AUSGEWÄHLTE KAPITEL AUS "NUMERISCHE METHODEN IN DER PHYSIK"

H. Sormann SS 2007

Thema ohne Prüfungscharakter: Least-Square-Erweiterungen

1. Einleitung
2. Erweiterung 1: y- und x-Werte fehlerbehaftet
3. Erweiterung 2: Einbeziehung von Zwangsbedingungen

Einleitung

Im 4. Hauptkapitel meines Vorlesungsskriptums "Numerische Methode in der Physik" behandle ich die "Least-Squares-Approximation". Eine der bekanntesten Verfahren für nicht-lineare Modellfunktionen ist das Gauss-Newton-Marquardt- oder Levenberg-Marquardt-Verfahren (s. Abschnitte 4.5.3-4.5.7).

Im folgenden stelle ich nun zwei Erweiterungen der GNM-Methode vor, die Sie relativ leicht in die vorliegenden Programme einbauen können.

Erweiterung 1: y- <u>und</u> x-Werte fehlerbehaftet

Bei der - nennen wir's mal so - üblichen Anwendung des LSQ-Verfahrens wird von folgender Situation ausgegangen: der Experimentator liefert eine Anzahl von Meßwerten y_k ($k = 1, \ldots, n$) für die Argumentwerte x_k. Die experimentelle Tätigkeit wird nun gewöhnlich so beschrieben: *man mißt zu jedem gewählten x_k eine physikalische Größe y_k, wobei diese Meßgröße eine statistische Unsicherheit enthält, sie man durch die jeweilige Standardabweichung beschreibt*:

Jeder Meßwert y_k hat die Standardabweichung σ_k^y.

Die Werte x_k, *an denen die Messungen durchgeführt werden*, werden ohne weitere Diskussion als <u>exakt</u> angenommen. Darauf beruht das konventionelle LSQ-Verfahren.

Obwohl die obige Vorgangsweise in vielen Fällen gerechtfertigt ist, kann i.a. aber keineswegs davon ausgegangen werden, daß die Argumentwerte x_k "fehlerfrei" sind, denn diese sind ja in der Regel auch physikalische Größen, die durch einen Meßprozess erhalten werden. Es gilt also auch der Satz:

Jeder Amplitudenwert x_k hat die Standardabweichung σ_k^x.

Das konventionelle LSQ-Verfahren geht also von der Annahme

$$\sigma_k^x << \sigma_k^y \qquad \text{für alle} \quad k = 1, \ldots, n \tag{1}$$

aus, die - wie gesagt - sehr oft gut erfüllt ist.

Die Frage, vor der man nun steht, lautet: *Wie ist die LSQ-Grundgleichung*

$$\chi^2 = \sum_{k=1}^{n} \left[\frac{y_k - f(x_k^0, \mathbf{a})}{\sigma_k^y} \right]^2 \quad \rightarrow \quad \text{Min.}. \tag{2}$$

zu variieren, wenn die Bedingung (1) <u>nicht</u> erfüllt ist, d.h. wenn (zumindest für <u>ein</u> k

$$\sigma_k^x \geq \sigma_k^y \tag{3}$$

gilt?

Das Verfahren, welches in diesem Fall angewendet werden kann, heißt *Effective Variance Method*. Die folgende Darstellung dieser Methode, habe ich einem *paper* von J. Orear entnommen[1].

Der erste Schritt ist die Erweiterung des Least-Squares-Ansatzes in bezug darauf, daß nicht nur die Modellfunktion $f(x, \mathbf{a})$ an die gemessenen y_k angepaßt werden muß, sondern auch die unbekannten (und daher ebenfalls als Fit-Parameter anzusehenden) Erwartungswerte x_k^0 an die gemessenen x_k:

$$\chi^2 = \sum_{k=1}^{n} \left\{ \left[\frac{x_k - x_k^0}{\sigma_k^x} \right]^2 + \left[\frac{y_k - f(x_k^0, \mathbf{a})}{\sigma_k^y} \right]^2 \right\} \quad \rightarrow \quad \text{Min.}. \tag{4}$$

[1] J. Orear, *Least squares when both variables have uncertainties*, Am. J. Phys. **50**, 912 (1982).

Beachten Sie, daß im Ansatz (4) $2n$ Datenwerte einzugeben sind, denen $n+q$ Fit-Parameter gegenüberstehen. Die Anzahl der Freiheitsgrade beträgt also

$$2n - (n+q) = n - q\,,$$

ist also gleich wie beim konventionellen LSQ-Ansatz.

Um die Fit-Parameter x_j^0 zu erhalten, wird Glg. (4) nach diesen Größen abgeleitet und Null gesetzt:

$$\frac{\partial \chi^2}{\partial x_j^0} = -2\left(\frac{x_j - x_j^0}{\sigma_j^{x2}}\right) - 2\left(\frac{y_j - f(x_j^0, \mathbf{a})}{\sigma_j^{y2}}\right)\left(\frac{\partial f(x_j^0, \mathbf{a})}{\partial x_j^0}\right) \stackrel{!}{=} 0\,. \tag{5}$$

Nun erfolgt eine weitere Linearisierung der Modellfunktion, diesmal aber in bezug auf das Argument x an der Stelle x_j^0:

$$f(x_j, \mathbf{a}) \approx f(x_j^0, \mathbf{a}) + \left(\frac{\partial f(x, \mathbf{a})}{\partial x}\right)_{x=x_j^0} (x_j - x_j^0)\,. \tag{6}$$

Mit der Abkürzung

$$\left(\frac{\partial f(x, \mathbf{a})}{\partial x}\right)_{x=x_j^0} \equiv f_j^{x0}$$

ergibt sich Einsetzen von Glg. (6) in Glg. (5) der Ausdruck

$$\left(\frac{x_j - x_j^0}{\sigma_j^{x2}}\right) + \frac{f^{x0}}{\sigma_j^{y2}}\left[y_j - f(x_j, \mathbf{a}) + \left(x_j - x_j^0\right) f_j^{x0}\right] = 0\,.$$

Löst man diese Gleichung nach den x auf, erhält man

$$x_j - x_j^0 = \frac{\sigma_j^{x2} f_j^{x0}\left[f(x_j, \mathbf{a}) - y_j\right]}{\sigma_j^2} \tag{7}$$

mit der Abkürzung

$$\sigma_j^2 = \left(f_j^{x0}\right)^2 \sigma_j^{x2} + \sigma_j^{y2}\,. \tag{8}$$

Setzt man nun die Gleichungen (7), (8) in Glg. (4) ein, ergibt sich - nach einigen elementaren Umformungen - die "neue" Least-Squares-Bedingung

$$\chi^2 = \sum_{k=1}^{n}\left[\frac{y_k - f(x_k, \mathbf{a})}{\sigma_k}\right]^2 \quad \rightarrow \quad \text{Min}\,. \tag{9}$$