Vorlesung: Numerik 1 für Ingenieure

Version 17.11.2014

Michael Karow

12. Vorlesung

Thema: Ausgleichsrechnung

Ausgleichsprobleme I

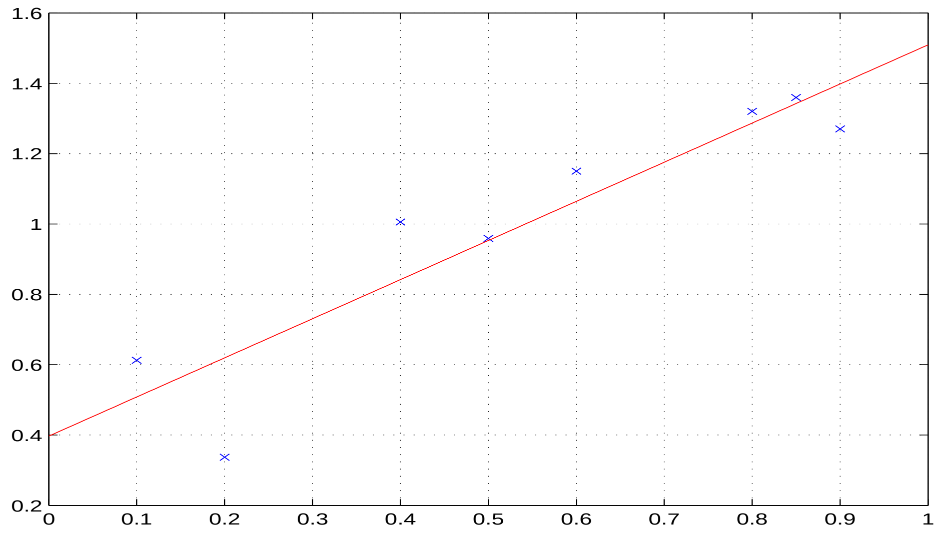
Problemstellung: Zwischen zwei (z.B. physikalischen) Größen x und y wird (z.B. aufgrund theoretischer Überlegungen) ein linearer Zusammenhang der Form

$$y = f(x) = c_1 + c_2\, x \qquad (*)$$

angenommen. Die unbekannten Parameter c_1, c_2 sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3), \quad \ldots, (x_m, y_m).$$

Aufgrund von Messfehlern oder weil die Ausgangshypothese $(*)$ nicht ganz korrekt ist, liegen die Messpunkte nicht auf einer Geraden. Was sind die besten Werte für c_1, c_2, die man in dieser Situation angeben kann?



Ausgleichsprobleme II

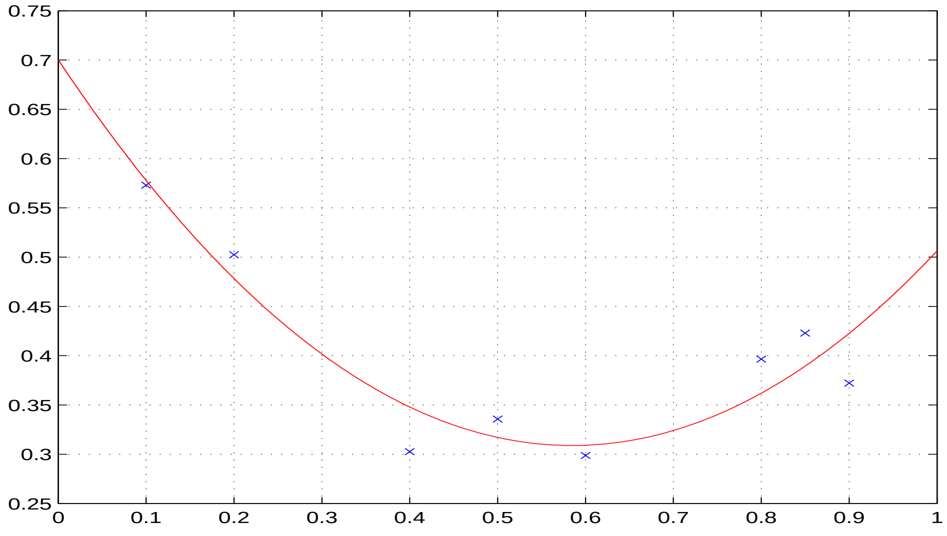
Problemstellung: Zwischen zwei Größen x und y wird ein quadratischer Zusammenhang der Form

$$y = f(x) = c_1 + c_2\,x + c_3\,x^2$$

angenommen. Die unbekannten Parameter c_1, c_2, c_3 sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3), \quad \ldots, (x_m, y_m).$$

Was sind die besten Werte für die c_k ?



Ausgleichsprobleme III

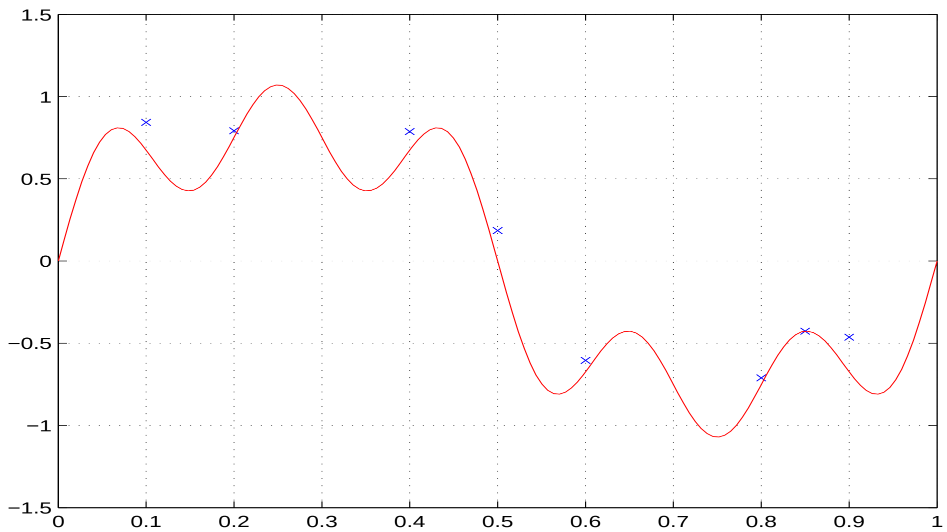
Problemstellung: Zwischen zwei Größen x und y wird ein Zusammenhang der Form

$$y = f(x) = c_1 \sin(2\pi\, x) + c_2 \sin(6\pi\, x) + c_3 \sin(10\pi\, x)$$

angenommen. Die unbekannten Parameter c_1, c_2, c_3 sollen durch Messungen bestimmt werden. Die Messungen ergeben die Wertepaare

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3), \quad \ldots (x_m, y_m).$$

Was sind die besten Werte für die c_k ?



Ausgleichsprobleme IV

Allgemeines lineares Ausgleichsproblem: Zwischen den Größen x und y wird ein Zusammenhang der Form

$$y = f(x) = c_1\,\beta_1(x) + c_2\,\beta_2(x) + \ldots + c_n\,\beta_n(x) \qquad (*)$$

angenommen. Dabei sind $\beta_1(x)$, . . . , $\beta_n(x)$ vorgegebene Funktionen (Basisfunktionen). Gegeben sind die Daten

$$(x_1, y_1), \quad (x_2, y_2), \quad (x_3, y_3), \quad \ldots (x_m, y_m),$$

wobei $m > n$. Wie muss man die freien Konstanten c_j wählen, so dass die Funktion $(*)$ die Daten optimal approximiert (annähert)?

Frage: Was heißt eigentlich 'optimal approximiert' ?

Auf diese Frage kann man verschiedene Antworten geben. In dieser VL besprechen wir die (in der Praxis am häufigsten vorkommende)

Optimalität im Sinne der kleinsten Fehlerquadrate

Ausgleichsprobleme V

Messwert zu x_j:	y_j
Funktionswert zu x_j:	$f(x_j) = c_1\,\beta_1(x_j) + c_2\,\beta_2(x_j) + \ldots + c_n\beta_n(x_j)$
Fehlerquadrat:	$(f(x_j) - y_j)^2 = (c_1\,\beta_1(x_j) + c_2\,\beta_2(x_j) + \ldots + c_n\beta_n(x_j) - y_j)^2$

Summe der Fehlerquadrate:

$$q = q(c_1, c_2, \ldots, c_n) = \sum_{j=1}^{m} (\, c_1\,\beta_1(x_j) + c_2\,\beta_2(x_j) + \ldots + c_n\beta_n(x_j) - y_j\,)^2.$$

Methode der kleinsten Fehlerquadrate (least squares method):

Bestimme die Parameter c_k so, dass q minimal wird.

$$q(c_1, c_2, \ldots, c_n) \quad \rightarrow \quad \text{min}$$

Zur Bearbeitung des Minimierungsproblems wird q zunächst in Matrix-Vektorform geschrieben (siehe nächste Seite).

Ausgleichsprobleme VI

Vektor der Daten: Vektor der Funktionswerte:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ y_m \end{bmatrix} \qquad F = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ f(x_m) \end{bmatrix} = \underbrace{\begin{bmatrix} \beta_1(x_1) & \beta_2(x_1) & \dots & \beta_n(x_1) \\ \beta_1(x_2) & \beta_2(x_2) & \dots & \beta_n(x_2) \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \beta_1(x_m) & \beta_2(x_m) & \dots & \beta_n(x_m) \end{bmatrix}}_{=:A} \underbrace{\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}}_{c}$$

Es ist

$$q(c) = q(c_1, c_2, \dots, c_n) = \sum_{j=1}^{m} (\, f(x_j) - y_j \,)^2 = \|F - y\|_2^2 = \|Ac - y\|_2^2.$$

Das Minimierungsproblem lautet in dieser Schreibweise

$$q(c) = \|Ac - y\|_2^2 \quad \to \quad \text{min}$$

Ausgleichsprobleme VII

Minimierungsproblem:

$$q(c) = \|Ac - y\|_2^2 \quad \rightarrow \quad \text{min} \qquad (*)$$

Man kann q folgendermaßen umschreiben.

$$\begin{aligned} q(c) &= \|Ac - y\|_2^2 \\ &= (Ac - y)^\top (Ac - y) \\ &= (c^\top A^\top - y^\top)(Ac - y) \\ &= c^\top A^\top A\, c - y^\top Ac - c^\top A^\top y + y^\top y \\ &= c^\top A^\top A\, c - 2\,(A^\top y)^\top c + y^\top y. \qquad (**) \end{aligned}$$

$q(c)$ ist somit eine **linear-quadratische Funktion** von c
(quadratischer Term: $c^\top A^\top A\, c$, linearer Term: $(A^\top y)^\top c$, konstanter Term: $y^\top y$.)

Die **allgemeine Form einer linear-quadratischen Funktion** ist:

$$\phi(x) = x^\top Ax - 2\,b^\top x + c, \qquad A = A^\top \in \mathbb{R}^{n\times n}, \quad b \in \mathbb{R}^n, \quad c \in \mathbb{R}. \tag{1}$$

Frage:

Unter welcher Bedingung hat eine solche Funktion ein Minimum, und wie findet man es?

Diese Frage wird auf den folgenden Seiten diskutiert.

Minimierung einer linear-quadratischen Funktion: 1. Methode. Taylorentwicklung

Linear-quadratische Funktion: $\phi(x) = x^\top Ax - 2\,b^\top x + c$.

Wir betrachten ϕ auf der Geraden $x + t\,v$, mit $v \in \mathbb{R}^n, t \in \mathbb{R}$.

Es ist

$$\begin{aligned}
\phi(x + t\,v) &= (x + t\,v)^\top A(x + t\,v) - 2\,b^\top(x + t\,v) + c \\
&= x^\top Ax + x^\top A(tv) + (tv)^\top Ax + (tv)^\top A(tv) - 2\,b^\top x - 2\,b^\top(tv) + c \\
&= \text{...(Terme sortieren, Symmetrie von } A \text{ ausnutzen)} \\
&= \phi(x) + 2\,[(Ax - b)^\top v]\,t + [v^\top Av]\,t^2.
\end{aligned}$$

Das ist bei festgehaltenem v die Gleichung einer Parabel (unabhängige Variable t).
Notwendig und hinreichend dafür, dass x Minimalstelle von ϕ ist, sind die Bedingungen

$$(1) \qquad 0 = \left.\frac{d}{dt}\right|_{t=0} \phi(x + t\,v) = 2\,(Ax - b)^\top v,$$

$$(2) \qquad 0 \leq \left.\frac{d^2}{dt^2}\right|_{t=0} \phi(x + t\,v) = 2\,v^\top Av.$$

Genau dann haben alle Parabeln nämlich ihr Minimum bei $t = 0$.

Die Bedingungen (1) und (2) sind genau dann erfüllt, wenn

$$Ax = b, \qquad \text{und} \qquad A \text{ positiv semidefinit.}$$

Minimierung einer linear-quadratischen Funktion: 2. Methode. Scheitelpunktsform

Linear-quadratische Funktion: $\phi(x) = x^\top A x - 2\, b^\top x + c$.

Skalarer Fall: $A = a \in \mathbb{R}$, $x \in \mathbb{R}$.

Scheitelpunktsform einer Parabelgleichung (wenn $a \neq 0$):

$$\phi(x) \;=\; a\, x^2 - 2\, b\, x + c \;=\; a \left(x - \frac{b}{a} \right)^2 + \left(c - \frac{b^2}{a} \right).$$

Folgerung: ϕ hat genau dann ein Minimum, wenn $a > 0$.

Das Minimum wird bei $x = b/a$ angenommen, und es ist $c - b^2/a$.

Scheitelpunktsform im vektoriellen Fall (wenn A invertierbar):

$$\phi(x) \;=\; x^\top A x - 2\, b^\top x + c \;=\; (x - A^{-1}b)^\top A (x - A^{-1}b) + \left(c - b^\top A^{-1} b \right).$$

Folgerung: ϕ hat genau dann ein Minimum, wenn A positiv definit ist.

Das Minimum wird bei $x = A^{-1}b$ angenommen, und es ist $c - b^\top A^{-1} b$.

Ausgleichsprobleme VIII

Minimierungsproblem: $q(c) = \|Ac - y\|_2^2 \quad \to \quad \text{min}$

Wir nehmen an, dass $A^\top A$ invertierbar (also positiv definit) ist
und bringen $q(c)$ in Scheitelpunktsform:

$$
\begin{aligned}
q(c) &= \|Ac - y\|_2^2 = (Ac - y)^\top (Ac - y) \\
&= (c^\top A^\top - y^\top)(Ac - y) \\
&= c^\top A^\top A\, c - y^\top Ac - c^\top A^\top y + y^\top y \\
&= c^\top A^\top A\, c - 2\,(A^\top y)^\top c + y^\top y \\
&= \left(c - (A^\top A)^{-1} A^\top y\right)^\top (A^\top A) \left(c - (A^\top A)^{-1} A^\top y\right) + \left(y^\top y - (A^\top y)^\top (A^\top A)^{-1} (A^\top y)\right).
\end{aligned}
$$

Folgerung: Das Minimum von q wird bei

$$c = (A^\top A)^{-1} A^\top y \qquad (*)$$

angenommen. Das Minimum ist

$$y^\top y - (A^\top y)^\top (A^\top A)^{-1} (A^\top y).$$

Bemerkungen zu (*):

(1) Die Matrix $A^+ := (A^\top A)^{-1} A^\top$ heißt **Moore-Penrose-Inverse** von A.
(Man kann sie auch für den Fall definieren, dass $(A^\top A)^{-1}$ nicht existiert.)

(2) Der Vektor c is Lösung der **Normalgleichung** $(A^\top A)c = A^\top y$.

Ausgleichsprobleme IX

Minimierungsproblem: $q(c) = \|Ac - y\|_2^2 \quad \to \quad \text{min}$

Lösung: $c = (A^\top A)^{-1} A^\top y. \qquad (*)$

Notation: Der von den Spalten von $A = [a_1 \ \ldots \ a_n]$ aufgespannte Unterraum von $\mathbb{R}^m$ ist

$$V \ := \ \{Az \mid z \in \mathbb{R}^n\} = \left\{ [a_1 \ \ldots \ a_n] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \,\middle|\, z_k \in \mathbb{R} \right\} = \left\{ \sum_{k=1}^n a_k z_k \,\middle|\, z_k \in \mathbb{R} \right\}.$$

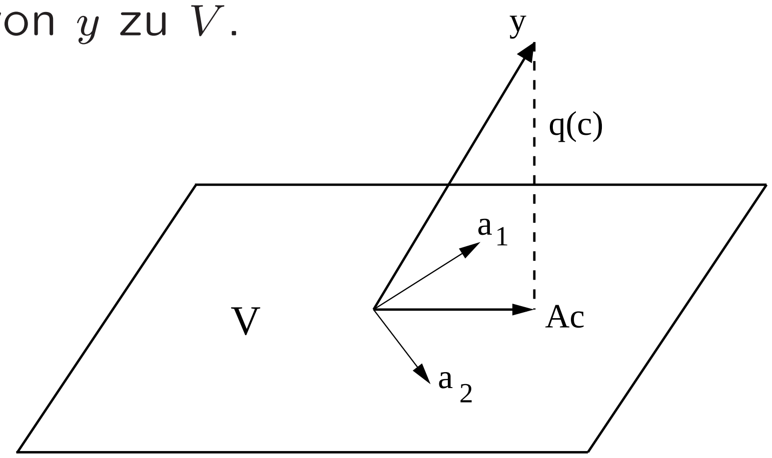
Geometrische Interpretation der Lösung des Minimierungsproblems:

Sei $c = (A^\top A)^{-1} A^\top y$ (Lösung der Normalgleichung). Dann gilt

- $Ac = \sum_{k=1}^n a_k c_k$ ist die **orthogonale Projektion** von y auf V,

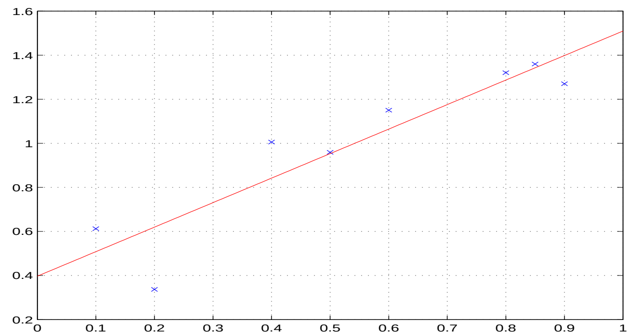
  denn $Ac - y$ steht senkrecht auf V, weil $(Az)^\top (Ac - y) = z^\top (A^\top A c - A^\top y) = 0$.

- $q(c)$ ist der Abstand von y zu V.



Ausgleichsprobleme X

Beispiel: Ausgleichsgerade zu den Wertepaaren $(x_j, y_j), \quad j = 1, \ldots, m$.



<u>Gesucht:</u> $f(x) = c_1 + c_2\,x$, so dass $q(c) = \sum_j (f(x_j) - y_j)^2$ minimal.

<u>Lösung:</u> $c = [c_1 \;\; c_2]^\top$ ist die Lösung der Normalgleichung

$$(A^\top A)\,c = A^\top y, \qquad \text{wobei} \qquad A = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}.$$

Es ist

$$A^\top A = \begin{bmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_m \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} = \begin{bmatrix} m & \sum_j x_j \\ \sum_j x_j & \sum_j x_j^2 \end{bmatrix},$$

$$A^\top y = \begin{bmatrix} 1 & \ldots & 1 \\ x_j & \ldots & x_j \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \sum_j y_j \\ \sum_j x_j y_j \end{bmatrix}.$$