

5 LINEARE AUSGLEICHSPROBLEME

Beispiel 5.0.1 Bestimmung eines unbekannten Widerstands x aus Messungen für die Stromstärke t und die Spannung b . Angenommen, es liegen m Messungen (b_i, t_i) , $i = 1, \dots, m$, mit $m \gg 1$,

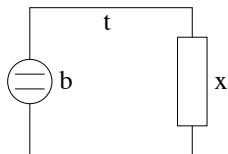


Abb. 5.1: Widerstandsbestimmung aus Messungen für Spannung und Stromstärke.

für Spannung und Stromstärke vor. Das **Ohm'sche Gesetz** aus der Physik besagt: $b = tx$, also

$$b_i = t_i x, \quad i = 1, \dots, m. \quad (5.1)$$

Da die Messdaten in der Regel (Mess-)Fehler beinhalten, kann man (5.1) nicht exakt für alle i erfüllen. Gesucht ist jetzt also ein Widerstand x , der „möglichst gut“ zu den Messdaten passt.

Motivation: Gesucht ist eine Gerade $g(x) = \alpha x + b$, deren y -Werte den kleinsten Quadratsummenabstand zu den vorgegebenen Daten $(x_i, f(x_i))$, $i = 1, \dots, n$ haben, die sogenannte **Ausgleichsgerade**. Die geometrische Veranschaulichung dazu ist in der nachfolgenden Abbildung dargestellt.

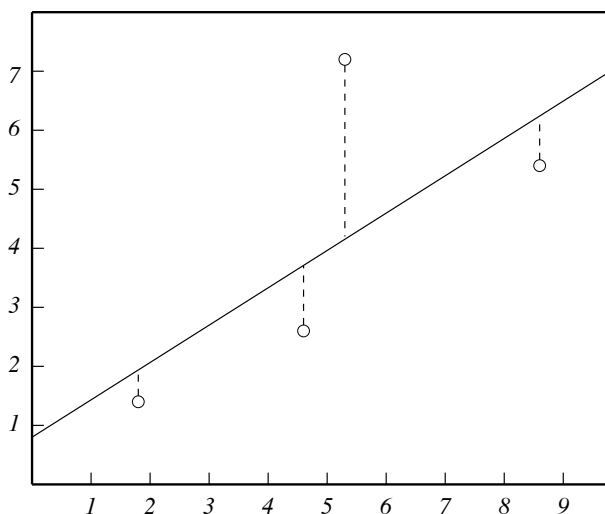


Abb. 5.2: Die Ausgleichsgerade, die die Quadratsumme der Abstände minimiert

5.1 Die Methode der kleinsten Quadrate

Gemäß obiger Motivation gilt es zu gegebenen Punktepaaaren $(x_1, f_1), (x_2, f_2), \dots, (x_n, f_n)$ die in der Geradengleichung $g(x) = mx + b$ freien Parameter m und b so zu bestimmen, dass gilt

$$F(m, b) := \sum_{i=1}^n (g(x_i) - f_i)^2 \rightarrow \min,$$

oder in alternativer Formulierung

$$\min_{(b,m) \in \mathbb{R}^2} \left\| A \begin{pmatrix} b \\ m \end{pmatrix} - d \right\|_2^2,$$

wobei wir nachfolgende Definitionen verwenden

$$A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad d = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

Hierzu bilden wir alle partiellen Ableitungen bzgl. m und b , woraus sich dann die kritischen Punkte der Funktion $F(m, b)$ wie folgt berechnen

$$\nabla F(m, b) = \left(\frac{\partial}{\partial b} F, \frac{\partial}{\partial m} F \right) = \vec{0}.$$

Im vorliegenden Fall erhalten wir damit

$$\begin{aligned} \frac{\partial}{\partial b} F(m, b) &= 2 \sum_{i=1}^n (mx_i + b - f_i) \cdot 1 \stackrel{!}{=} 0, \\ \frac{\partial}{\partial m} F(m, b) &= 2 \sum_{i=1}^n (mx_i + b - f_i) \cdot x_i \stackrel{!}{=} 0. \end{aligned}$$

Diese beiden Bestimmungsgleichungen können auch wie folgt geschrieben werden

$$\begin{aligned} 1^T \left(A \begin{pmatrix} b \\ m \end{pmatrix} - d \right) &= 0, \\ x^T \left(A \begin{pmatrix} b \\ m \end{pmatrix} - d \right) &= 0, \end{aligned}$$

wobei $1 = (1, \dots, 1)^T \in \mathbb{R}^n$ und $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ seien.
Es gilt also

$$\begin{aligned} \nabla F(m, b) = 0 &\Rightarrow 1^T \left(A \begin{pmatrix} b \\ m \end{pmatrix} - d \right) = 0, \\ &\quad x^T \left(A \begin{pmatrix} b \\ m \end{pmatrix} - d \right) = 0. \end{aligned}$$

Zusammengefasst folgt für potentielle Minimalstellen

$$\begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}^T \left(A \begin{pmatrix} b \\ m \end{pmatrix} - d \right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow A^T A \begin{pmatrix} b \\ m \end{pmatrix} = A^T d$$

Bemerkung 5.1.1 Falls in einem linearen Gleichungssystem die Anzahl der Gleichungen die der Unbekannten übersteigen sollte (wie es im vorliegenden Fall bei $Ax = d$ ist), so nennen wir das System **überbestimmt** und können es in der Regel nicht exakt lösen. Deshalb betrachten wir sinnvollerweise das Ausgleichsproblem der **Gaußschen Normalengleichung**

$$A^T A x = A^T d \tag{5.2}$$

und erhalten dafür genau dann eine eindeutige Lösung, falls A vollen Rang besitzt.

Bemerkung 5.1.2 Als einführendes Beispiel haben wir uns auf die Betrachtung von linear unabhängigen Basisvektoren des $\mathbb{P}_1$ in der Darstellung von g beschränkt, d.h. $1, x$. Als Nächstes werden wir zu allgemeineren Funktionen in der Funktionsvorschrift von g übergehen, wie etwa

$$g(x) = a \exp(x) + b \exp(-x) + c \log(x).$$

Damit erhalten wir folgendes Minimierungsproblem bei bekannten Daten (x_i, f_i) ($i = 1, \dots, n$)

$$\sum_{i=1}^n (g(x_i) - f_i)^2 \rightarrow \min,$$

welches wir völlig analog zu obigem Verfahren als Normalengleichung in der Form

$$A^T A \begin{pmatrix} a \\ b \\ c \end{pmatrix} = A^T d, \text{ mit } A = \begin{pmatrix} e^{x_1} & e^{-x_1} & \log x_1 \\ \vdots & \vdots & \vdots \\ e^{x_n} & e^{-x_n} & \log x_n \end{pmatrix} \text{ und } d = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

interpretieren können. Je komplexer also die Funktionsvorschrift von g , desto aufwendiger wird auch das Lösen der Gaußschen Normalengleichung (5.2). Im Folgenden werden wir ein Verfahren bereitstellen, das eine numerisch stabile und effiziente Lösung des Ausgleichsproblems garantiert.

5.2 Die QR-Zerlegung

Definition 5.2.1 Eine quadratische Matrix $Q \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, falls gilt:

$$QQ^T = Q^T Q = I.$$

Weiterhin bezeichne $\|x\|$ die vom **Euklidischen** Skalarprodukt im $\mathbb{R}^n$ induzierte Norm, vgl. Anhang B. Darüber hinaus halten wir noch fest, dass orthogonale Abbildungen stets längen- und winkeltreu sind, dass also gilt

$$\|Qx\|^2 = (Qx)^T Qx = x^T Q^T Qx = x^T x = \|x\|^2, \text{ d.h. } \|Qx\| = \|x\|.$$

Motivation: Zur Lösung des Minimierungsproblems $\min_{x \in \mathbb{R}^n} \|Ax - b\|^2$ setzen wir voraus, dass zu $A \in \mathbb{R}^{m \times n}$ ($\text{rg } A = n < m$) stets eine Zerlegung $A = QR$ existiert, wobei $Q \in \mathbb{R}^{m \times m}$ orthogonal ist und $R \in \mathbb{R}^{m \times n}$ folgende Gestalt besitzt

$$R = \begin{pmatrix} * & * & * \\ & * & * \\ & & * \\ 0 & & \end{pmatrix} = \begin{pmatrix} \hat{R} \\ \hline 0 \end{pmatrix},$$

$\hat{R}$ stellt dabei eine reguläre obere Dreiecksmatrix dar. Damit können wir unser bisheriges Minimierungsproblem wie folgt modifizieren (beachte die Faktorisierung $A = QR$):

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|^2 = \min_{x \in \mathbb{R}^n} \|Q^T(Ax - b)\|^2 = \min_{x \in \mathbb{R}^n} \|Rx - Q^T b\|^2.$$

Hierbei weisen wir vor allen Dingen nochmals auf die Gestalt von R hin:

$$Rx = \begin{pmatrix} * & * & * \\ & * & * \\ & & * \\ 0 & & \end{pmatrix} \cdot x = \begin{pmatrix} b_1 \\ \hline b_2 \end{pmatrix} = Q^T b$$

Dieser Faktorisierungsansatz führt uns nun zu folgenden Fragen:

- Zu welchen Matrizen $A \in \mathbb{R}^{m \times n}$ existiert eine derartige **QR**-Zerlegung?
- Wie bestimmen wir eine solche Zerlegung möglichst effizient?

Satz 5.2.2 Zu jeder Matrix $A \in \mathbb{R}^{m \times n}$ mit Maximalrang $n < m$ existiert eine orthogonale Matrix $Q \in \mathbb{R}^{m \times m}$ derart, dass

$$A = QR \text{ mit } R = \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix}, \hat{R} \in \mathbb{R}^{n \times n}$$

gilt, wobei $\hat{R}$ eine reguläre obere Dreiecksmatrix darstellt.

Beweis. Der Kern des Beweises liegt in der Verwendung von orthogonalen Drehmatrizen der Form

$$U(p, q; \varphi) := \begin{pmatrix} 1 & & & & & & & & 0 \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & \cos \varphi & & \sin \varphi & & & \\ & & & -\sin \varphi & & \cos \varphi & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ 0 & & & & & & & & 1 \end{pmatrix} \quad (5.3)$$

$\begin{matrix} \downarrow p\text{-te Spalte} & \downarrow q\text{-te Spalte} \end{matrix}$
 $\begin{matrix} \leftarrow p\text{-te Zeile} \\ \leftarrow q\text{-te Zeile} \end{matrix}$

Die Orthogonalität von $U(p, q; \varphi)$ ergibt sich unmittelbar, denn es gilt $U^T(p, q; \varphi) \cdot U(p, q; \varphi) = I$. Das Produkt $\tilde{A} := U^T(p, q; \varphi) \cdot A$ liefert dagegen für $j = 1, \dots, n$

$$\begin{pmatrix} 1 & & & & & & & & 0 \\ & \ddots & & & & & & & \\ & & \cos \varphi & & -\sin \varphi & & & & \\ & & \sin \varphi & & \cos \varphi & & & & \\ & & & \ddots & & & & & \\ 0 & & & & & & & & 1 \end{pmatrix} \cdot A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{p-1,1} & \cdots & a_{p-1,n} \\ a_{p1} \cos \varphi - a_{q1} \sin \varphi & \cdots & a_{pn} \cos \varphi - a_{qn} \sin \varphi \\ a_{p+1,1} & \cdots & a_{p+1,n} \\ \vdots & & \vdots \\ a_{q-1,1} & \cdots & a_{q-1,n} \\ a_{p1} \sin \varphi + a_{q1} \cos \varphi & \cdots & a_{pn} \sin \varphi + a_{qn} \cos \varphi \\ a_{q+1,1} & \cdots & a_{q+1,n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

$\begin{matrix} \downarrow p\text{-te Spalte} & \downarrow q\text{-te Spalte} \end{matrix}$

Demzufolge werden die Zeilen p und q von A durch Linearkombinationen der cos- und sin-Terme ersetzt.

Die sukzessive Multiplikation der Matrix A von links mit Rotationsmatrizen $U^T(p, q; \varphi)$ (φ dabei so gewählt, dass $\tilde{a}_{q,p} = 0$) mit den Rotationsindexpaaren

$$(1, 2), (1, 3), \dots, (1, m), (2, 3), (2, 4), \dots, (2, m), (3, 4), \dots, (n, m)$$

eliminiert dann die Matrixelemente $a_{2,1}, a_{3,1}, \dots, a_{m,n}$. Sukzessive werden dabei unterhalb des Diagonalelements Nullen erzeugt. Nach $k := \frac{1}{2}n(2m - n - 1)$ Transformationsschritten gilt dann $A = QR$ mit

$$\mathcal{U}_k^T \cdot \dots \cdot \mathcal{U}_2^T \mathcal{U}_1^T A = Q^T A = R \quad \text{oder} \quad A = QR.$$

Da die orthogonale Matrix Q regulär ist, ist der Rang von A und der Rang von R gleich n und folglich ist die Rechtecksdreiecksmatrix $\hat{R}$ regulär, da A nach Voraussetzung Maximalrang hat. $\square$

Der soeben erbrachte Existenzbeweis ist rein konstruktiv und liefert damit eine erste Möglichkeit eine Zerlegung QR zu bestimmen. Wir unterscheiden die folgenden beiden Vorgehensweisen:

- **Givens**¹-Rotation,
- **Householder**²-Spiegelung.

5.3 Givens-Rotationen

Es bezeichne im Folgenden stets $c := \cos \varphi$ und $s := \sin \varphi$. Die Grundidee der **Givens**-Rotation besteht darin, durch die Multiplikation mit einer geeigneten orthogonalen (Dreh-)Matrix einen Vektor $x \in \mathbb{R}^n$ so zu drehen, dass möglichst viele seiner Komponenten verschwinden.

Im Falle $(a, b) \in \mathbb{R}^2$ werden wir also $c, s \in \mathbb{R}$ so bestimmen, dass gilt

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}.$$

Setzen wir zuerst $a, b \neq 0$ voraus. Wir erhalten damit folgende Gleichung

$$as = bc \tag{5.4}$$

und mit $s^2 + c^2 = 1$ gilt

$$r^2 = (ac + bs)^2 = a^2 c^2 + abcs + abcs + b^2 s^2 \stackrel{(5.4)}{=} a^2 c^2 + a^2 s^2 + b^2 c^2 + b^2 s^2 = a^2 + b^2.$$

Weiter ergibt sich für $a \neq 0$

$$\begin{aligned} ac + bs = r &\Rightarrow c = \frac{1}{a}(r - bs) &\Rightarrow as - \frac{b}{a}(r - bs) = 0 \\ &\Leftrightarrow a^2 s - br + b^2 s = 0 \\ &\Leftrightarrow br = (a^2 + b^2)s \\ &\Leftrightarrow s = \frac{br}{a^2 + b^2} \stackrel{(r^2=a^2+b^2)}{=} \frac{b}{r}. \end{aligned}$$

Und damit erhalten wir dann auch für $b \neq 0$

$$bc = as \Rightarrow c = \frac{a}{b}s = -\frac{a}{r}.$$

Offen ist noch die Wahl des Vorzeichens des Vorzeichens $r = \pm\sqrt{a^2 + b^2}$. Setzen wir $r := \text{sign}(a)\sqrt{a^2 + b^2}$, so folgt $c > 0$, falls $a \neq 0$, und für $a = 0$ setzen wir $c = 0$ und $s = -1$. Somit sind auch s und c für die anderen Spezialfälle $b = 0$ und $a = b = 0$ wohldefiniert. Damit

¹Givens, W.

²Householder, A.

gilt (obwohl der Winkel bisher gar nicht explizit genannt wurde) $\varphi \in (-\frac{\pi}{2}, \frac{\pi}{2}]$ und somit lässt sich $\cos \varphi$ eindeutig aus $\sin \varphi$ mit $\cos \varphi = \sqrt{1 - \sin^2 \varphi}$ bestimmen.

Seien p und q zwei Zeilenindizes von A . Zur Eliminierung des Matricelements $a_{q,j}$ gehen wir daher wie folgt vor:

$$\begin{aligned} \text{Falls } a_{p,j} \neq 0: \quad \cos \varphi &= \frac{|a_{p,j}|}{\sqrt{a_{p,j}^2 + a_{q,j}^2}}, \\ \sin \varphi &= \frac{-\operatorname{sign}(a_{p,j}) a_{q,j}}{\sqrt{a_{p,j}^2 + a_{q,j}^2}}, \\ a_{p,j} = 0: \quad \cos \varphi &= 0, \quad \sin \varphi = 1. \end{aligned}$$

Da sich $\cos \varphi$ aus $\sin \varphi$ bestimmen lässt, ermöglicht das Verfahren eine effiziente Speicherung, denn es genügt $\sin \varphi$ an den entsprechenden freiwerdenden Stellen in A abzulegen:

$$\begin{aligned} A &= \begin{pmatrix} a_{1,1}^{(0)} & \cdots & a_{1,n}^{(0)} \\ \vdots & & \vdots \\ a_{m,1}^{(0)} & \cdots & a_{m,n}^{(0)} \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{1,1}^{(1)} & \cdots & a_{1,n}^{(1)} \\ \sin \varphi_{2,1} & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ a_{3,1}^{(0)} & a_{3,2}^{(0)} & \cdots & a_{3,n}^{(0)} \\ a_{4,1}^{(0)} & a_{4,2}^{(0)} & \cdots & a_{4,n}^{(0)} \\ \vdots & & & \vdots \\ a_{m,1}^{(0)} & \cdots & \cdots & a_{m,n}^{(0)} \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{1,1}^{(2)} & \cdots & a_{1,n}^{(2)} \\ \sin \varphi_{2,1} & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ \sin \varphi_{3,1} & a_{3,2}^{(2)} & \cdots & a_{3,n}^{(2)} \\ a_{4,1}^{(0)} & a_{4,2}^{(0)} & \cdots & a_{4,n}^{(0)} \\ \vdots & & & \vdots \\ a_{m,1}^{(0)} & \cdots & \cdots & a_{m,n}^{(0)} \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} a_{1,1}^{(3)} & \cdots & a_{1,n}^{(3)} \\ \sin \varphi_{2,1} & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ \sin \varphi_{3,1} & a_{3,2}^{(2)} & \cdots & a_{3,n}^{(2)} \\ \sin \varphi_{4,1} & a_{4,2}^{(3)} & \cdots & a_{4,n}^{(3)} \\ \vdots & & & \vdots \\ a_{m,1}^{(0)} & \cdots & \cdots & a_{m,n}^{(0)} \end{pmatrix} \rightsquigarrow \begin{pmatrix} a_{1,1}^{(m-1)} & \cdots & a_{1,n}^{(m-1)} \\ \sin \varphi_{2,1} & a_{2,2}^{(1)} & \cdots & a_{2,n}^{(1)} \\ \sin \varphi_{3,1} & a_{3,2}^{(2)} & \cdots & a_{3,n}^{(2)} \\ \sin \varphi_{4,1} & a_{4,2}^{(3)} & \cdots & a_{4,n}^{(3)} \\ \vdots & & & \vdots \\ \sin \varphi_{m,1} & a_{m,2}^{(m-1)} & \cdots & a_{m,n}^{(m-1)} \end{pmatrix} \rightsquigarrow \begin{pmatrix} & & & \\ \ddots & & & \hat{R} \\ & \ddots & & \\ \text{„Q“} & & \ddots & \end{pmatrix}. \end{aligned}$$

Die **QR-Zerlegung** liefert also QR in kompakter Form, gespeichert in A .

Der Aufwand dieser Vorgehensweise liegt bei ungefähr $\mathcal{O}(n^3)$ Operationen.

Bemerkung 5.3.1 Die Multiplikation eines gegebenen Vektors $x \in \mathbb{R}^n$ mit der Drehmatrix $\mathcal{U}(p, q; \varphi)$ ist äquivalent zur Drehung von x um einen Winkel φ entgegen dem Uhrzeigersinn in der Koordinatenebene. Vergleiche hierzu die nachfolgende Abbildung, die diesen geometrischen Sachverhalt verdeutlicht: