Ausgleichsrechnung

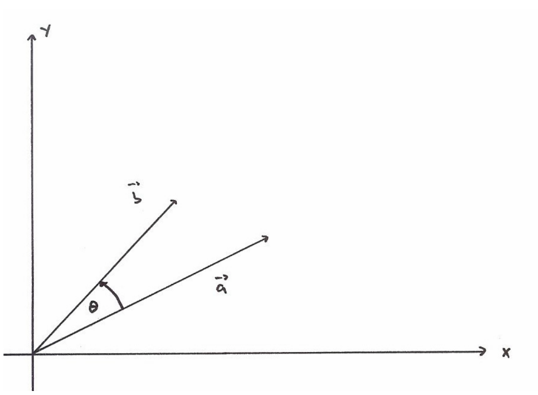
Roman Engeli, Florian Furrer, Yannik Drapela, Fabian Geiger

Einleitung

Die Ausgleichungsrechnung ist eine mathematische Optimierungsmethode, um für eine Reihe von Messdaten die unbekannten Parameter ihres geometrisch-physikalischen Modells oder die Parameter einer vorgegebenen Funktion zu bestimmen („zu schätzen"). Ziel der Ausgleichung ist, dass sich das endgültige Modell bzw. die Funktion den Daten und ihren unvermeidlichen kleinen Widersprüchen „bestmöglich" anpasst.

Die **Methode der kleinsten Quadrate** von Carl Friedrich Gauss (1777-1855) ist die am meisten verwendete Methode**.** Es wird versucht, eine Modellkurve durch eine Menge von Datenpunkten zu legen, so dass die Summe der Abstände der Punkte zur Kurve möglichst klein ist.

Skalarprodukt, Projektion



Gegeben sind zwei Vektoren a und b

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Skalarprodukt

Das Skalarprodukt ist definiert als

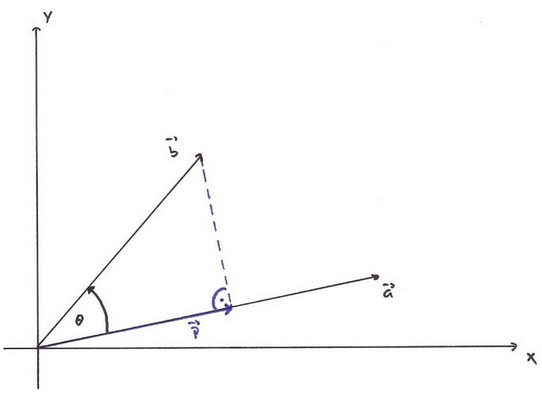
$$a \cdot b = a_1 b_1 + a_2 b_2 + a_3 b_3 = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a^T b \,.$$

Im Folgenden wird nur noch die Bezeichnung $a^T b$ verwendet.

Der Zusammenhang zwischen dem Zwischenwinkel und dem Skalarprodukt ist

$$\cos(\theta) = \frac{a^T b}{\|a\| \cdot \|b\|} \,.$$

Projektion



Die Projektion p von b auf a ist

$$p = \frac{a^T b}{a^T a} \cdot a \,.$$

Der Abstand der Endpunkte von b und p ist definiert als

$$d^2 = \frac{(a^T b)(a^T a) - (a^T b)^2}{(a^T a)} \,.$$

Die **Schwarz-Ungleichung** folgt aus der Tatsache, dass der Abstand zwischen zwei Punkten immer positiv ist.

$$\left| a^T b \right| \leq \|a\| \cdot \|b\|$$

Der Vektor a stellt einen Unterraum dar. Der Vektor b kann auch auf einen höher dimensionalen Unterraum, z.B. eine Ebene abgebildet werden. Der Vektor a wird dann zu der Matrix A mit den Spalteneinträgen der beiden Vektoren, die die Ebene aufspannen.

Methode der kleinsten Quadrate

Der Ausgleichsrechnung liegt immer dasselbe Problem zu Grunde. Man hat ein Gleichungssystem der Form $Ax = b$. Damit dieses System lösbar wird, muss b im Raum V liegen, der von A aufgespannt wird. Spannt A z.B. eine Ebene auf, muss b auf dieser Ebene liegen.
Stammt der Vektor b aus einer Messreihe, kann es sein, dass die Werte Messfehler beinhalten und b somit nicht genau im von A aufgespannten Raum liegt. Die Idee ist, dass man den Vektor b auf V projiziert um eine Lösung für das System zu erhalten. Dabei wird darauf geachtet, dass der Abstand zwischen v und b möglichst klein gehalten wird.

Normalengleichung

Die Normalengleichung beinhaltet die Lösung des inkonsistenten Gleichungssystems $Ax = b$.

$$A^T A\bar{x} = A^T b$$

Daraus folgt: $\bar{x} = \left(A^T A\right)^{-1} A^T b$.

Anwendung der Normalgleichung

Anhand des folgenden Beispieles soll eine Anwendung der Normalgleichung gezeigt werden. Weiter wird der Vergleich mit einer anderen Methode zur Fehlerminimalisierung gemacht.

Problem: Gegeben sei ein Dreieck mit den Eckpunkten A, B und C (Abb. 1). Die Aufgabe ist es, den Winkel x_1 bei A und x_2 bei B möglichst genau zu bestimmen.

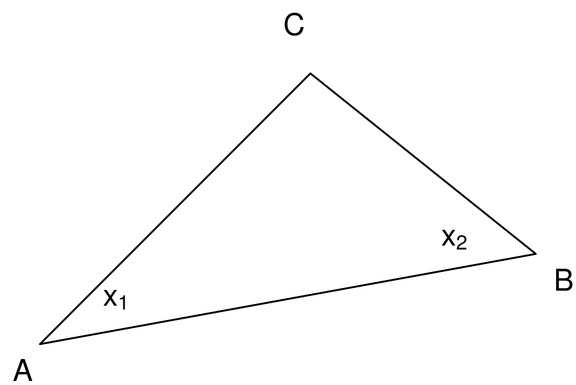


Abb. 1

Um die Genauigkeit der Messung zu erhöhen, wird auch der Winkel bei C gemessen.

Die Messresultate sind: $x_1 = 31°$

$x_2 = 62°$

Winkel bei C = 90° → $180 - (x_1 + x_2) = 90°$

Diese Messungen ergeben folgendes Gleichungssystem:

$$\begin{aligned} x_1 \quad\quad &= 31 \\ x_2 &= 62 \\ x_1 + x_2 &= 90 \end{aligned}$$

Es ist sofort ersichtlich, dass dieses Gleichungssystem widersprüchlich ist und keine Lösung besitzt. Deshalb werden nun die Messfehler mitberücksichtigt:

$$\begin{aligned} x_1 \quad\quad\; & - 31 = r_1 \\ x_2 & - 62 = r_2 \\ x_1 + x_2 & - 90 = r_3 \end{aligned}$$

Anders geschrieben wird daraus:

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \cdot x_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot x_2 - \begin{pmatrix} 31 \\ 62 \\ 90 \end{pmatrix} = r$$

Das Ziel ist es nun, x_1 und x_2 so zu bestimmen, dass der Vektor r minimal wird. Dies wird nun auf zwei verschiedene Methoden durchgeführt.

Methode 1: Minimierung der Quadrate der Komponenten von r:

$$E = (x_1 - 31)^2 + (x_2 - 62)^2 + (x_1 + x_2 - 90)^2$$

Die Funktion E ist minimal wenn gilt: $\frac{dE}{dx_1} = \frac{dE}{dx_2} = 0$

Dies führt zu folgendem Gleichungssystem:

$$\begin{aligned} 4x_1 + 2x_2 &= 242 \\ 2x_1 + 4x_2 &= 304 \end{aligned}$$

Dies führt zu den Lösungen: $x_1 = 30$

$x_2 = 61$

Methode 2: Normalgleichung

Die Normalgleichung lautet: $A^T Ax = A^T b$

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \cdot x_1 + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \cdot x_2 - \begin{pmatrix} 31 \\ 62 \\ 90 \end{pmatrix} = r$$

Lässt sich schreiben als:

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}}_{\text{A}} x - \underbrace{\begin{pmatrix} 31 \\ 62 \\ 90 \end{pmatrix}}_{\text{b}} = r$$

$$A^T A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad A^T b = \begin{pmatrix} 121 \\ 152 \end{pmatrix}$$

Dies ergibt das Gleichungssystem:

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 121 \\ 152 \end{pmatrix}$$

Aufgelöst ergibt dies die Lösungen: $x_1 = 30$
$x_2 = 61$

Somit ist ersichtlich, dass die Normalgleichung und die Minimierung der Quadrate dieselben Lösungen ergeben.

QR-Zerlegung

Als QR-Zerlegung bezeichnet man die Zerlegung einer Matrix A in das Produkt

$$A = Q \cdot R$$

wobei Q eine orthogonale ($QQ^T = Q^TQ = QQ^{-1} = I_n$) und R eine obere Dreiecksmatrix ist.

Eine solche Zerlegung existiert immer und kann mit verschiedenen numerischen Methoden berechnet werden. Die bekanntesten davon sind das Gram-Schmidtsche Orthogonalisierungsverfahren, die Householdertransformation und die Givens-Rotation.

Gram-Schmidtsches Orthogonalisierungsverfahren

Das Gram-Schidtsche Orthogonalisierungsverfahren berechnet aus einer Menge von linear unabhängigen Vektoren $a_1, a_2, ..., a_n$ eine Menge von paarweis orthogonalen, resp. orthonormalen Vektoren $v_1, v_2, ..., v_n$, die denselben Untervektorraum aufspannen.

Die einzelnen Vektoren $v_1, v_2, ..., v_n$ des Orthogonalsystems berechnen sich wie folgt:

$$v_1 = a_1$$

$$v_2 = a_2 - \frac{v_1{}^T a_2}{v_1{}^T v_1} \cdot v_1$$

$$v_3 = a_3 - \frac{v_1{}^T a_3}{v_1{}^T v_1} \cdot v_1 - \frac{v_2{}^T a_3}{v_2{}^T v_2} \cdot v_2$$

$$\vdots$$

$$v_n = a_n - \frac{v_1{}^T a_n}{v_1{}^T v_1} \cdot v_1 - \frac{v_2{}^T a_n}{v_2{}^T v_2} \cdot v_2 - ...... - \frac{v_{n-1}{}^T a_n}{v_{n-1}{}^T v_{n-1}} \cdot v_{n-1} = a_n - \sum_{i=1}^{n-1} \frac{v_i{}^T a_n}{v_i{}^T v_i} \cdot v_i$$

Um aus den orthogonalen Vektoren $v_1, v_2, ..., v_n$ ein Orthonormalsystem zu erhalten, müssen sie anschliessend noch durch ihre Länge dividiert werden.

$$q_n = \frac{v_n}{\|v_n\|}$$

Um aus diesen orthonormalen Vektoren die Matrizen Q und R zu erhalten, gibt man die Vektoren a_i in abhängigkeit von v und q an:

$$a_1 = |v_1| \cdot q_1$$

$$a_2 = \frac{v_1{}^T a_2}{v_1{}^T v_1} \cdot |v_1| \cdot q_1 + |v_2| \cdot q_2$$

$$\vdots$$

$$a_n = \frac{v_1{}^T a_n}{v_1{}^T v_1} \cdot |v_1| \cdot q_1 + ... + \frac{v_{n-1}{}^T a_n}{v_{n-1}{}^T v_{n-1}} \cdot |v_{n-1}| \cdot q_{n-1} + |v_n| \cdot q_n = \sum_{i=1}^{n-1} \frac{v_i{}^T a_n}{v_i{}^T v_i} \cdot |v_i| \cdot q_i + |v_n| \cdot q_n$$

Dieses Gleichungsystem kann nun folgendermassen in Matrizenform geschrieben werden:

$$\begin{bmatrix} | & & | \\ a_1 & ... & a_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ q_1 & ... & q_n \\ | & & | \end{bmatrix} \cdot \begin{bmatrix} |v_1| & \frac{v_1{}^T a_2}{v_1{}^T v_1} \cdot |v_1| & ... & \frac{v_1{}^T a_n}{v_1{}^T v_1} \cdot |v_1| \\ 0 & |v_2| & ... & \frac{v_2{}^T a_n}{v_2{}^T v_2} \cdot |v_2| \\ ... & ... & ... & ... \\ 0 & ... & 0 & |v_n| \end{bmatrix}$$

$$\mathbf{A} \quad = \quad \mathbf{Q} \quad \cdot \quad \mathbf{R}$$

Dabei wird ersichtlich, dass die orthonormalen Vektoren q die Spalten der orthonormalen Matrix Q bilden und dass die Matrix R eine obere Dreiecksmatrix bildet.

Bedeutung für das Ausgleichsproblem

Wendet man die QR-Zerlegung auf unser Augleichsproblem

$$\overline{x} = (A^T A)^{-1} A^T b$$

an, so vereinfacht sich die Lösung wie folgt:

$$\overline{x} = (R^T Q^T QR)^{-1} R^T Q^T b = (R^T R)^{-1} R^T Q^T b$$

Daraus folgt:

$\overline{x} = R^{-1} Q^T b$ Obwohl sich das Ausgleichsproblem vereinfacht, ist der Aufwand aufgrund der zuvor nötigen QR-Zerlegung grösser als wenn man die Normalengleichung direkt löst. Der Vorteil des Lösungsweges mit der QR-Zerlegung ist, dass das Verfahren im Gegensatz zum Lösen der Normalengleichung numerisch stabil ist. Das heisst, dass Rechnungsfehler aufgrund begrenzter Genauigkeit nicht zu falschen Lösungen führt. Dies ist vor allem bei Computerberechnungen wichtig.

Sinn und Zweck der QR-Zerlegung

Das Lösen der Ausgleichrechnung führt im wesentlichen dazu, ein lineares Gleichungssystem zu lösen., die sogenannte Normalengleichung.
Wenn die Spalten der Fehlergleichungsmatrix linear unabhängig sind, dann besitzt die Normalengleichung eine eindeutige Lösung. Die Normalengleichung kann somit mit dem Gauss-Algorithmus oder mit einer Variante dieses Algorithmus der LR-Zerlegung gelöst werden.

Allerdings ergeben sich in der Praxis beim Lösen der Normalengleichung oft Genauigkeitsprobleme, falls diese nicht von Hand und somit mit algebraischer Genauigkeit gelöst wird. Wird die Normalengleichung mit dem Computer gelöst, kann es im Resultat zu grossen Fehlern kommen. Dies liegt daran, dass der Computer mit endlicher Genauigkeit arbeitet. Konkret bedeutet dies, dass der Computer z.B. nur 32 Stellen zur Verfügung hat um eine Zahl zu speichern. Wenn dann eine Zahl wie $\sqrt{2}$ gespeichert wird, entstehen Rundungsfehler, welche durch das Lösen der Gleichung zusätzlich verstärkt werden können. Ein weiteres Problem, das beim Lösen von Gleichungssystemen mit der LR-Zerlegung oder dem Gaussalgorithmus beobachtet wird, nennt sich Auslöschung. Es entsteht, wenn z.B. gerundet wird, oder ein „schlechtes" Pivotelement gewählt wird um die Gleichung zu lösen. Beim Auslöschen bekommt ein Element einen völlig falschen Wert, da ein extremer Genauigkeitsverlust bei der Subtraktion zweier fast gleich grossen Zahlen entsteht.

Diese Probleme kann man aber voraussehen, indem man die Kondition der Matrix berechnet. Die Kondition ist eine hier nicht weiter definierte Grösse, welche uns erlaubt einzuschätzen, ob mit endlicher Genauigkeit eine genaue Lösung gefunden werden kann.

Falls die Matrix schlecht konditioniert ist, ist das Lösen der Gleichung mit dem Gauss-Algorithmus (oder der LR-Zerlegung) keine gute Idee. Bei diesen Algorithmen werden schlecht konditionierte Matrizen keine gute Resultate liefern, da diese Algorithmen die Fehler in diesem Fall verstärken. Normalengleichungen sind in der Regel schlecht konditionierte Matrizen, was dazuführt, dass Computer nicht in der Lage sind die Lösung genau zu berechnen.

Die Lösung des Problems liegt also darin, ein numerisch stabiles Verfahren zu finden, welche den Fehler trotz der endlichen Genauigkeit sehr klein lässt.

Die QR-Zerlegung ist ein solches Verfahren, und deshalb wird für numerische Zwecke sehr oft dieser Algorithmus gebraucht.

Der Rechenaufwand wird durch die QR-Zerlegung nicht kleiner, jedoch wird die Genauigkeit verbessert. Deshalb ist die QR-Zerlegungen für Rechnungen, welche von Hand gelöst werden nicht sehr empfehlenswert, bzw. nicht nötig, da keine Approximationen gemacht werden (algebraische Genauigkeit).

Beispiele

Historisch gesehen wurde die Methode der kleinsten Quadrate für Himmelskörper entwickelt, deshalb wird hier auch ein Beispiel aus der Astrologie genommen.

Ein Astrophysiker beobachtet drei Sterne, welche er originellerweise a, b und c nennt. Diese 3 Sterne liegen alle schön auf einer Gerade. Nun misst er die Distanzen zwischen den Sternen. Dabei erhält er folgende Messdaten:

- Distanz (a,b) = 1000
- Distanz (a,c) = 3800
- Distanz (b,c) = 3000

Der Astrophysiker bemerkt, dass er wohl einen Messfehler gemacht hat, denn falls die Sterne auf einer Geraden liegen, so sollte Distanz (a,c) = Distanz (a,b) + Distanz (b,c) sein. Dies ist aber nicht erfüllt! (Für müde Köpfe: 1000 + 3000 = 4000 ≠3800...)
Dummerweise weiss er nicht, was er falsch gemacht hat. Allerdings kann er nicht einfach seine Messung verwerfen oder verändern, denn es könnte die richtige Messung sein.

Also stellt er folgendes (überbestimmtes) Gleichungssystem auf:

$y_1 = 1000$ (= Distanz (a,b))
$y_2 = 3000$ (= Distanz (b,c))
$y_3 = 3800$ (= Distanz (a,c))
$y_1 + y_2 - y_3 = 0$

Daraus entsteht folgende Fehlergleichung:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} - \begin{pmatrix} 1000 \\ 3000 \\ 3800 \\ 0 \end{pmatrix} = r$$

Lösung mit QR-Zerlegung

Zuerst wird aus A Q und R berechnet. Dafür gibt es mehrere Wege (z.B. Gram-Schmidt Verfahren, Givens-Rotationen) Dabei gilt A = QR

Alle nachfolgenden Rechnungen wurden mithilfe von MATLAB gemacht.

Man erhält

$$Q = \begin{pmatrix} \frac{-\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{-\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & \frac{\sqrt{6}}{3} & \frac{-\sqrt{3}}{6} & -\frac{1}{2} \\ 0 & 0 & \frac{-\sqrt{3}}{2} & \frac{1}{2} \\ \frac{-\sqrt{2}}{2} & \frac{-\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2} \end{pmatrix} \qquad R = \begin{pmatrix} -\sqrt{2} & \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{-\sqrt{6}}{2} & \frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{-2\sqrt{3}}{3} \\ 0 & 0 & 0 \end{pmatrix}$$

Danach berechnet man den Vektor d mit $Q^T \cdot c$

$$d = Q^T \cdot \begin{pmatrix} 1000 \\ 3000 \\ 3800 \\ 0 \end{pmatrix} \text{ wobei } d = 1000 \cdot \begin{pmatrix} -0.7071 \\ -2.0421 \\ -4.4456 \\ -0.1 \end{pmatrix}$$

danach löst man die Gleichung $R_0 x = d_0$. Das bedeutet, das nur die ersten n Zeilen gebraucht werden, in diesem Fall also 3.

Die Lösung erhält man durch Rückwärtseinsetzen. Somit erhält man dann für y_1 950 y_2 2950 und y_3 3850.
Dies ist die Ausgleichslösung für dieses Problem.

Lösung mit Gauss-Algorithmus

Dabei wird die Gleichung $A^T A \cdot y = A^{T} \cdot c$ mit dem Gauss-Algorithmus gelöst.
Siehe Beispiel mit Normalengleichung.

Quellen:

Lineare Algebra 5. Auflage K. Nipp und D. Stoffer 2002
Strang-linear Algebra and applications

http://de.wikipedia.org/wiki/Carl_Friedrich_Gau%C3%9F
http://de.wikipedia.org/wiki/Satz_von_Gau%C3%9F-Markov
http://de.wikipedia.org/wiki/Ausgleichsrechnung
http://de.wikipedia.org/wiki/Methode_der_kleinsten_Quadrate