

Fit und Fehler

Anleitung zur Fehlerrechnung für das Kernphysikalische Praktikum
(18.Sep.2006, K. Huber)

1	Fehlerrechnung für Anfänger	2
1.1	Fehlerquellen	2
1.1.1	Statistische Fehler	2
1.1.2	Messfehler	2
1.1.3	Systematische Fehler	3
1.2	Die Fehlerfortpflanzungsformel	3
1.3	Beispiele	3
1.3.1	Beispiel 1: Addition von Messungen	3
1.3.2	Beispiel 2: Achtung Falle!	3
1.3.3	Beispiel 3: Fläche unter einer Geraden	4
1.3.4	Beispiel 4: Schnittpunkt einer Geraden mit der X-Achse	6
2	Fit und Fehlerrechnung für Experten	8
2.1	Linearer Fit nach der Methode der kleinsten Quadrate	8
2.2	Fehlerfortpflanzung	9
2.3	Beispiele	10
2.3.1	Beispiel 1: Fehlerfortpflanzung	10
2.3.2	Beispiel 2: Mittelwertbildung	11
2.3.3	Beispiel 3: Geraden-Fit zur Energieeichung beim Spektroskopieversuch	12

1 Fehlerrechnung für Anfänger

Dies hier ist der Versuch, möglichst kurz gefasst (weil es sonst ja doch keiner liest) das wichtigste Wissen über die Fehlerrechnung zu vermitteln.

In der Physik versteht man unter Fehlerangaben üblicherweise die Standardabweichung. Da man davon ausgeht, dass Messfehler normalverteilt (Gauss-Verteilung, Glockenkurve) sind (Fehlermodell von Laplace), gibt die Standardabweichung an, dass man bei einer Wiederholung der Messung mit 68,2% Wahrscheinlichkeit ein Ergebnis innerhalb der Fehlergrenzen erhält. Ängstliche Gemüter geben gelegentlich auch die doppelte oder dreifache Standardabweichung an und erhalten damit 95,4% bzw. 99,8% Wahrscheinlichkeit. Auf eine solche Vervielfachung der Standardabweichung sollte man aber auf jeden Fall ausdrücklich hinweisen, um Missverständnisse zu vermeiden!

Jede Fehlerrechnung hat folgende Ziele:

- a) Die ursprünglichen Fehler einer Messung so zu bestimmen, dass sie als Standardabweichung behandelt werden können.
- b) Durch Fehlerfortpflanzungsrechnung für das Endergebnis einer nachfolgenden Auswertung einen der Standardabweichung adäquaten Fehler zu ermitteln.

Wie wir sehen werden, können diese Ziele manchmal mit recht einfachen Mitteln erreicht werden, oft jedoch bereiten sie auch ziemliches Kopferbrechen und werden deshalb als lästiges Anhängsel ignoriert.

1.1 Fehlerquellen

1.1.1 Statistische Fehler

Für zufällige Ereignisse, wie z.B. dem radioaktiven Zerfall, lässt sich aus der Verteilungsfunktion (Poissonverteilung) die Standardabweichung berechnen als Wurzel aus dem Mittelwert der Verteilung. Leider kennt man den Mittelwert aber nicht und eine Messung ergibt nur eine Näherung für ihn. Für die Fehlerrechnung muss dies aber ausreichen und man nimmt deshalb die Wurzel aus dem Messwert M als Standardabweichung.

$$\sigma = \sqrt{M}$$

Warnung: dies gilt nur für Zählergebnisse und nicht für Zählraten, die das Ergebnis der Rechnung Zählergebnis/Messzeit sind und deshalb einer Fehlerfortpflanzungsrechnung bedürfen (s.u.)!

1.1.2 Messfehler

Haben Sie z.B. mit einer Stoppuhr eine Zeitmessung durchgeführt, so ist es sicherlich nicht ohne weiteres möglich einen mit der Standardabweichung vergleichbaren Fehler dazu anzugeben. Sie könnten aber eine ausreichende Anzahl von Wiederholungen solcher Messungen machen, daraus das Mittel bilden und aus der Streuung (Verteilung) der Messwerte sich die Standardabweichung berechnen lassen (z.B. mit Origin), bzw. selber berechnen. Dazu summieren Sie die quadratischen Abweichungen der Messwerte x_i vom Mittelwert M und dividieren durch die Anzahl der Freiheitsgrade, das ist hier die Anzahl der Messwerte n minus 1.

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - M)^2}$$

1.1.3 Systematische Fehler

Apparative oder systematische Fehler sind nur recht schwierig in den Griff zu bekommen. Bei Präzessionsmessgeräten kann man möglicherweise beim Hersteller die Toleranzen erfahren. Bei komplexen Apparaturen hilft oft nur ein vorsichtiges Abschätzen.

1.2 Die Fehlerfortpflanzungsformel

Die allgemeine Formel der Fehlerfortpflanzung für die Funktion

$$y = f(a, b, \dots)$$

lautet :

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial a} \Delta a\right)^2 + \left(\frac{\partial y}{\partial b} \Delta b\right)^2 + \dots}$$

und hat als Bedingung, dass die Parameter a, b, ... statistisch unabhängig sind, d.h. sie sind das Ergebnis unabhängiger Messungen. Ein Missachten dieser Bedingung kann zu völlig falschen Ergebnissen führen, wie in den folgenden Beispielen gezeigt wird.

Diese Fehlerfortpflanzungsformel stellt eine Reihenentwicklung bis zum linearen Glied dar, was für kleine Fehler ausreichend ist, für große Fehler allerdings zu unbrauchbaren Ergebnissen führen kann.

1.3 Beispiele

1.3.1 Beispiel 1: Addition von Messungen

Sie haben zwei Zählmessungen an einem radioaktiven Präparat mit

$$z_1 = 95 \pm \sqrt{95}$$

$$z_2 = 105 \pm \sqrt{105}$$

durchgeführt und addieren diese beiden Messergebnisse mit Fehlerrechnung

$$y = z_1 + z_2 = 95 + 105 = 200$$

$$\Delta y = \sqrt{(\Delta z_1)^2 + (\Delta z_2)^2} = \sqrt{200}$$

$$\Delta y / y = 1 / \sqrt{200}$$

Wir erhalten für den Fehler das gleiche Ergebnis wie wenn wir beide Messungen zusammen an einem Stück gemacht hätten, was ja durchaus einleuchtend ist. Der relative Fehler verbessert sich beim Summieren.

1.3.2 Beispiel 2: Achtung Falle!

Sie haben eine Zählmessung an einem radioaktiven Präparat mit

$$z = 100 \pm \sqrt{100}$$

durchgeführt und addieren dieses Messergebnis zu sich selber mit Fehlerrechnung:

$$y = z + z = 100 + 100 = 200$$

$$\Delta y = \sqrt{4 \cdot 100} = \sqrt{400}$$

$$y = 200 \pm \sqrt{400}$$

$$\Delta y / y = 1 / \sqrt{100}$$

Im Gegensatz zu Beispiel 1 verbessert sich durch das Vervielfältigen der Messung der relative Fehler nicht!

Leider sind solche Abhängigkeiten nicht immer offensichtlich wie die folgende Rechnung zeigt. Aus ihr kann man auch ersehen, dass das Ignorieren der Abhängigkeiten sowohl zu einem zu großen als auch zu einem zu kleinen Fehler führen kann.

$$z \pm \Delta z$$

$$a = f_a(z); \quad \Delta a = \sqrt{\left(\frac{\partial f_a}{\partial z} \Delta z\right)^2}$$

$$b = f_b(z); \quad \Delta b = \sqrt{\left(\frac{\partial f_b}{\partial z} \Delta z\right)^2}$$

$$y = a + b$$

falsche Rechnung:

$$\Delta y \neq \sqrt{\left(\frac{\partial y}{\partial a} \Delta a\right)^2 + \left(\frac{\partial y}{\partial b} \Delta b\right)^2} = \sqrt{\left(\left(\frac{\partial f_a}{\partial z}\right)^2 + \left(\frac{\partial f_b}{\partial z}\right)^2\right) (\Delta z)^2}$$

richtige Rechnung:

$$\Delta y = \sqrt{\left(\frac{\partial y}{\partial z} \Delta z\right)^2} = \sqrt{\left(\left(\frac{\partial f_a}{\partial z}\right) + \left(\frac{\partial f_b}{\partial z}\right)\right)^2 (\Delta z)^2}$$

1.3.3 Beispiel 3: Fläche unter einer Geraden

Für die beiden unabhängigen Variablen x_1 und x_2 haben wir zwei fehlerbehaftete Messwerte y_1 und y_2 , durch die wir eine Gerade legen, um die Fläche F darunter zu bestimmen:

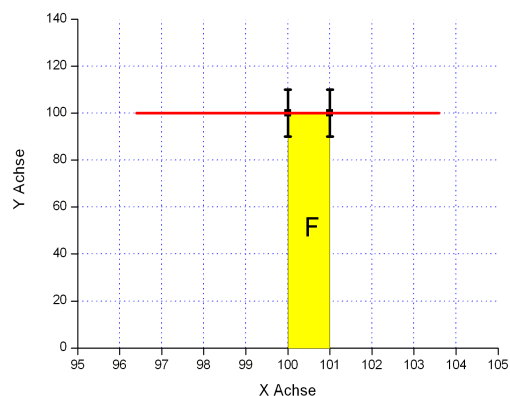
$$x_1 = 100; \quad y_1 = 100 \pm 10$$

$$x_2 = 101; \quad y_2 = 100 \pm 10$$

Zunächst der einfache Lösungsweg:

$$F = \frac{y_1 + y_2}{2} (x_2 - x_1) = 100$$

$$\Delta F = \sqrt{\frac{(\Delta y_1)^2 + (\Delta y_2)^2}{4}} (x_2 - x_1) = \sqrt{50}$$



Nun soll die Fläche aber durch Integration der Geraden errechnet werden. Zuerst bestimmen wir die Geradengleichung, die sich hier aus nur zwei Messpunkten errechnet, im Allgemeinen aber sich durch eine Ausgleichsrechnung (Fit) aus vielen Punkten ergibt. Dann wäre eine einfache Flächenberechnung wie zuvor nicht möglich:

$$y = a + bx$$

$$a = (y_1 x_2 - y_2 x_1) / (x_2 - x_1) = 100$$

$$b = (y_2 - y_1) / (x_2 - x_1) = 0$$

$$\Delta a = \sqrt{((x_2 \Delta y_1)^2 + (x_1 \Delta y_2)^2) / (x_2 - x_1)^2} = \sqrt{2020100}$$

$$\Delta b = \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2) / (x_2 - x_1)^2} = \sqrt{200}$$

Der Fehler des y-Achsabschnittes Δa gerät ziemlich groß, da wir weit von dem Nullpunkt der x-Achse entfernt sind.

Jetzt integrieren wir die Fläche unter der Geraden und bestimmen deren Fehler durch Fehlerfortpflanzung aus den Geradenparametern. Dies kann allerdings nicht gut gehen, da a und b statistisch nicht unabhängig sind.

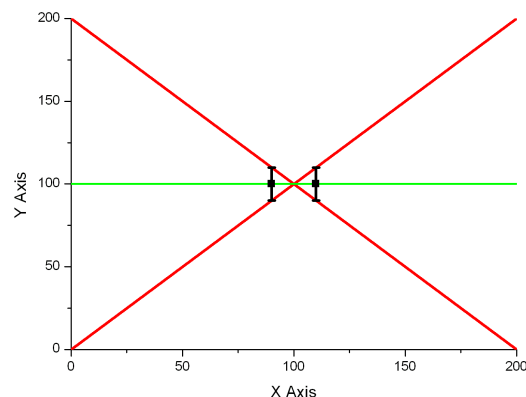
$$F = a(x_2 - x_1) + \frac{b}{2}(x_2^2 - x_1^2) = 100$$

$$\Delta F \neq \sqrt{\left(\frac{\partial F}{\partial a} \Delta a\right)^2 + \left(\frac{\partial F}{\partial b} \Delta b\right)^2} = \sqrt{((x_2 - x_1) \Delta a)^2 + \left(\frac{1}{2}(x_2^2 - x_1^2) \Delta b\right)^2} \approx 2010$$

Dies ist offensichtlich ein total unsinniges Resultat für den Fehler der Fläche. Ferner ist auffällig, dass das Ergebnis der Fehlerrechnung abhängig von der Position der Fläche auf der x-Achse ist, da nicht nur Differenzen von x_1 und x_2 auftreten. Diese Rechnung ist ohne Zweifel gründlich falsch!

Wie müssten wir nun vorgehen, um ein richtiges Ergebnis zu erhalten?

- Im vorliegenden Fall können wir die Fehlerrechnung auf die ursprünglichen Fehler Δy_1 und Δy_2 zurückführen statt mit Δa und Δb zu rechnen, um die statistische Abhängigkeit (Kovarianzen) von a und b zu umgehen. Dies ist jedoch so einfach nicht möglich, wenn wir die Gerade z.B. aus mehreren Messwerten durch Ausgleichsrechnung (Fit) bestimmt haben.
- Wir rechnen die Fehlerfortpflanzung mit voller Berücksichtigung der Kovarianzen, so wie dies im Kapitel 2 beschrieben ist.
- Es gibt aber auch noch einen interessanten Kniff um dieses Problem zu umgehen. Wie die folgende Abbildung zeigt, können bei einem großen Abstand der Messpunkte vom Nullpunkt der x-Achse bereits geringe Unsicherheiten in der Steigung der Geraden zu einem großen Fehler des y-Achsabschnitts führen. Bei einer professionellen Fehlerrechnung wird dieses Verhalten durch die Kovarianzen berücksichtigt. Tun wir hingegen so als wären diese Fehler unabhängig voneinander, dann addieren sie sich zu unsinnigen Werten.



Aus der obigen Abbildung kann man aber auch erkennen, dass wir dieses Problem durch eine geeignete Wahl des Nullpunktes der x-Achse in den Griff bekommen können: man legt den Nullpunkt genau zwischen die beiden Messwerte.

$$x_0 = (x_1 + x_2) / 2$$

$$y = a + b(x - x_0)$$

$$a = (y_1(x_2 - x_0) - y_2(x_1 - x_0)) / (x_2 - x_1) = (y_1 + y_2) / 2 = 100$$

$$b = (y_2 - y_1) / (x_2 - x_1) = 0$$

$$\Delta a = \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2) / 4} = \sqrt{50}$$

$$\Delta b = \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2) / (x_2 - x_1)^2} = \sqrt{200}$$

$$F = a(x_2 - x_1) = 100$$

$$\Delta F = \sqrt{((x_2 - x_1)\Delta a)^2} = \sqrt{50}$$

Das Problem hat sich hier schon auf einfache Weise dadurch erledigt, dass bei der Flächenberechnung der Parameter b nicht mehr auftritt. Es hätte jedoch in jedem Falle funktioniert, da durch die Nullpunktverschiebung die beiden Geradenparameter statistisch unabhängig werden.

Dieses Verfahren ist auch anwendbar, wenn die Gerade durch einen Fit an mehrere Messwerte entsteht. Dann wird man vor dem Fit eine Transformation des Nullpunktes der x-Achse in den Schwerpunkt S_x der Messdaten vornehmen. Für einen gewichteten Fit sind natürlich die einzelnen Gewichte g_i bei der Schwerpunktsberechnung zu berücksichtigen:

$$x_0 = S_x = \frac{\sum_i x_i g_i}{\sum_i g_i}$$

Wie in Kapitel 2 gezeigt wird, verschwinden durch eine solche Transformation des Nullpunktes der x-Achse die Kovarianzen zwischen a und b.

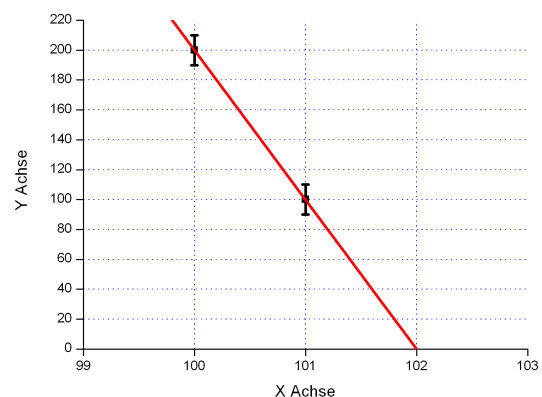
1.3.4 Beispiel 4: Schnittpunkt einer Geraden mit der X-Achse

Für die beiden unabhängigen Variablen x_1 und x_2 haben wir zwei fehlerbehaftete Messwerte y_1 und y_2 , durch die wir eine Gerade legen. Zu bestimmen ist der Schnittpunkt der Geraden mit der x-Achse sowie der Fehler des Schnittpunktes.

$$x_1 = 100; \quad y_1 = 200 \pm 10$$

$$x_2 = 101; \quad y_2 = 100 \pm 10$$

Als erstes transformieren wir den Nullpunkt der x-Achse in den Schwerpunkt der Messwerte und berechnen die Geradengleichung:



$$\begin{aligned}
x_0 &= (x_1 + x_2)/2 = 100.5 \\
y &= a + b(x - x_0) \\
a &= (y_1(x_2 - x_0) - y_2(x_1 - x_0))/(x_2 - x_1) = (y_1 + y_2)/2 = 150 \\
b &= (y_2 - y_1)/(x_2 - x_1) = -100 \\
\Delta a &= \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2)/4} = \sqrt{50} \\
\Delta b &= \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2)/(x_2 - x_1)^2} = \sqrt{200}
\end{aligned}$$

Nun die Berechnung des Schnittpunktes mit der x-Achse:

$$\begin{aligned}
y = 0 &\rightarrow x = -\frac{a}{b} + x_0 = 102 \\
\Delta x &= \sqrt{\left(-\frac{1}{b}\Delta a\right)^2 + \left(\frac{a}{b^2}\Delta b\right)^2} = \sqrt{0.045 + 0.005} \approx 0.224
\end{aligned}$$

Die gleiche Rechnung ohne Nullpunktverschiebung liefert wegen der vernachlässigten Kovarianzen einen viel zu großen Fehler:

$$\begin{aligned}
y &= a + bx \\
a &= (y_1x_2 - y_2x_1)/(x_2 - x_1) = 10200 \\
b &= (y_2 - y_1)/(x_2 - x_1) = -100 \\
\Delta a &= \sqrt{((x_2\Delta y_1)^2 + (x_1\Delta y_2)^2)/(x_2 - x_1)^2} = \sqrt{2020100} \\
\Delta b &= \sqrt{((\Delta y_1)^2 + (\Delta y_2)^2)/(x_2 - x_1)^2} = \sqrt{200} \\
y = 0 &\rightarrow x = -\frac{a}{b} = 102 \\
\Delta x &\neq \sqrt{\left(-\frac{1}{b}\Delta a\right)^2 + \left(\frac{a}{b^2}\Delta b\right)^2} = \sqrt{202.01 + 416.16} \approx 24.9
\end{aligned}$$

Dieses Beispiel behandelt einen Teil der Fehlerrechnung der Kurie-Plot-Auswertung von Versuch 4 des Kernphysik-Praktikums.

2 Fit und Fehlerrechnung für Experten

Hier wird nur das notwendige Handwerkszeug vorgestellt. Wer sich für Herleitungen und Hintergründe interessiert, wird in folgender Literatur fündig:

[1] *Sigmund Brandt: Statistische Methoden der Daten Analyse*

[2] *Philip R. Bevington: Data Reduction and Error Analysis for the Physical Sciences*

2.1 Linearer Fit nach der Methode der kleinsten Quadrate

Die folgende Betrachtung beschränkt sich auf lineare Funktionen. Beliebige Funktionen fittet man durch Taylorentwicklung bis zur ersten Ableitung und iterativem Anwenden des linearen Formalismus (siehe [2]).

Es seien y_i n unabhängige Messungen, die von r unbekannten a_j und deren Koeffizienten x_{ij} sowie von n Konstanten c_i abhängen:

$$y_1 = x_{11}a_1 + x_{12}a_2 + \dots + x_{1r}a_r + c_1$$

...

$$y_n = x_{n1}a_1 + x_{n2}a_2 + \dots + x_{nr}a_r + c_n$$

In Matrixschreibweise wird dies übersichtlicher:

$$\mathbf{Y} = \mathbf{XA} + \mathbf{C}$$

$$\begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & \dots & x_{1r} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nr} \end{pmatrix} \begin{pmatrix} a_1 \\ \dots \\ a_r \end{pmatrix} + \begin{pmatrix} c_1 \\ \dots \\ c_n \end{pmatrix}$$

Die Fehler der n Messungen seien σ_i . Da die Messungen statistisch unabhängig sein sollen, können wir die Varianzen σ_i^2 zu einer diagonalen Kovarianzmatrix \mathbf{C}_Y bzw. Gewichtsmatrix \mathbf{G}_Y zusammenfassen:

$$\mathbf{C}_Y = \begin{pmatrix} \sigma_1^2 & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \sigma_n^2 \end{pmatrix} \quad \mathbf{G}_Y = \mathbf{C}_Y^{-1} = \begin{pmatrix} 1/\sigma_1^2 & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & 1/\sigma_n^2 \end{pmatrix}$$

Ein Fit nach der Methode der kleinsten Fehlerquadrate führt dann zu folgendem Ergebnis für die unbekannten a_j :

$$\mathbf{A} = (\mathbf{X}^T \mathbf{G}_Y \mathbf{X})^{-1} \mathbf{X}^T \mathbf{G}_Y (\mathbf{Y} - \mathbf{C})$$

Von Interesse ist jetzt natürlich noch, wie sich die Fehler σ_i der Messwerte y_i auf die Parameter a_j fortpflanzen:

$$\mathbf{C}_A = (\mathbf{X}^T \mathbf{G}_Y \mathbf{X})^{-1}$$

Die Kovarianzmatrix \mathbf{C}_A enthält auf der Diagonalen die Varianzen der Parametern a_j aber im Allgemeinen auch noch von Null verschiedene Werte für die Kovarianzen, die bei einer weitergehenden Fehlerrechnung mit den Fehlern der a_j nicht vernachlässigt werden dürfen!

Mit den so gewonnenen a_j können wir nun die „verbesserten“ η_i berechnen sowie deren Fehler (Kovarianzmatrix \mathbf{C}_η), die natürlich kleiner sein sollten als die ursprünglichen Fehler der Messwerte y_i , weshalb dieses Verfahren auch den Namen „Ausgleichsrechnung“ erhalten hat:

$$\boldsymbol{\eta} = \mathbf{X}\mathbf{A} + \mathbf{C}$$

$$\mathbf{C}_\eta = \mathbf{X}\mathbf{C}_A\mathbf{X}^T$$

Was aber tun, wenn die Messfehler unbekannt sind?

Für den Fall, dass alle Messfehler gleich sind ($\sigma_i = \sigma$), kann man die gemeinsame Varianz aus der Gewichtsmatrix \mathbf{G}_Y herausziehen und es bleibt nur die Einheitsmatrix \mathbf{E} .

$$\mathbf{G}_Y = 1/\sigma^2 * \mathbf{E}$$

$$\mathbf{A} = (\mathbf{X}^T \mathbf{E} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{E} (\mathbf{Y} - \mathbf{C})$$

$$\mathbf{C}_A = \sigma^2 * (\mathbf{X}^T \mathbf{E} \mathbf{X})^{-1}$$

Die Formel für den Fit der \mathbf{A} bleibt dabei unverändert, da sich die gemeinsame Varianz rauskürzt. Für die Berechnung der Kovarianzmatrix \mathbf{C}_A benötigen wir sie aber. Sie lässt sich abschätzen aus der Streuung der Messwerte y_i um die Fit-Gerade:

$$\sigma^2 = \sum_{i=1}^n (y_i - \eta_i)^2 / F$$

$$F = n - r$$

Dies ist die Fehlerquadratsumme dividiert durch die Anzahl der Freiheitsgrade F (reduzierte Fehlerquadratsumme) mit der Anzahl n der Messwerte y_i und Anzahl r der Parameter a_j .

Obwohl dieses Verfahren nur begründet ist, wenn alle Varianzen gleich sind, funktioniert es ebenfalls erstaunlich gut in anderen Fällen. Es ist ferner ein guter Rat, beide Fehlerrechnungen durchzuführen und die Ergebnisse zu vergleichen. Sollte z.B. die Fehlerabschätzung der Messwerte missraten sein, so würde das dabei zu Tage treten. Im Zweifel nimmt man den größeren der beiden Fehler.

2.2 Fehlerfortpflanzung

Wir gehen wieder von n linearen Funktionen b_i von r Variablen a_j mit den Koeffizienten t_{ij} aus, die die Transformation der Parameter \mathbf{A} nach den Parametern \mathbf{B} darstellen:

$$b_1 = t_{11}a_1 + t_{12}a_2 + \dots + t_{1r}a_r + c_1$$

...

$$b_n = t_{n1}a_1 + t_{n2}a_2 + \dots + t_{nr}a_r + c_n$$

In Matrixschreibweise:

$$\mathbf{B} = \mathbf{T}\mathbf{A} + \mathbf{C}$$

$$\begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix} = \begin{pmatrix} t_{11} & \dots & t_{1r} \\ \dots & \dots & \dots \\ t_{n1} & \dots & t_{nr} \end{pmatrix} \begin{pmatrix} a_1 \\ \dots \\ a_r \end{pmatrix} + \begin{pmatrix} c_1 \\ \dots \\ c_n \end{pmatrix}$$

Die Fehlerfortpflanzung, d.h. die Transformation der Kovarianzmatrix \mathbf{C}_A der Parameter a_j in die Kovarianzmatrix \mathbf{C}_B der Parameter b_i sieht dann folgendermaßen aus:

$$\mathbf{C}_B = \mathbf{T} \mathbf{C}_A \mathbf{T}^T$$

Dieser Zusammenhang gilt für lineare Funktionen. Er lässt sich auf beliebige Funktionen ausdehnen, wenn man von verhältnismäßig kleinen Fehlern ausgeht, eine Taylorentwicklung der Funktionen durchführt und nur die erste partielle Ableitung verwendet:

$$\mathbf{T} = \begin{pmatrix} \partial b_1 / \partial a_1 & \dots & \partial b_1 / \partial a_r \\ \dots & \dots & \dots \\ \partial b_n / \partial a_1 & \dots & \partial b_n / \partial a_r \end{pmatrix}$$

2.3 Beispiele

Die Beispiele wurden mit den Matrixfunktionen von Excel gerechnet. Es empfiehlt sich Excel durch das freie Excel-Matrix-Add-In von **digilander.libero.it** zu ergänzen. Dies erlaubt z.B. aus einem Vektor eine Diagonalmatrix zu erzeugen oder mit höherer Genauigkeit zu rechnen, was beim Invertieren einer größeren Matrix schnell notwendig wird.

2.3.1 Beispiel 1: Fehlerfortpflanzung

Es seien z_1, \dots, z_n statistisch unabhängige Messungen mit den Fehlern $\sigma_1, \dots, \sigma_n$ und y eine Funktion dieser n Messungen mit dem Fehler σ_y . Da die Messungen unabhängig sein sollen, ist die zugehörige Kovarianzmatrix \mathbf{C}_z eine Diagonalmatrix mit den Varianzen auf der Diagonalen. Die Transformationsmatrix enthält die partiellen Ableitungen.

$$y = f(z_1, \dots, z_n)$$

$$\mathbf{C}_Y = \mathbf{T} \mathbf{C}_Z \mathbf{T}^T$$

$$\mathbf{C}_Y = \begin{pmatrix} \partial y / \partial z_1 & \dots & \partial y / \partial z_n \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \sigma_n^2 \end{pmatrix} \begin{pmatrix} \partial y / \partial z_1 \\ \dots \\ \partial y / \partial z_n \end{pmatrix}$$

$$\mathbf{C}_Y = \sigma_y^2 = (\partial y / \partial z_1)^2 \cdot \sigma_1^2 + \dots + (\partial y / \partial z_n)^2 \cdot \sigma_n^2$$

Als Ergebnis erhalten wir die gewohnte Fehlerfortpflanzungsformel für statistisch unabhängige Eingangsfehler.

Wir betrachten jetzt die Summen y_1 und y_2 , in denen beiden z_3 enthalten ist. Wir erwarten deshalb Kovarianzen zwischen beiden Summen. Auf der Hauptdiagonalen finden wir die Varianzen für y_1 und y_2 , die wir auch mit der üblichen Fehlerrechnung erhalten hätten und abseits der Diagonalen die Kovarianzen. Kovarianzmatrizen sind immer symmetrisch bezüglich der Hauptdiagonalen.

$$y_1 = z_1 + z_3$$

$$y_2 = z_2 + z_3$$

$$\mathbf{C}_Y = \begin{pmatrix} \partial y_1 / \partial z_1 & \partial y_1 / \partial z_2 & \partial y_1 / \partial z_3 \\ \partial y_2 / \partial z_1 & \partial y_2 / \partial z_2 & \partial y_2 / \partial z_3 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \dots & \sigma_2^2 & \dots \\ 0 & \dots & \sigma_3^2 \end{pmatrix} \begin{pmatrix} \partial y_1 / \partial z_1 & \partial y_2 / \partial z_1 \\ \partial y_1 / \partial z_2 & \partial y_2 / \partial z_2 \\ \partial y_1 / \partial z_3 & \partial y_2 / \partial z_3 \end{pmatrix}$$

$$\mathbf{C}_Y = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \dots & 0 \\ \dots & \sigma_2^2 & \dots \\ 0 & \dots & \sigma_3^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\mathbf{C}_Y = \begin{pmatrix} \sigma_1^2 + \sigma_3^2 & \sigma_3^2 \\ \sigma_3^2 & \sigma_2^2 + \sigma_3^2 \end{pmatrix}$$

Wenn wir jetzt y_1 und y_2 addieren wollen, so müssen wir bei der Fehlerrechnung die Kovarianzen zwischen beiden berücksichtigen:

$$u = y_1 + y_2$$

$$\mathbf{C}_U = \sigma_u^2 = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 + \sigma_3^2 & \sigma_3^2 \\ \sigma_3^2 & \sigma_2^2 + \sigma_3^2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \sigma_1^2 + \sigma_2^2 + 4\sigma_3^2$$

Das gleiche Ergebnis hätten wir natürlich auch mit der üblichen Fehlerfortpflanzung für $u = z_1 + z_2 + 2z_3$ erhalten, für $u = y_1 + y_2$ jedoch nur $\sigma_1^2 + \sigma_2^2 + 2\sigma_3^2$.

2.3.2 Beispiel 2: Mittelwertbildung

Wiederum seien z_1, \dots, z_n statistisch unabhängige Messungen mit den Fehlern $\sigma_1, \dots, \sigma_n$ und M der Mittelwert für diese n Messungen mit dem Fehler σ_M . Zur Mittelwertrechnung benötigen wir die Gewichtsmatrix \mathbf{G}_z , die einfach die invertierte Kovarianzmatrix \mathbf{C}_z ist und sich im Falle einer Diagonalmatrix besonders einfach rechnet:

$$\mathbf{C}_M = (\mathbf{X}^T \mathbf{G}_z \mathbf{X})^{-1}$$

$$\mathbf{C}_M = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \begin{pmatrix} g_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & g_n \end{pmatrix} \begin{bmatrix} 1 \\ \dots \\ 1 \end{bmatrix}^{-1}$$

$$\mathbf{C}_M = \sigma_M^2 = 1 / (g_1 + \dots + g_n)$$

$$\mathbf{M} = \mathbf{C}_M \mathbf{X}^T \mathbf{G}_Z \mathbf{Z}$$

$$\mathbf{M} = \mathbf{C}_M \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} g_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & g_n \end{pmatrix} \begin{pmatrix} z_1 \\ \dots \\ z_n \end{pmatrix}$$

$$\mathbf{M} = (z_1 g_1 + \dots + z_n g_n) / (g_1 + \dots + g_n)$$

Wir haben das gewichtete Mittel \mathbf{M} erhalten.

!!Anmerkung zur Wichtung bei der Mittelwertrechnung !!

Für zufällige Ereignisse, wie z.B. dem radioaktiven Zerfall, lässt sich aus der Verteilungsfunktion (Poissonverteilung) die Standardabweichung berechnen als Wurzel aus dem Mittelwert der Verteilung. Leider kennt man den Mittelwert aber nicht und eine Messung ergibt nur eine Näherung für ihn. Trotzdem nimmt man als Fehler die Wurzel aus dem Messwert und verwendet ihn als Wichtung. Macht man nun am selben Präparat 10 gleiche Messungen, um anschließend den Mittelwert zu bilden, so gehen diese Messungen alle mit unterschiedlicher Wichtung in die Rechnung ein, obwohl sie alle aus der gleichen statistischen Gesamtheit abstammen. Dies wird im Allgemeinen sogar den Mittelwert verfälschen wie folgendes Beispiel zeigt:

Der 'echte' Mittelwert sei $z_0 = 100$. Wir machen zwei Messungen mit $z_1 = 90$ und $z_2 = 110$ und bilden daraus das gewichtete Mittel mit $g_i = 1/\sigma_i^2 = 1/z_i$

$$\mathbf{M} = (90/90 + 110/110) / (1/90 + 1/110) = 99$$

$$\sigma_M = \sqrt{1/(g_1 + g_2)} = \sqrt{99/2}$$

während sich für gleiche Wichtung die Mittelung

$$\mathbf{M} = (90 + 110) / 2 = 100$$

$$\sigma_M = \sqrt{1/(2g_0)} = \sqrt{100/2}$$

mit dem aus dem Mittelwert errechneten $g_0 = 1/\sigma_0^2 = 1/M$ ergibt, wobei M natürlich auch wieder nur eine, wenn auch verbesserte Näherung für z_0 ist.

Wenn also die Messwerte ersichtlich aus der gleichen statistischen Gesamtheit abstammen, macht es keinen Sinn mit unterschiedlicher Wichtung zu mitteln!

2.3.3 Beispiel 3: Geraden-Fit zur Energieeichung beim Spektroskopieversuch

Wir gehen davon aus, dass die Literaturwerte E_i für die Energien so genau sind, dass wir sie als fehlerfrei ansehen können, während die ermittelten Kanalwerte K_i fehlerbehaftet sind.

Deshalb schreiben wir die Eichgerade in der folgenden Form, um den obigen Formalismus anwenden zu können:

$$K_i = a_1 * E_i + a_2$$

Wir haben $n = 6$ Messwertepaare und wollen damit $r = 2$ Parameter bestimmen. Die Koeffizienten der a_1 sind die E_i , die der a_2 sind 1 (Koeffizientenmatrix \mathbf{E}_M) und die c_i sind 0. Die 6 Messungen sind statistisch unabhängig, so dass die Fehler ΔK_i der Kanäle nur zur Diagonalen der Gewichtsmatrix \mathbf{G}_K beitragen.

Kanalnummer		Koeffizienten Matrix E_M			Gewicht zu K	
Kanal K_i	Fehler ΔK_i	Energie E_i			$1 / (\Delta K_i)^2$	
512,81	0,50	624,16	1		4,0000E+00	
538,98	0,50	655,61	1		4,0000E+00	
396,96	0,50	481,60	1		4,0000E+00	
457,65	0,50	553,74	1		4,0000E+00	
793,41	0,50	975,60	1		4,0000E+00	
852,34	0,50	1047,74	1		4,0000E+00	
Gewichtsmatrix G_K der Kanalnummern						
$1 / \Delta K_i^2$ auf der Diagonalen						
4,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	
0,0000E+00	4,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	
0,0000E+00	0,0000E+00	4,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	
0,0000E+00	0,0000E+00	0,0000E+00	4,0000E+00	0,0000E+00	0,0000E+00	
0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	4,0000E+00	0,0000E+00	
0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	0,0000E+00	4,0000E+00	

Die Kovarianzmatrix C_A der Parameter A errechnet sich dann zu

$$C_A = (E_M^T G_K E_M)^{-1}$$

und die Parameter A fitten zu

$$A = C_A E_M^T G_K K$$

Nun ist es aber so, dass uns eigentlich die inverse Beziehung

$$E_i = b_1 * K_i + b_2 = K_i / a_1 - a_2 / a_1$$

interessiert, um gemessene Kanäle in Energien umrechnen zu können. Deshalb wurden auch b_1 und b_2 berechnet und mit Hilfe der Fehlertransformationsmatrix T_B die zugehörige Kovarianzmatrix C_B . Lediglich zur Demonstration ist C_B wieder auf C_A zurückgerechnet. Hätten wir dabei nur die Varianzen berücksichtigt, so wären wir zu einem völlig anderen Ergebnis gekommen.

Kovarianzmatrix zu A		Geraden-Fit zu $K_i = a_1 * E_i + a_2$			
$C_A = (E_M^T G_K E_M)^{-1}$		$A = C_A E_M^T G_K K$			
9,2423E-07	-6,6829E-04	$a_1 =$	8,0091E-01		
-6,6829E-04	5,2489E-01	$a_2 =$	1,2907E+01		
				Fehlertransf. T_B (A->B)	
Kovarianzmatrix zu B		inverse Gerade		$-1/a_1^2$	0
$C_B = (T_B C_A T_B^T)$		$E_i = b_1 * K_i + b_2 = K_i / a_1 - a_2 / a_1$		a_2 / a_1^2	$-1/a_1$
2,2462E-06	-1,3298E-03	$b_1 =$	1,2486E+00	-1,5590E+00	0
-1,3298E-03	8,5223E-01	$b_2 =$	-1,6115E+01	2,0121E+01	-1,2486E+00
Rücktransformation als Test				Fehlertransf. T_A (B->A)	
Kovarianzmatrix zu A				$-1/b_1^2$	0
$C_A = (T_A C_B T_A^T)$				b_2 / b_1^2	$-1/b_1$
9,2423E-07	-6,6829E-04			-6,4145E-01	0
-6,6829E-04	5,2489E-01			-1,0337E+01	-8,0091E-01

Jetzt können wir die „verbesserten“ Kanalwerte Kc_i berechnen und deren Fehler ΔKc_i mit

$$Kc_i = E_M A$$

$$\Delta Kc_i = E_M C_A E_M^T$$

Die Berechnung der ΔKc_i^* , bei der nur die Varianzen berücksichtigt werden, zeigt, dass dies zu völlig anderen Fehlern führt.

Ferner wird die alternative Fehlerrechnung durchgeführt, bei der die Wichtung eins gesetzt ist und das σ^2 aus der Streuung abgeschätzt wird. Wie man sieht, ist die Fehlerabschätzung für die Messwerte (Kanäle) offensichtlich zu optimistisch. Die tatsächliche Streuung entspricht etwa einem ΔK_i von einem Kanal (hier nicht gezeigt). Natürlich kann dies aber auch ein Zufall sein. Wie wahrscheinlich ein solcher Zufall ist, darüber gibt der χ^2 -Test (Chi²-Test) Auskunft: < 0,001, d.h. die Fehler unserer Messwerte sind wohl doch eher missraten. (dazu demnächst mehr)

origin. Kanalnummer		korrig. Kanalnummer		vernachlässigte Kovarianzen!!!		
Kanal K_i	Fehler ΔK_i	$Kc_i = E_M^* \cdot A$	ΔKc_i		Δkc_i^*	
512,81	0,50	512,80	0,23		0,94	
538,98	0,50	537,99	0,21		0,96	
396,96	0,50	398,62	0,31		0,86	
457,65	0,50	456,40	0,26		0,90	
793,41	0,50	794,27	0,32		1,19	
852,34	0,50	852,05	0,37		1,24	
alternative Fehlerrechnung mit Wichtung 1						
σ^2			ΔKc_i			
1,5411E+00			0,56			
			0,53			
Kovarianzmatrix zu A			0,77			
$C_A = \sigma^2 \cdot (E_M^T \cdot E \cdot E_M)^{-1}$			0,65			
5,6973E-06	-4,1196E-03		0,79			
-4,1196E-03	3,2356E+00		0,93			

Wie mehrfach gezeigt wurde führt das Vernachlässigen der Kovarianzen zu falschen Ergebnissen. Man muss also stets die ganze Matrixrechnung durchführen. Für das vorliegende Beispiel ist es jedoch auch möglich, durch eine einfache Transformation die Kovarianzen verschwinden zu lassen. Die Gewichtsmatrix der Parameter **A** errechnet sich folgendermaßen:

$$G_A = (E_M^T G_K E_M)$$

$$\begin{pmatrix} \sum g_i E_i^2 & \sum g_i E_i \\ \sum g_i E_i & \sum g_i \end{pmatrix} = \begin{pmatrix} E_1 & \dots & E_n \\ 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} g_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & g_n \end{pmatrix} \begin{pmatrix} E_1 & 1 \\ \dots & \dots \\ E_n & 1 \end{pmatrix}$$

Wenn wir die Gewichtsmatrix **G_A** mit einer geeigneten Transformation zu einer Diagonalmatrix machen können, so wird auch die Kovarianzmatrix **C_A** eine Diagonalmatrix, d.h. die Kovarianzen verschwinden. Versuchen wir's mit einer Nullpunktverschiebung E_0 der Energie:

$$\sum g_i (E_i - E_0) = 0$$

$$E_0 = \sum g_i E_i / \sum g_i$$

Dies stellt gerade eine Verlegung des Energie-Nullpunktes in den Schwerpunkt bezüglich der Gewichte g_i dar.

Als nächstes soll aus einem Kanalwert **K** mit Hilfe der Eichungen **A** bzw. **B** ein Energiewert **E** mit Fehler ΔE berechnet werden unter Berücksichtigung der Kovarianzmatrizen **C_A** bzw.

C_B der Eichung und dem Fehler des Kanalwertes ΔK . Wenn wir alles richtig machen, sollte es gleichgültig sein, ob wir den Parametersatz **A** oder **B** verwenden. Durch die entsprechenden partiellen Ableitungen erhalten wir die Fehlertransformationen T_A bzw. T_B :

$$\begin{aligned} E &= (K - a_2) / a_1 & E &= b_1 * K + b_2 \\ \partial E / \partial a_1 &= -K / a_1^2 & \partial E / \partial b_1 &= K \\ \partial E / \partial a_2 &= -1 / a_1 & \partial E / \partial b_2 &= 1 \\ \partial E / \partial K &= 1 / a_1 & \partial E / \partial K &= b_1 \\ T_A &= \begin{pmatrix} -K / a_1^2 & -1 / a_1 & 1 / a_1 \end{pmatrix} & T_B &= \begin{pmatrix} K & 1 & b_1 \end{pmatrix} \end{aligned}$$

Jetzt fehlen uns noch die zugehörigen Kovarianzmatrizen C_A bzw. C_B :

$$C_A = \begin{pmatrix} \sigma_{A11}^2 & \sigma_{A12}^2 & 0 \\ \sigma_{A21}^2 & \sigma_{A22}^2 & 0 \\ 0 & 0 & \sigma_K^2 \end{pmatrix} \quad C_B = \begin{pmatrix} \sigma_{B11}^2 & \sigma_{B12}^2 & 0 \\ \sigma_{B21}^2 & \sigma_{B22}^2 & 0 \\ 0 & 0 & \sigma_K^2 \end{pmatrix}$$

Die σ_{Aij} bzw. σ_{Bij} sind die Elemente der Kovarianzmatrizen der Parameter **A** bzw. **B** und σ_K ist der Messfehler des Kanals K. Da die Messung des Kanals unabhängig ist von der Eichung, sind die entsprechenden Kovarianzen null.

Mit Ausführen der Fehlerfortpflanzung erhält man:

$$\begin{aligned} C_E &= T C T^T \\ \sigma_E^2 &= \sigma_{A11}^2 * K^2 / a_1^4 + \sigma_{A12}^2 * K / a_1^3 + \sigma_{A21}^2 * K / a_1^3 + \sigma_{A22}^2 / a_1^2 + \sigma_K^2 / a_1^2 \\ \sigma_E^2 &= \sigma_{B11}^2 * K^2 + \sigma_{B12}^2 * K + \sigma_{B21}^2 * K + \sigma_{B22}^2 + \sigma_K^2 * b_1^2 \end{aligned}$$

Bei verschwindenden Kovarianzen der Eichung führt die Rechnung zur „Standard“-Fehlerfortpflanzung: das quadratische Addieren der Fehler. Ferner erkennt man, dass der Fehler des Kanals nicht mit Kovarianzen verknüpft ist. Wir hätten also auch den Fehler der Eichung getrennt berechnen können, um anschließend den Kanalfehler quadratisch zu addieren.

Kanalnummer						
Kanal K	Fehler ΔK					
800,00	0,50					
Fehlertransformation A				Fehlertransformation B		
-K/a ₁ ²	-1/a ₁	1/a ₁		K	1	b ₁
-1,2472E+03	-1,2486E+00	1,2486E+00		8,0000E+02	1,0000E+00	1,2486E+00
Kovarinzmatrix A				Kovarinzmatrix B		
9,2423E-07	-6,6829E-04	0		2,2462E-06	-1,3298E-03	0
-6,6829E-04	5,2489E-01	0		-1,3298E-03	8,5223E-01	0
0	0	0,25		0	0	0,25
E	ΔE			E	ΔE	
982,750	0,564			982,750	0,552	
Rechnung ohne Berücksichtigung der Kovarianzen						
	2,646				2,680	

Auch hier zeigt sich, dass das Ignorieren der Kovarianzen zu unsinnigen Ergebnissen führen kann. Der Unterschied in den beiden korrekt gerechneten ΔE erklärt sich durch die

Beschränkung auf die erste partielle Ableitung in der Taylorentwicklung bei den Fehlertransformationen.