0	-1
$s_{3,4}$	0	0	1	-1	0
$s_{3,5}$	0	0	1	0	-1
$s_{4,5}$	0	0	0	1	-1

Bildet man mit dieser Kanten-Knoten-Matrix den Ausdruck

$$\bar{C} = C^T C, \quad (5.5)$$

erhält man eine symmetrische Knoten-Knoten-Matrix. Diese besitzt in dem dargestellten Beispiel die Dimension 5×5 , auf der Hauptdiagonalen steht der Wert „4“, die übrigen Elemente sind mit dem Wert „-1“ belegt.

Die Elemente der Knoten-Knoten-Matrix \bar{C} haben die folgende Bedeutung:

- \bar{C}_{ii} = Anzahl der im Punkt i zusammentreffenden Kanten,
- \bar{C}_{ij} = -1, wenn eine Kantenverbindung zwischen Punkt i und Punkt j besteht,
- \bar{C}_{ij} = 0, wenn keine Kantenverbindung zwischen Punkt i und Punkt j besteht.

Für die Suche nach möglichen kongruenten Punktgruppen werden nur die Streckenverbindungen in die Matrix C eingetragen, deren Abweichung akzeptiert wurde. Dann wird die zugehörige Knoten-Knoten-Matrix \bar{C} gebildet, in der die Suche wie folgt durchgeführt wird:

1. Suche über alle \bar{C}_{ii} das maximale Element $\bar{C}_{ii} = \max$.
2. Suche über alle \bar{C}_{ij} , wie viele Elemente k mit $\bar{C}_{ij} \geq \max$ gefunden werden. Ist $k \geq \max + 1$, können kongruente Punktgruppen mit $\max + 1$ Punkten vorliegen.
3. Ist die Anzahl der gefundenen Punkte k gleich dem Wert $\max + 1$, kann getestet werden, ob zwischen diesen Punkten alle Streckenverbindungen existieren. Ist dies der Fall, dann ist diese Gruppe ein Kandidat für eine mögliche kongruente Punktgruppe.
4. Ist $k > \max + 1$, dann sind mit den gefundenen Punkten alle möglichen Kombinationen zu bilden und zu prüfen, ob für diese Kombinationen alle Streckenverbindungen vorhanden sind. Die Punktgruppen, in denen alle Streckenverbindungen existieren, sind Kandidaten für eine kongruente Punktgruppe.
5. Die Kandidaten für eine kongruente Punktgruppe werden unter Anwendung des Globaltests untersucht. Konnte eine stabile Gruppe gefunden werden, sind, unter der Annahme dass unterschiedliche kongruente Gruppen keine gemeinsamen Punkte aufweisen, die entsprechenden Zeilen und Spalten in der Matrix \bar{C} zu streichen und die Suche nach weiteren Gruppen kann mit Schritt 1 fortgesetzt werden.

6. Konnte keine kongruente Punktgruppe gefunden werden, wird $max = max - 1$ gesetzt und die Lokalisierung mit Schritt 2 fortgesetzt.

Diese Suche soll nun an einem einfachen Beispiel veranschaulicht werden. In dem in Abb. 5.2 dargestellten Netz wurden in zwei Epochen alle Strecken berechnet. Die Beurteilung der Streckendifferenzen anhand des Signal-Rausch-Verhältnisses liefert die Testgrößen $q_{dl_{ij}}$ (5.4), die in Tab. 5.2 in sortierter Reihenfolge aufgelistet sind.

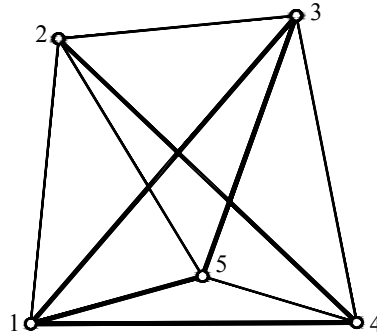


Abb. 5.2: Netz mit fünf Punkten. Strecken, deren Abweichung akzeptiert wird, sind in dickerer Strichstärke dargestellt.

Tab. 5.2: Testgrößen $q_{dl_{ij}}$

Zeile	Testgröße
1	$q_{dl_{2,5}} = 6.8$
2	$q_{dl_{1,2}} = 5.9$
3	$q_{dl_{2,3}} = 5.5$
4	$q_{dl_{4,5}} = 5.4$
5	$q_{dl_{3,4}} = 5.2$
6	$q_{dl_{1,4}} = 1.7$
7	$q_{dl_{3,5}} = 1.3$
8	$q_{dl_{1,3}} = 1.1$
9	$q_{dl_{2,4}} = 0.9$
10	$q_{dl_{1,5}} = 0.4$

Da die ersten fünf Werte mit $q_{dl_{ij}} > 5$ über dem Grenzwert liegen, beschränkt sich die Suche nach der maximalen Gruppe stabiler Punkte auf ein Suchfenster von Zeile 6 bis 10. Nur diese Streckenverbindungen werden in eine Kanten-Knoten-Matrix \mathbf{C} eingetragen, die sich dann wie in Tab. 5.3 dargestellt ergibt. Die zugehörige Knoten-Knoten-Matrix $\bar{\mathbf{C}}$ ist in Tab. 5.4 dargestellt.

Tab. 5.3: Kanten-Knoten-Matrix

	Pkt. 1	Pkt. 2	Pkt. 3	Pkt. 4	Pkt. 5
$s_{1,3}$	1	0	-1	0	0
$s_{1,4}$	1	0	0	-1	0
$s_{1,5}$	1	0	0	0	-1
$s_{2,4}$	0	1	0	-1	0
$s_{3,5}$	0	0	1	0	-1

Tab. 5.4: Knoten-Knoten-Matrix

	Pkt. 1	Pkt. 2	Pkt. 3	Pkt. 4	Pkt. 5
Pkt. 1	3	0	-1	-1	-1
Pkt. 2	0	1	0	-1	0
Pkt. 3	-1	0	2	0	-1
Pkt. 4	-1	-1	0	2	0
Pkt. 5	-1	0	-1	0	2

Die Lokalisierung möglicher kongruenter Punktgruppen anhand der Knoten-Knoten-Matrix \bar{C} geschieht in diesem Beispiel wie folgt:

- Das maximale Diagonalelement ist $\bar{C}_{11} = \max = 3$, was auf das Vorhandensein eines Vierecks deutet, da im Eckpunkt eines Vierecks drei Seiten auftreten.
- Es kann nur $k = 1$ Element mit $\bar{C}_{ii} \geq \max$ gefunden werden, so dass kein Viereck vorliegen kann.
- Setze $\max = \max - 1 = 2$.
- Es können $k = 4$ Elemente mit $\bar{C}_{ii} \geq \max$ gefunden werden, somit können mehrere Dreiecke vorliegen.
- Mit den gefundenen Punkten 1, 3, 4 und 5 lassen sich folgende Dreiecke bilden: 1/3/4, 1/3/5, 1/4/5, 3/4/5. Die Überprüfung, ob in jedem Dreieck alle Streckenverbindungen vorhanden sind, entspricht einer Untersuchung der entsprechenden Knoten-Knoten-Matrizen, die sich aus Tab. 5.4 ergeben. In Tab. 5.5 und Tab. 5.6 sind exemplarisch zwei der vier Knoten-Knoten-Matrizen aufgeführt.

Tab. 5.5: Knoten-Knoten-Matrix zur Überprüfung des Dreiecks 1/3/4

	Pkt. 1	Pkt. 3	Pkt. 4
Pkt. 1	3	-1	-1
Pkt. 3	-1	2	0
Pkt. 4	-1	0	2

Tab. 5.6: Knoten-Knoten-Matrix zur Überprüfung des Dreiecks 1/3/5

	Pkt. 1	Pkt. 3	Pkt. 5
Pkt. 1	3	-1	-1
Pkt. 3	-1	2	-1
Pkt. 5	-1	-1	2

Da die Matrix \bar{C} symmetrisch ist, kann man jeweils anhand der Elemente oberhalb der Hauptdiagonalen erkennen, ob in der zu untersuchenden Punktgruppe alle Streckenverbindungen vorhanden sind. Dies ist nur für die Kombination der Punkte 1, 3, 5 der Fall, so dass ein kongruentes Dreieck vorliegen kann. Die Hypothese der Kongruenz kann dann mit einem Globaltest überprüft werden. In diesem Beispiel wird angenommen, dass die Nullhypothese der Kongruenz angenommen werden kann.

- Möchte man untersuchen, ob sich weitere Gruppen finden lassen, sind die Zeilen und Spalten der lokalisierten Punkte 1, 3 und 5 aus der ursprünglichen Matrix \bar{C} in Tab. 5.4 zu streichen und man erhält die in Tab. 5.7 dargestellte Knoten-Knoten-Matrix als Ausgangspunkt für eine weitere Suche.

Tab. 5.7: Knoten-Knoten-Matrix

	Pkt. 2	Pkt. 4
Pkt. 2	1	-1
Pkt. 4	-1	2

- Ohne Darstellung der Zwischenschritte wird als Ergebnis eine weitere Gruppe bestehend aus den Punkten 2 und 4 gefunden.

Vergleicht man den Aufwand, mit dem die größte kongruente Punktgruppe gefunden wurde, ist zunächst zu erwähnen, dass man bereits in diesem kleinen Beispiel bei einer Berechnung aller Kombinationen insgesamt

$$K = \binom{5}{4} + \binom{5}{3} = 15 \quad (5.6)$$

Ausgleichungen hätte berechnen müssen. Die Voruntersuchung führt also zu einer erheblichen Reduzierung des Rechenaufwandes, so dass die kombinatorische Suche nach der maximalen Gruppe stabiler Punkte auch in größeren Netzen praktikabel ist.

Sind Maßstabsunterschiede zwischen den Vergleichsepochen zu erwarten, ist eine Voruntersuchung anhand von Streckendifferenzen nicht möglich. In diesem Fall kann eine Voruntersuchung anhand von Maßstabsverhältnissen, siehe Neitzel (2004, S. 117 ff.), durchgeführt werden. Das Ergebnis dieser Untersuchung kann ebenfalls mit Hilfe von Topologiematrizen weiterverarbeitet werden.

Liegen mehr als zwei Messepochen vor, besteht die Aufgabe, eine Aussage zu treffen, welche Punkte über den gesamten Untersuchungszeitraum als stabil eingestuft werden können. Dies kann mit Hilfe eines kumulativen Mehrepochevergleichs oder durch einen mehrfachen Vergleich mit einer Nullepoche erfolgen. In beiden Fällen kann die MSS-Methode angewendet werden, wobei das Ergebnis der Voruntersuchung wieder mit Hilfe von Topologiematrizen weiterverarbeitet werden kann, um Kandidaten für eine kongruente Punktgruppe zu finden. Eine detaillierte Darstellung des Mehrepochevergleichs ist in Neitzel (2004, S. 120 ff.) aufgeführt.

5.3 Numerisches Beispiel

Als numerisches Beispiel wird wieder das Netz aus Abschnitt 2.5 betrachtet, bei dem die Einzelpunktanalyse dazu geführt hat, dass der Punkt 9 fälschlicherweise als verschoben identifiziert wurde. Nun soll versucht werden, die kongruente Punktgruppe mit Hilfe der MSS-Methode zu identifizieren.

Im ersten Schritt werden, gemäß Abschnitt 5.1, diejenigen Punktkombinationen ausgeschlossen, bei denen keine Aussicht besteht, dass sie eine kongruente Gruppe bilden können. Dazu werden aus den ausgeglichenen Koordinaten in Tab. 2.1 in jeder Epoche mit (5.1) und (5.2) jeweils alle möglichen Strecken und die dazugehörigen Kofaktorenmatrizen berechnet. Daraus wird mit (5.3) der Differenzvektor und die zugehörige Kofaktorenmatrix berechnet. Die einzelnen Elemente des Differenzvektors sind in Tab. 5.8 aufgeführt.

Tab. 5.8: Streckendifferenzen zwischen Epoche 1 und 2

dl_{ij} [m]	$j = 2$	3	4	5	6	7	8	9	10
$i = 1$	-0.429	0.055	-1.230	-0.944	-0.635	-1.096	-0.955	2.906	<u>-0.018</u>
2		0.095	-2.891	-1.841	-0.937	-2.475	-2.673	-3.291	-0.902
3			1.667	1.362	1.158	2.558	3.023	2.317	0.444
4				-0.281	-0.594	-0.903	-0.501	-2.405	-1.265
5					-0.309	-0.625	-0.356	-1.611	-0.975
6						-0.340	-0.294	-0.466	-0.664
7							<u>-0.003</u>	<u>0.007</u>	-0.604
8								<u>-0.001</u>	-1.894
9									3.178

Bildet man zur Beurteilung der Streckendifferenzen die Quotienten (5.4), ergibt sich (ohne Darstellung der Zwischenschritte), dass lediglich die in Tab. 5.8 unterstrichenen Werte zu einem „Signal-Rausch-Verhältnis“ von $q_{dl_{ij}} \leq 5$ führen. Nur diese Strecken können somit Bestandteil einer (oder mehrerer) kongruenter Punktgruppen sein. Diese Information wird nun, analog zu Tab. 5.3, in eine Kanten-Knoten-Matrix \mathbf{C} eingetragen, mit der sich aus (5.5) die in Tab. 5.9 dargestellte Knoten-Knoten-Matrix $\bar{\mathbf{C}}$ berechnen lässt.

Tab. 5.9: Knoten-Knoten-Matrix auf Grundlage der als unverändert angenommenen Strecken

	Pkt. 1	Pkt. 7	Pkt. 8	Pkt. 9	Pkt. 10
Pkt. 1	1	0	0	0	-1
Pkt. 7	0	2	-1	-1	0
Pkt. 8	0	-1	2	-1	0
Pkt. 9	0	-1	-1	2	0
Pkt. 10	-1	0	0	0	1

Wendet man den Suchalgorithmus aus Abschnitt 5.2 an, ist sofort ersichtlich, dass die Punkte 7, 8, 9 ein kongruentes Dreieck bilden können, da auf der Hauptdiagonalen der Wert „2“ steht und die übrigen Elemente mit dem Wert „-1“ belegt sind. Eine weitere kombinatorische Suche ist in diesem Beispiel nicht erforderlich. Um die Kongruenz dieses Teilnetzes zu testen, wird der Wert R_F berechnet. Mit den Werten

$$\mathbf{dl} = \begin{bmatrix} -0.0029 \text{ m} \\ 0.0007 \text{ m} \\ -0.0006 \text{ m} \end{bmatrix} \text{ und } \mathbf{Q}_{\text{dl}} = \begin{bmatrix} 0.5770 & -0.1762 & 0.1293 \\ -0.1762 & 1.0087 & 0.1746 \\ 0.1293 & 0.1746 & 1.1686 \end{bmatrix}$$

erhält man aus (2.11) den Wert $R_F = 5.5680 \cdot 10^{-5} \text{ m}^2$. Mit $h_F = 3$ und der empirischen Einheitsvarianz $s_0^2 = 1.2519 \cdot 10^{-4} \text{ m}^2$ kann man die Testgröße

$$T_F = \frac{R_F/h_F}{s_0^2} = \frac{5.5680 \cdot 10^{-5} \text{ m}^2 / 3}{1.2519 \cdot 10^{-4} \text{ m}^2} = 0.148 \quad (5.7)$$

aufstellen und mit dem Grenzwert der FISHER-Verteilung vergleichen. Mit den Freiheitsgraden $h_F = 3$ und $f = 56$ und einer Irrtumswahrscheinlichkeit von $\alpha = 5\%$ ergibt sich für $F_{h_F, f, 1-\alpha} = 2.77$ und somit kann die Nullhypothese der Kongruenz für diese Punktgruppe nicht verworfen werden.

Nachdem die erste kongruente Punktgruppe gefunden wurde, soll nun nach weiteren Gruppen gesucht werden. Dazu werden in Tab. 5.9 die Zeilen und Spalten der gefundenen Punkte gestrichen und man erhält die Knoten-Knoten-Matrix in Tab. 5.10.

Tab. 5.10: Knoten-Knoten-Matrix

	Pkt. 1	Pkt. 10
Pkt. 1	1	-1
Pkt. 10	-1	1

Es ist zu erkennen, dass es sich hierbei um die Strecke zwischen Punkt 1 und 10 handelt, die auf Kongruenz überprüft werden kann. Mit den Werten

$$\mathbf{dl} = [-0.0179 \text{ m}] \text{ und } \mathbf{Q}_{\text{dl}} = [0.5214]$$

erhält man aus (2.11) den Wert $R_F = 6.1481 \cdot 10^{-4} \text{ m}^2$. Mit $h_F = 1$ und der empirischen Einheitsvarianz $s_0^2 = 1.2519 \cdot 10^{-4} \text{ m}^2$ kann man die Testgröße

$$T_F = \frac{R_F/h_F}{s_0^2} = \frac{6.1481 \cdot 10^{-4} \text{ m}^2 / 1}{1.2519 \cdot 10^{-4} \text{ m}^2} = 4.911 \quad (5.8)$$

aufstellen. Mit den Freiheitsgraden $h_F = 1$, $f = 56$ und $\alpha = 5\%$ ergibt sich für $F_{h_F, f, 1-\alpha} = 4.01$. Mit $4.91 > 4.01$ kann die Nullhypothese der Kongruenz für diese Punktgruppe nicht angenommen werden

Abschließend lässt sich feststellen, dass in diesem Beispiel, in dem die sukzessive Einzelpunktanalyse versagt hat, die kongruente Punktgruppe 7, 8, 9 mit Hilfe der MSS-Methode gefunden werden konnte. Die Voruntersuchung mit Hilfe von Topologiematrizen hat dabei zu einer erheblichen Verringerung des Rechenaufwandes beigetragen. Hätte man tatsächlich alle Kombinationen gebildet, wären in diesem Beispiel bis zu

$$K = \binom{10}{9} + \binom{10}{8} + \dots + \binom{10}{3} = 967 \quad (5.9)$$

Untersuchungsschritte erforderlich gewesen. In diesem Beispiel hat sich sogar der „Idealfall“ der kombinatorischen Suche ergeben. Aufgrund der Voruntersuchungen wurden die „unmöglichen“ Kombinationen schon soweit eliminiert, dass die verbliebene Kombination direkt die Lösung darstellt.

6 Schlussbetrachtung und Ausblick

Mit der „MSS-Methode“ steht nun ein plausibles und unmittelbar einsichtiges Verfahren für die Datenanalyse zur Verfügung, das sich gerade im Hinblick auf die stetig zunehmende Rechnerleistung als echte Alternative zu den bisherigen Auswertestrategien anbietet. Die entscheidenden Vorteile sind:

- Der Anteil inkonsistenter Daten darf wesentlich größer sein, als bei den resistenten und robusten Schätzverfahren, die einen maximalen Bruchpunkt von 50% aufweisen.
- Aufgrund der kombinatorischen Suche ist die „MSS-Methode“ von Verschmierungseffekten und der Hebelpunkt-Problematik befreit.
- Die bekannten statistischen Testverfahren können angewendet werden.
- Die kombinatorische Suche kann problemlos in bereits bestehende Auswertesoftware integriert werden.
- Das Endergebnis kann direkt als endgültige Parameterschätzung verwendet werden.

Der Nachteil, dass unter Umständen ein hoher Rechenaufwand durch eine Vielzahl von Kombinationen entsteht, kann dadurch behoben werden, dass „unmögliche“ Kombinationen von vornherein aus der Untersuchung ausgeschlossen werden. Für die Kongruenzuntersuchung geodätischer Netze wurde diese Voruntersuchung detailliert ausgearbeitet und die MSS-Methode anhand eines numerischen Beispiels mit einem Anteil verschobener Punkte von 70% verifiziert.

Die Untersuchungen auf dem Gebiet der Kongruenzuntersuchung ermutigen dazu, die „MSS-Methode“ auch auf andere Fragestellungen der Datenanalyse zu übertragen, wie zum Beispiel der Suche nach groben Fehlern oder der automatischen Formenerkennung.

Literaturverzeichnis

- Berber, S., Hekimoglu, S.: What is the Reliability of Robust Estimators in Networks? In: First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS. A. Carosio, H. Kutterer (editors), Swiss Federal Institute of Technology Zurich (ETH), Institute of Geodesy and Photogrammetry, IGP-Bericht Nr. 295, S. 61-66, 2001.
- Caspary, W., Borutta, H.: Geometrische Deformationsanalyse mit robusten Schätzverfahren. Allgemeine Vermessungs-Nachrichten 93, Nr. 8-9, S. 315-326, 1986.
- Caspary, W.: Anmerkungen zum robusten Schätzen. Allgemeine Vermessungs-Nachrichten 103, Nr. 7, S. 287-289, 1996.
- Gründig, L.: Datenverwaltungskonzepte für typische Aufgaben aus der Ingenieurgeodäsie. Deutsche Geodätische Kommission, Reihe C, Nr. 336, München, 1988.
- Hekimoglu, S.: Local Reliability of the Least Median of Squares Method. In: First International Symposium on Robust Statistics and Fuzzy Techniques in Geodesy and GIS. A. Carosio, H. Kutterer (editors), Swiss Federal Institute of Technology Zurich (ETH), Institute of Geodesy and Photogrammetry, IGP-Bericht Nr. 295, S. 67-72, 2001.
- Hekimoglu, S., Koch, K.-R.: How can reliability of the robust methods be measured. Proceedings of the Third Turkish-German Joint Geodetic Days, Istanbul, pp. 179-196, 1999.
- Helmert, F. R.: Die Ausgleichsrechnung nach der Methode der kleinsten Quadrate. 3. Auflage, Teubner Verlag, Leipzig, 1924.
- Huber, P. J.: Robust Statistics. John Wiley & Sons, New York, 1981.
- Kampmann, G.: Auswertetechniken bei der überbestimmten Koordinatentransformation. BDVI-FORUM 3/1993, S. 139-152, 1993.
- Neitzel, F.: Identifizierung konsistenter Datengruppen am Beispiel der Kongruenzuntersuchung geodätischer Netze. Deutsche Geodätische Kommission, Reihe C, Nr. 565, München, 2004.
- Niemeier, W.: Ausgleichsrechnung. Walter de Gruyter, Berlin, New York, 2002.
- Petrovic, S.: Geometry of the Correlation Coefficient and its Application in Geodesy. Mitteilungen der geodätischen Institute der Technischen Universität Graz, Folge 71, 1991.
- Reinking, J.: Geodätische Analyse inhomogener Deformationen mit nichtlinearen Transformationsfunktionen. Deutsche Geodätische Kommission, Reihe C, Nr. 413, München, 1994.

Rousseeuw, P. J.: Least Median of Squares Regression. Journal of the American Statistical Association, Volume 79, Number 388, pp. 871-880, 1984.

Welsch, W., Heunecke, O., Kuhlmann, H.: Auswertung geodätischer Überwachungsmessungen. In: Möser, M., Müller, G., Schlemmer, H., Werner, H. (Hrsg.), Handbuch Ingenieurgeodäsie, Herbert Wichmann Verlag, Heidelberg, 2000.

Wicki, F.: Robuste Schätzverfahren für die Parameterschätzung in geodätischen Netzen. Institut für Geodäsie und Photogrammetrie an der Eidgenössischen Technischen Hochschule Zürich, Mitteilungen Nr. 67, 1999.

Anschrift des Autors

Dr.-Ing. Frank Neitzel
Institut für Geodäsie und
Geoinformationstechnik
Technische Universität Berlin
Straße des 17. Juni 135
D-10623 Berlin
Tel.: ++ 49-30/314-24054
Fax: ++ 49-30/314-21973
E-Mail: frank@mca.bv.tu-berlin.de