Zur Lösung des nichtlinearen Gauß-Markov-Modells

Lothar Lenzmann und Enno Lenzmann

Zusammenfassung
Die Lösung des nichtlinearen Gauß-Markov-Modells ist auf zwei unterschiedlichen Wegen möglich; einmal dadurch, dass man die nichtlinearen Funktionen vorab linearisiert und dann im linearen Modell die Minimumsaufgabe der Methode der kleinsten Quadrate löst. Der andere Weg ergibt sich, indem man zuerst die Minimumsaufgabe löst und dann die Auflösung des so erhaltenen nichtlinearen Gleichungssystems numerisch vornimmt. Dieses Vorgehen stellt zudem mit der Hesse-Matrix ein Mittel zur Verfügung, mit dem man die Existenz eines lokalen Minimums überprüfen kann. Darüber hinaus gewinnt man über die Hesse-Matrix einen Iterationsalgorithmus, der deutlich schneller konvergiert als der herkömmliche.

Summary
There are two different ways for solving problems of nonlinear adjustment. In the first way, one linearizes all nonlinear functions and solves the minimization problem of least squares for this approximated, linear model. This usual procedure leads to correct solutions for Gauß-Markov-Models. In this article, we present an alternative way by taking into account the nonlinear minimization problem without any previous linearizations. The resulting nonlinear equations then are treated by numerics leading to the correct solution within desired precision. In the case of the Gauß-Markov-Model, our method incorporates the computation of the Hessian, and we obtain an algorithm that rapidly converges and provides a practical test for minima.

1 Einleitung

Die Auswertung der nichtlinearen Ausgleichungsprobleme kann auf zwei unterschiedlichen Wegen vorgenommen werden. Beim Standardverfahren (vgl. z. B. Helmert 1924, Wolf 1968, Höpcke 1980) linearisiert man zunächst die nichtlinearen Funktionsgleichungen und löst dann die Minimumsaufgabe der Methode der kleinsten Quadrate im linearen Modell. Diese Vorgehensweise wird als Gauß-Newton-Verfahren bezeichnet. Auf ähnlicher Grundlage basierend, ermöglicht auch der Levenberg-Marquardt-Algorithmus die Lösung des nichtlinearen Gauß-Markov-Modells. Obwohl beide Verfahren auf Grund der Linearisierung der Beobachtungsgleichungen streng genommen nur das lineare Ersatzproblem minimieren, führen sie dennoch auf iterativem Wege zur numerisch exakten Lösung. Sie zählen zu den lokalen Optimierungsverfahren (vgl. Mautz 2001).

Um das nichtlineare Gauß-Helmert-Modell und die nichtlineare Ausgleichung bedingter Beobachtungen zu lösen, wird in Lenzmann et al. (2004) ein anderer Weg beschritten. Und zwar wird zunächst die Lösung der Minimumsaufgabe gemäß der Methode der kleinsten Quadrate vorgenommen und erst darauf mittels numerischer Verfahren das so erhaltene nichtlineare Gleichungssystem nach den gesuchten Schätzern aufgelöst. Diese Vorgehensweise wird hier auch auf das nichtlineare Gauß-Markov-Modell angewendet. Dabei ergeben sich einige – auch für die praktische Berechnung – nützliche Erkenntnisse.

2 Die Minimumsaufgabe im Gauß-Markov-Modell

Das Gauß-Markov-Modell wird in der Literatur auch als die Ausgleichung vermittelnder Beobachtungen bezeichnet. Dieser Ausgleichungsansatz spielt in der Praxis eine überragende Rolle, weil er programmiertechnisch leicht zu beherrschen ist und weil durch geschickte Wahl der Unbekannten sich alle Ausgleichungsprobleme in dieser Weise darstellen lassen. Zudem muss die Minimumsaufgabe nicht über den Umweg der Lagrange-Funktion formuliert werden. Wir wählen die Bezeichnungen in Anlehnung an die DIN 18 709, Teil 4. Mit φ_i als der Funktion, die die i-te Beobachtung L_i aus dem Unbekanntenvektor $\boldsymbol{X} = (X_k)$ mit $k = 1 \ldots u$ berechnet, und mit v_i als i-ter Verbesserung lautet das Ausgleichungsmodell

$$L_i + v_i = \varphi_i(\boldsymbol{X}), \quad i = 1 \ldots n. \tag{1}$$

Hierbei ist n die Anzahl der Beobachtungen. Aus (1) ergeben sich die in v_i linearen Verbesserungsgleichungen

$$v_i = \varphi_i(\boldsymbol{X}) - L_i. \tag{2}$$

Mit $\boldsymbol{v} = (v_i)$ ist gemäß der Methode der kleinsten Quadrate die Funktion

$$\Omega := \boldsymbol{v}^{\mathrm{T}}\mathbf{Q}^{-1}\boldsymbol{v} = (\boldsymbol{\varphi}(\boldsymbol{X}) - \boldsymbol{L})^{\mathrm{T}}\mathbf{Q}^{-1}(\boldsymbol{\varphi}(\boldsymbol{X}) - \boldsymbol{L}). \tag{3}$$

zu minimieren. Hierbei gilt $\boldsymbol{\varphi} = (\varphi_i)$, und $\mathbf{Q}^{-1}$ ist die symmetrische Gewichtsmatrix zu $\boldsymbol{L}$. Mit der Jacobimatrix (Modellmatrix)

$$\mathbf{A} := \frac{\partial \boldsymbol{\varphi}(\boldsymbol{X})}{\partial \boldsymbol{X}} \tag{4}$$

lautet das Gleichungssystem, das der Minimierer $\hat{\boldsymbol{X}}$ erfüllen muss, wie folgt:

$$\frac{\partial \Omega}{\partial \boldsymbol{X}} = 2\mathbf{A}^{\mathrm{T}}\mathbf{Q}^{-1}(\boldsymbol{\varphi}(\hat{\boldsymbol{X}}) - \boldsymbol{L}) = 0. \tag{5}$$

Die symmetrische Hesse-Matrix mit

$$\mathbf{H} := \frac{\partial^2 \Omega}{\partial \boldsymbol{X}^2} \tag{6}$$

erlaubt zu überprüfen, ob das nichtlineare Gleichungssystem (5) ein lokales Minimum beschreibt. Bei nichtlinearen Funktionen gibt es leider kein allgemeingültiges Kriterium, um ein globales Minimum nachzuweisen. Allerdings ist im Falle der Konvexität der Funktion Ω ein lokales Minimum zugleich das globale Minimum von Ω. Ein numerisches Verfahren zur Bestimmung eines globalen Optimums wird in Xu (2003) angegeben.

3 Numerische Lösung

Bei der numerischen Auflösung des exakten Gleichungssystems (5) wird man im Allgemeinen iterativ vorgehen müssen. Hier werden zwei Wege beschrieben, die man beschreiten kann.

3.1 Lösung mittels Approximation von $\boldsymbol{\varphi}(\boldsymbol{X})$

Im Folgenden sei die Anwendung des Newton-Verfahrens auf $\boldsymbol{\varphi}(\boldsymbol{X})$ in der Bestimmungsgleichung (5) dargelegt. Wir bezeichnen die Ergebnisse der j-ten Iteration mit $\boldsymbol{X}^{(j)}$. Als Startwerte $\boldsymbol{X}^{(0)}$ nimmt man Näherungswerte. Die Jacobimatrix an der Stelle $\boldsymbol{X}^{(j)}$ sei $\mathbf{A}^{(j)}$. Das Newton-Verfahren (vgl. z. B. Bronstein et al. (1993), S. 608) schreibt für $\boldsymbol{\varphi}(\boldsymbol{X})$ folgende lineare Approximation vor:

$$\mathbf{A}^{(j)} \cdot (\boldsymbol{X}^{(j+1)} - \boldsymbol{X}^{(j)}) + \boldsymbol{\varphi}(\boldsymbol{X}^{(j)}) = 0. \tag{7}$$

Mit dem gekürzten Beobachtungsvektor (vgl. DIN 18 709, Teil 4)

$$-\boldsymbol{l}^{(j)} = \boldsymbol{\varphi}(\boldsymbol{X}^{(j)}) - \boldsymbol{L} \tag{8}$$

gewinnt man auf Grund von (7) aus (5) das geläufige lineare Gleichungssystem

$$\begin{aligned}&(\mathbf{A}^{(j)})^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{A}^{(j)} \cdot (\boldsymbol{X}^{(j+1)} - \boldsymbol{X}^{(j)})\\ &-(\mathbf{A}^{(j)})^{\mathrm{T}}\mathbf{Q}^{-1} \cdot \boldsymbol{l}^{(j)} = 0\end{aligned} \tag{9}$$

zur Iteration des Unbekanntenvektors $\hat{\boldsymbol{X}}$.

3.2 Lösung mittels Approximation von $\partial\Omega/\partial\boldsymbol{X}$

Der zweite Lösungsweg ergibt sich, indem man das Newton-Verfahren zur Nullstellenbestimmung auf die Bestimmungsgleichung (5) direkt anwendet. Die lineare Approximation lautet dann an der Stelle $\boldsymbol{X}^{(j)}$

$$\left.\frac{\partial^2 \Omega}{\partial \boldsymbol{X}^2}\right|_{\boldsymbol{X}=\boldsymbol{X}^{(j)}} \cdot (\boldsymbol{X}^{(j+1)} - \boldsymbol{X}^{(j)}) + \left.\frac{\partial \Omega}{\partial \boldsymbol{X}}\right|_{\boldsymbol{X}=\boldsymbol{X}^{(j)}} = 0. \tag{10}$$

Mit der Jacobimatrix $\mathbf{A}^{(j)}$ zu $\boldsymbol{\varphi}(\boldsymbol{X})$ und der symmetrischen Hesse-Matrix $\mathbf{H}^{(j)}$ zu Ω – beide an der Stelle $\boldsymbol{X}^{(j)}$ –, der Bestimmungsgleichung (5) und dem gekürzten Beobachtungsvektor (8) gewinnt man daraus den Iterationsalgorithmus für $\hat{\boldsymbol{X}}$ zu

$$\mathbf{H}^{(j)} \cdot (\boldsymbol{X}^{(j+1)} - \boldsymbol{X}^{(j)}) - 2(\mathbf{A}^{(j)})^{\mathrm{T}}\mathbf{Q}^{-1} \cdot \boldsymbol{l}^{(j)} = 0. \tag{11}$$

4 Beispiel

Wir wählen das Zahlenbeispiel aus Lenzmann et al. (2004), um zu demonstrieren, dass man das Gauß-Markov-Modell nicht nur auf die herkömmliche Weise mittels Normalgleichungsmatrix numerisch lösen kann. Man gewinnt mit Hilfe der Hesse-Matrix einen schneller konvergierenden Algorithmus. Die im Konvergenzpunkt positiv definite Hesse-Matrix ist zudem ein Beweis dafür, dass dort zumindest ein lokales Minimum vorliegt. Es handelt sich bei diesem Beispiel um eine Parabel durch den Nullpunkt, bei der an zwei Stellen gleich genaue Ordinaten- und Abszissenwerte gemessen wurden. Mit den Unbekannten a, x_1 und x_2 lauten die Funktionen φ_i $(i = 1 \ldots 4)$, nach denen die Beobachtungen l_i berechnet werden,

$$\varphi_1 = x_1, \quad \varphi_2 = x_2, \quad \varphi_3 = ax_1^2, \quad \varphi_4 = ax_2^2, \tag{12}$$

und das System der Verbesserungsgleichungen mit den Verbesserungen v_i gewinnt die Gestalt

$$\begin{aligned}&v_1 = x_1 - l_1, \quad v_2 = x_2 - l_2,\\ &v_3 = ax_1^2 - l_3, \quad v_4 = ax_2^2 - l_4.\end{aligned} \tag{13}$$

Der Vektor der gleichgewichteten Beobachtungen ist

$$\boldsymbol{L}^{\mathrm{T}} = \begin{bmatrix} 2.5 & 4.0 & 4.8 & 5.0 \end{bmatrix}. \tag{14}$$

Gemäß der Forderung der Methode der kleinsten Quadrate ist zur Bestimmung der Schätzer die Funktion $\Omega = \boldsymbol{v}^{\mathrm{T}}\boldsymbol{v}$ aus (3) zu minimieren. Die Jacobimatrix (Modellmatrix) $\mathbf{A}$ zu $\boldsymbol{\varphi} = (\varphi_i)$ und die Normalgleichungsmatrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ lauten

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ x_1^2 & 2ax_1 & 0 \\ x_2^2 & 0 & 2ax_2 \end{bmatrix} \tag{15}$$

und

$$\mathbf{A}^{\mathrm{T}}\mathbf{A} = \begin{bmatrix} x_1^4 + x_2^4 & 2ax_1^3 & 2ax_2^3 \\ 2ax_1^3 & 1 + 4a^2x_1^2 & 0 \\ 2ax_2^3 & 0 & 1 + 4a^2x_2^2 \end{bmatrix}. \tag{16}$$

Wie die Berechnung zeigt, ist die Hesse-Matrix von Ω in diesem Beispiel gegeben durch

$$\mathbf{H} = 2\mathbf{A}^{\mathrm{T}}\mathbf{A} + 2\sum_{i=1}^{4} v_i \mathbf{H}_i, \tag{17}$$

wobei $\mathbf{H}_i$ die Hesse-Matrix der jeweiligen Funktion φ_i ist. Explizit erhält man in unserem Fall

$$\mathbf{H} = 2\begin{bmatrix} x_1^4 + x_2^4 & 2ax_1^3 + v_3 2x_1 & 2ax_2^3 + v_4 2x_2 \\ 2ax_1^3 + v_3 2x_1 & 1 + 4a^2x_1^2 + v_3 2a & 0 \\ 2ax_2^3 + v_4 2x_2 & 0 & 1 + 4a^2x_2^2 + v_4 2a \end{bmatrix}. \tag{18}$$

Mit den Startwerten $a^0 = 0.5$; $x_1^0 = 2.5$; $x_2^0 = 4.0$ und $v_i^0 = 0.0$ gewinnt man über den herkömmlichen Algorithmus mittels Normalgleichungsmatrix gemäß (9) nach 15 Iterationen das numerisch auf 15 Stellen stabile Endergebnis. Über (11) ergibt sich mit Hilfe der Hesse-Matrix bereits nach sechs Iterationen dasselbe Endergebnis. Zur Reduktion wird das Cholesky-Verfahren benutzt, das eine positiv definite Matrix $\mathbf{H}$ notwendig voraussetzt. Die erfolgreiche Berechnung ist deshalb zugleich ein Beleg dafür, dass es sich im Konvergenzpunkt zumindest um ein lokales Minimum handelt. Für die Unbekannten erhält man die folgenden Schätzwerte:

$$\begin{aligned} \hat{a} &= 0.456218634812259; \\ \hat{x}_1 &= 3.16489918245210; \\ \hat{x}_2 &= 3.37683009883002. \end{aligned} \tag{19}$$

5 Schlussbemerkung

In aller Regel wird in den Lehrbüchern für alle Formen nichtlinearer Ausgleichungsaufgaben der Weg über das Vorablinearisieren der Bedingungs- bzw. Beobachtungsgleichungen gewählt. Im Falle vermittelnder Beobachtungen (Gauß-Markov-Modell) gewinnt man so einen Iterationsalgorithmus, mit dem man das korrekte Zahlenergebnis beliebig genau berechnen kann. Für die Ausgleichung bedingter Beobachtungen und den Allgemeinfall der Ausgleichungsrechnung (Gauß-Helmert-Modell) ist das nicht möglich, wenn man die in den Lehrbüchern geläufige Linearisierung wählt. Das Standardverfahren vernachlässigt nämlich höhere als die erste Potenz der Verbesserungen v (vgl. Helmert 1924, S. 248). Eine Linearisierung, wie sie in Lenzmann et al. (2004) angegeben wird, führt allerdings ebenfalls zum korrekten Ergebnis (vgl. auch Kupferer 2005 und Lenzmann et al. 2005).

In diesem Beitrag wird zunächst die Minimumsaufgabe nach der Methode der kleinsten Quadrate streng gelöst, und zwar ohne das Hilfsmittel der Linearisierung. Für die Berechnung der Zahlenergebnisse wird ein Verfahren der numerischen Mathematik notwendig. Das ist z. B. das Newton-Verfahren. Dieser Weg führt bei entsprechendem Vorgehen direkt zu dem bekannten Iterationsverfahren für das Gauß-Markov-Modell. Er kann aber auch so aufbereitet werden, dass er einen deutlich schneller konvergierenden Lösungsalgorithmus ermöglicht und quasi nebenbei den Nachweis des lokalen Minimums erlaubt. Die schnellere Konvergenz wird durch den zusätzlichen Aufwand zur Ermittlung der Hesse-Matrix erkauft. Allerdings deutet sich bereits im dargelegten Zahlenbeispiel an, dass man die Hesse-Matrix durch die Normalgleichungsmatrix zuzüglich additiver Terme für die Berechnung aufbereiten kann (vgl. (17)). Es werden also erheblich mehr Summationsschritte notwendig sein als beim herkömmlichen Verfahren; im Gegenzug ergeben sich aber deutlich weniger Iterationen, und man erspart sich somit eine entsprechende Anzahl von rechenaufwändigen Inversionen.

Literatur

Bronstein, I. N., Semendjajew, K. A., Musiol, G., Mühlig, H.: Taschenbuch der Mathematik. Verlag Harri Deutsch, Thun und Frankfurt am Main, 1993.

DIN 18 709: Begriffe, Kurzzeichen und Formelzeichen im Vermessungswesen, Teil 4, 1984.

Helmert, F. R.: Die Ausgleichungsrechnung nach der Methode der kleinsten Quadrate mit Anwendungen auf die Geodäsie, Physik und die Theorie der Messinstrumente. Dritte Auflage. Teubner, Leipzig, 1924.

Höpcke, W.: Fehlerlehre und Ausgleichsrechnung. Walter de Gruyter, Berlin, New York 1980.

Kupferer, S.: Zur korrekten Linearisierung von nichtlinearen GH-Modellen. AVN 111, S. 394–396, 2004.

Lenzmann, L., Lenzmann, E.: Strenge Auswertung des nichtlinearen Gauß-Helmert-Modells. AVN 111. S. 68–73, 2004.

Lenzmann, L., Lenzmann, E.: Stellungnahme zu dem Beitrag »Zur korrekten Linearisierung von nichtlinearen GH-Modellen«. AVN 111. S. 114, 2005.

Mautz, R.: Zur Lösung nichtlinearer Ausgleichungsprobleme bei der Bestimmung von Frequenzen in Zeitreihen. Dissertation. Fachbereich 9, TU Berlin, 2001.

Wolf, H.: Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. Dümmler, Bonn 1968.

Xu, P.: A hybrid global optimization method: The multi-dimensional case. Journal of Computational and Applied Mathematics, Volume 115, Issue 2, June 2003, p. 423–446.

Anschrift der Autoren

Prof. Dr.-Ing. Lothar Lenzmann
Kurt-Weill-Weg 3, 45657 Recklinghausen
llenzmann@t-online.de

Dr. sc. math. Enno Lenzmann
M. I. T., Department of Mathematics
77 Massachusetts Avenue, Office 2-230
Cambridge, MA 02139-4307, USA
lenzmann@math.mit.edu