

AUTOMATISCHE ELIMINATION GROBER DATENFEHLER IM ERWEITERTEN BLOCKAUSGLEICHUNGSPROGRAMM PAT-M

H. Klein, Stuttgart

Manuelle Datenbereinigung

Jeder, der sich in der Praxis mit der Blockausgleichung beschäftigt, weiß, daß die Datenbereinigung einen großen Aufwand an Bearbeitungszeit erfordert und in den meisten Fällen nur von qualifizierten Bearbeitern ausgeführt werden kann.

Die Bereinigung wird durch folgende Probleme erschwert:

1. Ein grober Fehler zeigt sich nur teilweise in der Verbesserung der fehlerhaften Beobachtung, der Rest wird auf benachbarte Beobachtungen verteilt.
2. Grobe Fehler beeinflussen sich gegenseitig.
3. Die Elimination grober Datenfehler erfolgt vollständig, d.h. durch wirkliche Elimination oder die Einführung des Gewichts $P = 0$.

Diese Probleme machen eine schrittweise Datenbereinigung entsprechend der Größe der Fehler notwendig und führen zu Fehlinterpretationen und Fehlentscheidungen, die nur durch mehrere Probeläufe und jeweils vollständig neue Ausgleichungen zu vermeiden sind.

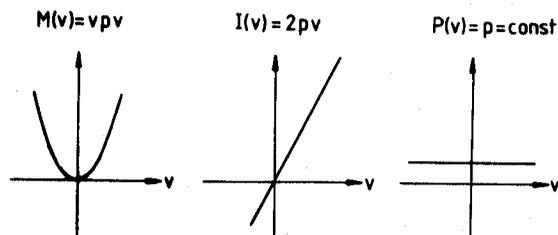
Insgesamt ist die Qualität der manuellen Fehlerbereinigung bei sorgfältiger Bearbeitung jedoch gut, denn alle überhaupt lokalisierbaren groben Fehler können auch lokalisiert werden, sie erfordert jedoch einen enormen Aufwand an Bearbeitungszeit, in der Praxis zwischen 10 Minuten und 30 Minuten pro Modell.

Automatische Fehlersuche

Eine Automatisierung des manuellen Verfahrens scheidet aus, da die Erfahrung des Bearbeiters nur sehr schwer zu formulieren ist und außerdem zu denselben oben genannten Problemen führt.

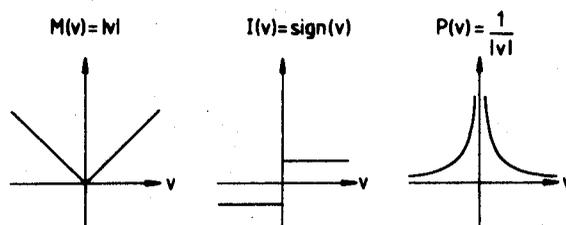
Betrachten wir zunächst die Methode der kleinsten Quadrate wie sie in der Blockausgleichung angewendet wird:

LEAST SQUARES ADJUSTMENT:



Jede Beobachtung erhält ein konstantes Gewicht p . Durch Einsetzen des konstanten Gewichts in die Minimumfunktion und Bilden der ersten Ableitung nach v erhält man die Einflußfunktion $I(v)$. Sie kann als Einfluß einer fehlerhaften Beobachtung auf das Ergebnis einer Ausgleichung interpretiert werden. Bei der Ausgleichung nach der Methode der kleinsten Quadrate ist der Einfluß proportional zur Größe des Fehlers, eine Eigenschaft, die für die Fehlersuche ungeeignet ist.

MINIMUM NORM METHOD:



Das Gewicht einer Beobachtung ist nicht konstant sondern wird nach jedem Interaktionsschritt als Funktion der Verbesserungen $P(v) = 1/|v|$ neu bestimmt. Dies führt zu der Einflußfunktion $I(v) = \text{sign}(v)$. Alle Beobachtungen haben unabhängig von der Größe des Fehlers den gleichen Einfluß auf das Ergebnis, d.h. der Einfluß eines groben Fehlers auf das Ergebnis ist immer noch zu groß.

Robuste Ausgleichung

Wählt man die Gewichtsfunktion so, daß der Einfluß eines groben Fehlers auf das Ergebnis mit anwachsendem Fehler zu Null wird, spricht man von Robusten Ausgleichungen (robust gegenüber dem Einfluß grober Fehler).

Nach langen Untersuchungen haben wir folgende Gewichtsfunktionen gewählt:

$$P(v_i) = P_i \cdot \frac{1}{1+(\alpha_i |v_i|)^d} \quad \text{für } Q \geq 2 \quad (1)$$

$$P(v_i) = P_i \cdot e^{-(\alpha_i |v_i|)^d} \quad \text{für } Q < 2 \quad (2)$$

und $\frac{P(v_i)}{P_i} \geq PMIN_i$

$$P(v_i) = P_i \frac{PMIN_i}{\alpha_i |v_i| + 10 \cdot PMIN_i} \quad \text{für } Q < 2 \quad (3)$$

für $\frac{P(v_i)}{P_i} < PMIN_i$

Hierbei wird $PMIN_i$ durch Gleichsetzen der Funktionen (2) und (3) bestimmt.

Hierbei ist:

$$\alpha_i = \frac{\sqrt{P_i}}{\sqrt{r_i} \cdot \hat{\sigma}_0 \cdot k} = \frac{1}{\hat{\sigma}_{v_i} \cdot k}$$

$$d = dmin + (dmax - dmin) \cdot \frac{82}{81+Q^4}$$

$$k = kmin + (kmax - kmin) \cdot \frac{101}{100+Q^2}$$

$$Q = \frac{\hat{\sigma}_0}{\sigma_{apriori}}$$

v_i = Verbesserung der Beobachtung i

r_i = Redundanzanteil der Beobachtung i

$\hat{\sigma}_{v_i}$ = Genauigkeit der Verbesserung v_i

P_i = a priori Gewicht der Beobachtung i

$\hat{\sigma}_0$ = geschätztes Sigma Null

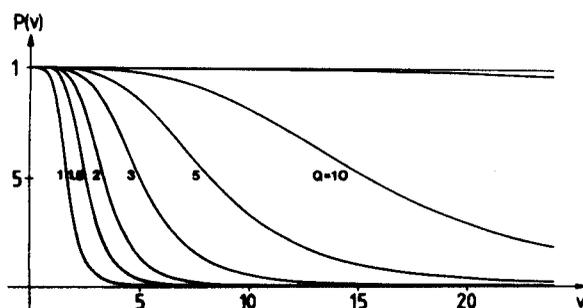
dmin, dmax, kmin, kmax = Grenzen für d und k

Die Gewichtsfunktionen (1), (2) und (3) führen zu den Einflußfunktionen (4), (5) und (6):

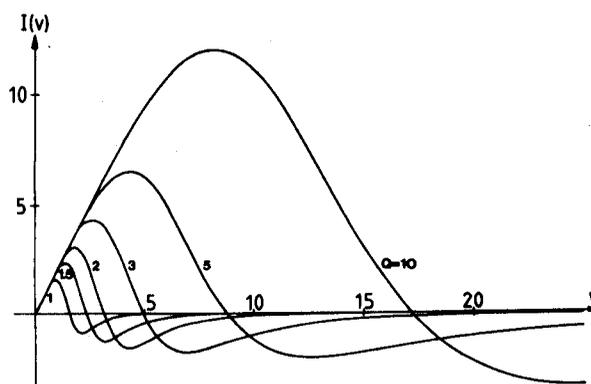
$$I(v_i) = \frac{|v_i|}{1+(\alpha_i|v_i|)^d} \cdot \left(2 - \frac{d \cdot (\alpha_i|v_i|)^d}{1+(\alpha_i|v_i|)^d}\right) \quad (4)$$

$$I(v_i) = |v_i| e^{-(\alpha_i|v_i|)^d} \cdot (2 - d(\alpha_i|v_i|)^d) \quad (5)$$

$$I(v_i) = \frac{|v_i| \cdot \text{PMIN}_i}{(\alpha_i|v_i| + 10 \cdot \text{PMIN}_i)^2} (\alpha_i|v_i| + 20 \cdot \text{PMIN}_i) \quad (6)$$



PATM-43: Gewichtsfunktionen



PATM-43: Einflußfunktionen

Eigenschaften

Der Wert Q konvergiert bei richtiger Wahl der Genauigkeiten der Beobachtungen mit fortschreitender Iterationszahl und schwindendem Einfluß grober Fehler gegen 1.

Die Exponentialfunktion (2) führt zu großen Werten von v_i zu schnell zu kleinen Gewichten, sie wird deshalb bei großen Werten von v_i durch (3) ersetzt. Für die ersten Iterationsschritte wird aus demselben Grund die hyperbolische Funktion (1) gewählt.

1. Über die Werte d und k passen sich die Funktionen im Laufe des Iterationsprozesses dem noch vorhandenen Einfluß grober Datenfehler und der Konvergenz des Iterationsprozesses an (Abhängigkeit von Q und $\hat{\sigma}_0$).

Die Gewichtsfunktionen verlaufen am Anfang des Iterationsprozesses relativ flach, sodaß Fehlentscheidungen verursacht durch grobe Fehler und verfälschte Anfangswerte (berechnet mit Kleinste-Quadrate-Ausgleichung) leichter rückgängig gemacht werden können. Gegen Ende des Iterationsprozesses werden die Kurven steiler, sodaß eine klarere Trennung der Beobachtungen mit groben Fehlern von den übrigen Beobachtungen erfolgt. Damit können auch grobe Datenfehler in der Größenordnung von mehreren Basislängen lokalisiert werden.

2. Mit $\hat{\sigma}_{v_i} = (\sqrt{r_i} \cdot \hat{\sigma}_0) / \sqrt{p_i}$ wird für die Berechnung des Gewichts einer einzelnen Beobachtung die Geometrie und damit die Genauigkeit der Verbesserung v_i berücksichtigt, sodaß auch Datenfehler im Grenzbereich der Lokalisierbarkeit noch aufgedeckt werden können.

3. Die robuste Ausgleichung kann auch als a priori Gewichtsschätzung interpretiert werden.

Das Verfahren ist iterativ, da die Verbesserungen nach einer ersten (kleinste-Quadrat-)Ausgleichung zur Berechnung von Anfangswerten noch nicht den wahren Fehlern entsprechen und verfälscht sind. Im Laufe des Iterationsprozesses konvergieren die Verbesserungen gegen die wahren Fehler und die Gewichte vom a priori Gewicht zum Gewicht welches dem wahren Fehler entspricht.

Programmablauf

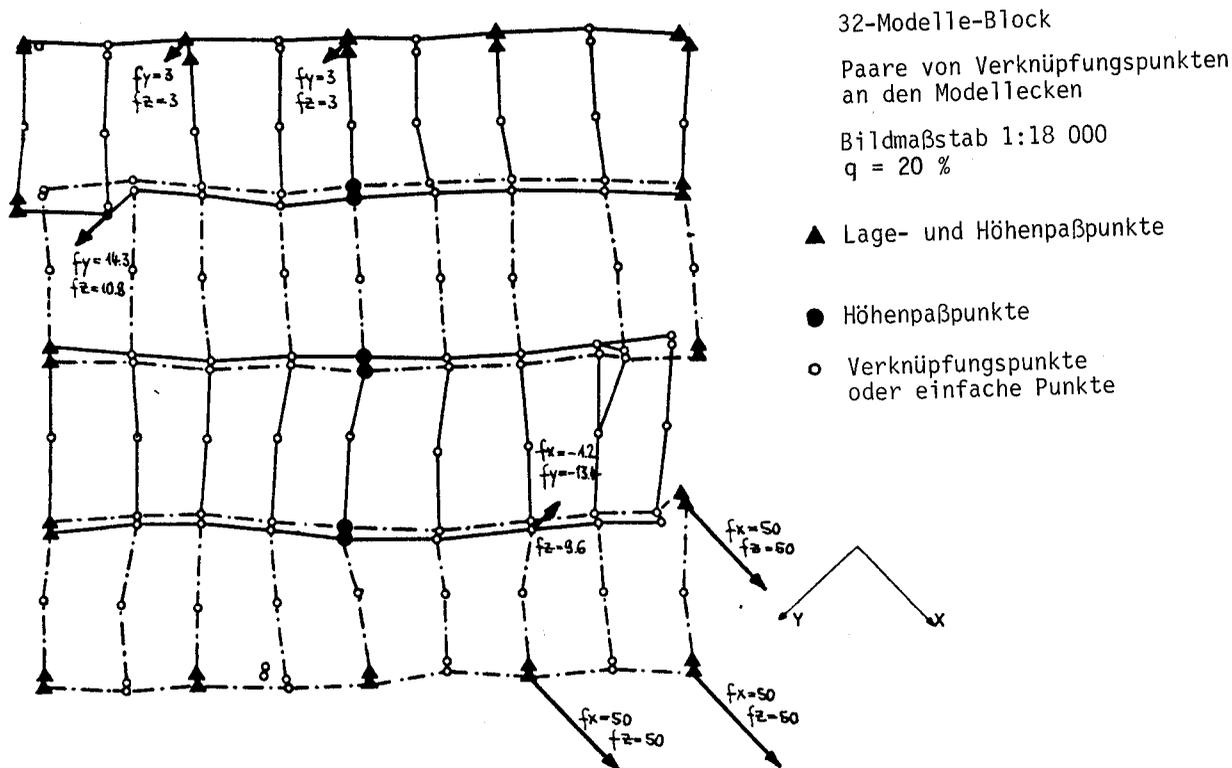
Die robuste Ausgleichung wird nur zur Fehlersuche verwendet.

Nach der Berechnung von Näherungswerten durch eine Lageblockausgleichung nach Helmert und einer reinen Verschiebungsausgleichung in der Höhe erfolgen robuste Iterationsschritte bis alle groben Fehler lokalisiert sind. Für die noch folgenden nicht robusten Iterationsschritte erhalten die lokalisierten groben Fehler das Gewicht $p = 10^{-10}$, alle übrigen ihr a priori in die Ausgleichung eingeführtes Gewicht.

Die Anzahl aller Iterationsschritte (Lage und Höhe getrennt) ist abhängig von der Anzahl und Größe der groben Fehler, aber selbst in ungünstigsten Fällen höchstens 3mal so groß wie bei einer einzigen Ausgleichung ohne Fehlersuche, sodaß die Rechenzeit höchstens 2 1/2mal größer ist als bei einer Ausgleichung ohne Fehlersuche. Wiederholungsausgleichungen werden jedoch weitgehend vermieden.

Das folgende Beispiel zeigt einen praktischen Block, der manuell bereinigt wurde und an dem die in der Übersicht gezeigten Fehler angebracht wurden.

Die Auszüge aus den Ausdrucken (Seiten 5, 6 und 7) stellen das Ergebnis der Ausgleichung ohne Fehlersuche dem Ergebnis der Ausgleichung mit automatischer Fehlersuche gegenüber.



Zum Beispiel: Ausschnitt aus dem Printout
 nach Ausgleichung mit automatischer Fehlersuche:
 Ende der robusten Ausgleichung und Elimination
 der groben Fehler

END OF ERROR DETECTION IN ELEVATION
 SIGMA REACHED = 0.2186

VERTICAL CONTROL POINT		327	PC 2	ELIMINATED	V=	49.933
VERTICAL CONTROL POINT		406	PC 2	ELIMINATED	V=	50.019
VERTICAL CONTROL POINT		412	PC 2	ELIMINATED	V=	50.381
VERTICAL CONTROL POINT		3002	PC 2	ELIMINATED	V=	2.760
MODEL	20872089	POINT	119	TP 3	ELIMINATED	V= 9.731
MODEL	40784080	POINT	316	TP 4	ELIMINATED	V= 9.457
MODEL	40784076	POINT	316	TP 4	ELIMINATED	V= 9.777

 ITERATION STEP 9.....HORIZONTAL ADJUSTMENT

ITERATION STEP FOR ERROR DETECTION

MAXIMUM NUMBER OF MODELS PER SUBMATRIX =10

NUMBER OF HYPERROWS = 4
 MAXIMUM NUMBER OF SUBMATRICES PER HYPERROW OF DIAGONAL BAND = 2

MAXIMUM COORDINATE DIFFERENCES OF TRANSFORMED COORDINATES
 IN THE TERRAIN SYSTEM

IN X AT POINT	227	0.121
IN Y AT POINT	131	0.207
IN Z AT POINT	392	0.001

END OF ERROR DETECTION IN PLANIMETRY
 SIGMA REACHED = 0.1526

HORIZONTAL CONTROL POINT		131	PC 2	ELIMINATED	V=	8.009
HORIZONTAL CONTROL POINT		327	PC 2	ELIMINATED	V=	49.957
HORIZONTAL CONTROL POINT		406	PC 2	ELIMINATED	V=	49.871
HORIZONTAL CONTROL POINT		412	PC 2	ELIMINATED	V=	49.636
HORIZONTAL CONTROL POINT		3002	PC 2	ELIMINATED	V=	3.083
MODEL	30083006	POINT	310	TP 4	ELIMINATED	V= 0.522

 ITERATION STEP 10.....HORIZONTAL ADJUSTMENT

MAXIMUM NUMBER OF MODELS PER SUBMATRIX =10

NUMBER OF HYPERROWS = 4
 MAXIMUM NUMBER OF SUBMATRICES PER HYPERROW OF DIAGONAL BAND = 2

MAXIMUM COORDINATE DIFFERENCES OF TRANSFORMED COORDINATES
 IN THE TERRAIN SYSTEM

IN X AT POINT	306	0.117
IN Y AT POINT	229	0.120
IN Z AT POINT	392	0.001

 ITERATION STEP 11.....VERTICAL ADJUSTMENT

ZUSAMMENFASSUNG

Eines der Hauptprobleme bei der praktischen Durchführung der Blockausgleichung ist die Elimination grober Datenfehler. Im erweiterten Blockausgleichungsprogramm PAT-M wird dieses Problem mit Hilfe der Methode der robusten Schätzer automatisch gelöst.

Bei diesem Verfahren werden im Laufe eines Iterationsprozesses die Gewichte der einzelnen Beobachtungen sukzessive den Datenfehlern angepaßt. Nach erfolgter Konvergenz sind grobe Datenfehler ohne Einfluß auf das Ergebnis und können automatisch eliminiert werden.

Im Vortrag wird die Wahl der verwendeten Gewichts- und Einflußfunktionen erläutert, sowie die Leistungsfähigkeit als auch die Grenzen des Verfahrens mit Hilfe von Beispielen aufgezeigt.

AUTOMATIC ELIMINATION OF GROSS DATA ERRORS IN THE EXTENDED PAT-M BLOCK ADJUSTMENT PROGRAM

Abstract

One of the major problems in practical block adjustment is the elimination of gross data errors. In the extended PAT-M block adjustment program this problem is solved automatically by means of the robust estimation method.

With this method, the weights of the individual observations are successively matched to the data errors during iteration. If convergence is established, gross data errors do not affect the result and can be eliminated automatically.

The paper describes the selection of the weighting functions used as well as the potential and limitations of the method by means of examples.

ELIMINATION AUTOMATIQUE DES ERREURS DE DONNEES GRAVES DANS LE PROGRAMME DE COMPENSATION DE BLOCS ELARGI PAT-M

Résumé

L'un des principaux problèmes qui se posent lors de l'exécution pratique de la compensation de blocs, réside dans l'élimination des erreurs de données graves. Ce problème est dorénavant résolu de façon automatique dans la version élargie du programme de compensation de blocs PAT-M, par la méthode des estimateurs.

Le procédé utilisé consiste à adapter successivement aux erreurs de données les valeurs pondérales des différentes observations en appliquant une méthode itérative. Une fois la convergence réalisée, les erreurs de données graves n'ont plus d'influence sur les résultats et peuvent être éliminées automatiquement.

L'exposé précise les différentes fonctions de poids et d'influence qui ont été choisies à cette fin, ainsi que les possibilités et les limites de ce procédé en présentant des exemples précis.

ELIMINACION AUTOMATICA DE ERRORES GRAVES DE DATOS EN EL PROGRAMA AMPLIADO DE COMPENSACION DE BLOQUES PAT-M

Resumen

La eliminación de los errores graves de datos constituye uno de los principales problemas que plantea la realización práctica de la compensación de bloques. El programa ampliado de compensación de bloques lo resuelve con ayuda del método de la estimación poco sensible ("robust estimation").

Al aplicar este método, el peso de las distintas observaciones se adapta sucesivamente y durante un proceso de iteración a los errores de los datos. Una vez conseguida la convergencia, los errores graves de datos ya no ejercen influencia en el resultado y pueden eliminarse de modo automático.

En la conferencia se explica la elección de las funciones de peso y influencia y se demuestran el rendimiento a la vez que las limitaciones del método, recurriendo a ejemplos.

Dipl.-Math. Hermann Klein
Universität Stuttgart, Institut für Photogrammetrie,
Keplerstraße 11, 7000 Stuttgart 1