8. Lineare Ausgleichsprobleme

8.1 Problemstellung

Von zwei skalaren Größen y und t sei bekannt, dass zwischen ihnen ein Zusammenhang

$$y = \phi(t; x_1, \dots, x_n) = \phi(t; x) \quad \underline{(8.1)}$$

besteht.

$x_1, \dots, x_n$: unbekannte Parameter

$\phi(t; x)$: Modellfunktion

Beispiele (8.2)

a) "Linearer Zusammenhang"

$$y = \phi(t; x_1, x_2) = x_1 + t\, x_2$$

b) "Polynomialer Zusammenhang"

$$y = \phi(t; x) = \sum_{i=0}^{n-1} x_{i+1} t^i$$

c) "Wachstumsfunktion"

$$y = \phi(t; \alpha, \lambda) = \alpha\, e^{\lambda t}$$

d) "Ellipse" Kurve $(x(t), y(t))$

$$\frac{x(t)^2}{a^2} + \frac{y(t)^2}{b^2} = 1$$

$\boxed{\text{1. Idee:}}$ Man könnte (t_i, y_i) n mal messen und $x = (x_1, .., x_n)^T$ aus den n Gleichungen

$$y_i - \phi(t_i; x) = 0 \quad (i = 1, \ldots, n)$$

bestimmen!

Für die obigen Beispiele würde das bedeuten:

a):
$$\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

b): ϕ: Interpolationspolynom zu (t_i, y_i)

oder:
$$\begin{bmatrix} 1 & t_1 & \ldots & t_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & t_n & \ldots & t_n^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

c)
$$\alpha e^{\lambda t_1} - y_1 = 0$$
$$\alpha e^{\lambda t_2} - y_2 = 0$$

(nichtlineares Gleichungssystem)

d) $$\begin{pmatrix} x_1^2 & y_1^2 \\ x_2^2 & y_2^2 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \; ; \; \alpha = \frac{1}{a^2} \, , \; \beta = \frac{1}{b^2}$$

Nachteil: Aufgrund der Messfehler in den y_i erhält man nur ungenaue x-Werte.

Abhilfe: Man verwende $m \gg n$ Messwerte und löse das überbestimmte Gleichungs=system möglichst genau!

Methode der kleinsten Quadrate (Gauß)

Setze zu $x \in \mathbb{R}^n$:

$$r(x) := \begin{bmatrix} \phi(t_1; x) - y_1 \\ \vdots \\ \phi(t_m; x) - y_m \end{bmatrix} \in \mathbb{R}^m \quad \text{(Residuum)}$$

und bestimme x so, dass

$$\|r\|_2 = \left(\sum_{j=1}^{m} (\phi(t_j; x) - y_j)^2 \right)^{1/2} \qquad (8.2)$$

minimal ist!

Hier: Einschränkung auf den Fall linearer Ansatzfunktionen

$$\phi(t,x) = \sum_{i=1}^{n} a_i(t)\, x_i \qquad (8.3)$$

Damit lautet das Residuum:

$$r(x) = \underbrace{\begin{bmatrix} a_1(t_1) & \dots & a_n(t_1) \\ \vdots & & \vdots \\ a_1(t_m) & \dots & a_n(t_m) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}}_{b}$$

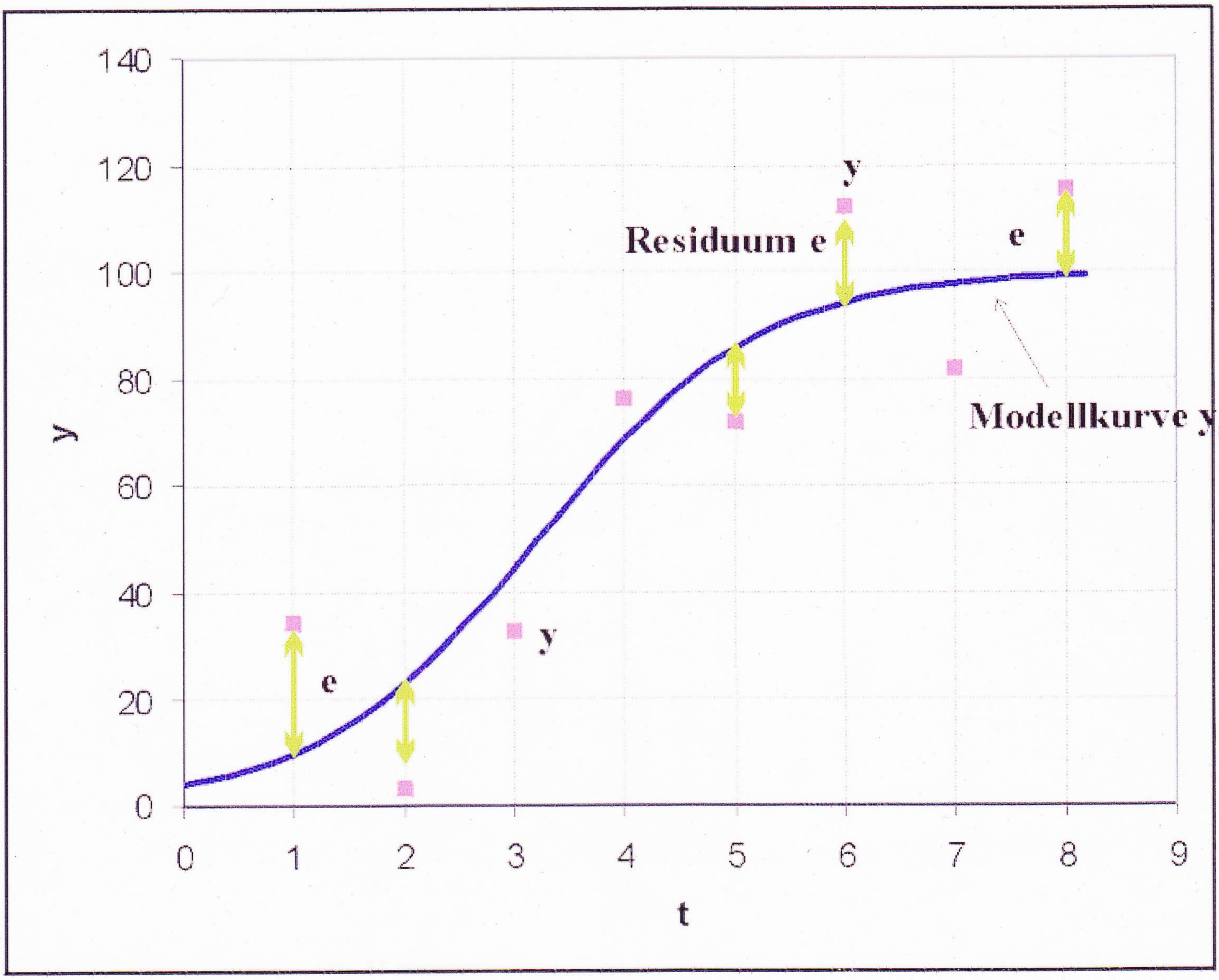
Problemstellung (8.4):

gegeben: $A \in \mathbb{R}^{(m,n)}$, $b \in \mathbb{R}^m$, $m \geq n$

gesucht: $x \in \mathbb{R}^n$ mit

$$\|r\|_2 = \|Ax - b\|_2 \text{ minimal!}$$

Aufgabe: Formulieren Sie die Beispiele a), b) und d) in der Form (8.4).



140
120
100
80
60
40
20
0
0
1
2
3
4
5
6
7
8
9
y
t
Residuum e
e
e
y
y
y
Modellkurve y_m


DEUTSCHE BUNDESBANK
Banknote
10
1777–1855 Carl Friedr. Gauß
10
AY7831976K1

Wilkinson - Givens - Young - Householder - Henrici - Bauer

8.2. Normalgleichung

Das Problem (8.4) ("lineares Ausgleichsproblem") ist eine typische Aufgabe der Approximation:

Der Vektor $b \in \mathbb{R}^m$ ist „möglichst gut" durch ein Element $z = Ax$ des linearen Teilraumes

$$U := \mathrm{Bild}(A) = \mathrm{Spann}(a^1, \dots, a^n)$$

zu approximieren:

Bestimme $z^* \in U$ mit

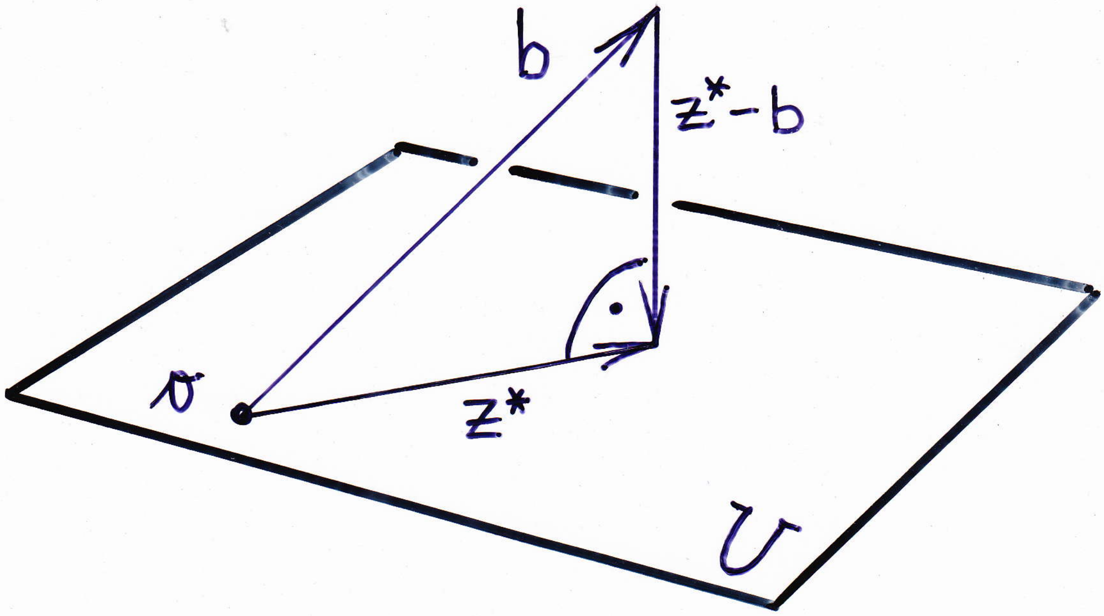
$$\|z^* - b\|_2 = \inf_{z \in U} \|z - b\|_2 \qquad (8.5)$$

Charakt.satz (8.6)

$(E, \langle .,. \rangle)$ Euklid. Vektorraum, $\|.\|$ zug. Norm.
$b \in E$, U linearer Teilraum von E.
Es sind äquivalent: $(z^* \in U)$

a) $\|z^* - b\| \leq \|z - b\| \quad (\forall z \in U)$

b) $\forall v \in U : \langle z^* - b, v \rangle = 0$



(Vgl. auch Satz (4.18) über Splines!)

Beweis:

a) $\Rightarrow$ b) (indirekt)

Wäre für ein $v \in U$: $\delta := \langle z^* - b, v \rangle \neq 0$,

so folgt für $z := z^* - \delta \frac{v}{\|v\|^2} \in U$:

$$\begin{aligned}
\|z-b\|^2 &= \langle z^* - \delta v/\|v\|^2 - b,\ z^* - \delta v/\|v\|^2 - b\rangle \\
&= \|z^*-b\|^2 - 2\frac{\delta}{\|v\|^2}\langle z^*-b, v\rangle \\
&\qquad + \frac{\delta^2}{\|v\|^4}\langle v,v\rangle \\
&= \|z^*-b\|^2 - \frac{\delta^2}{\|v\|^2} \\
&< \|z^*-b\|^2
\end{aligned}$$

Widerspruch!

b) $\Rightarrow$ a) Für $z \in U$ sei $v := z - z^*$.

$$\|z-b\|^2 = \|z^*-b+v\|^2$$

$$= \|z^*-b\|^2 + \underbrace{2\langle z^*-b, v\rangle}_{=0} + \|v\|^2$$

$$> \|z^*-b\|^2 \quad \text{für } v = z - z^* \neq 0$$

$\blacksquare$

Bemerkungen (8.7)

a) Der obige Beweis zeigt zugleich die Eindeutigkeit der Bestapproximation.
Die Existenz ist gesichert, falls U ein endlich dimensionaler Teilraum von E ist.

b) Die Abbildung $P: E \to U$, $P(b) = z^*$ heißt die orthogonale Projektion von E auf U.

c) Für das Problem (8.5) lautet die Orthogonalitätsbedingung:

$$\forall v \in \text{Bild}(A): \langle Ax^* - b, v\rangle = 0$$

$$\Leftrightarrow \forall x \in \mathbb{R}^n : (Ax^* - b)^T A x = 0$$

$$\Leftrightarrow \forall x \in \mathbb{R}^n : x^T \left[A^T A x^* - A^T b \right] = 0$$

$$\Leftrightarrow \quad \boxed{A^T A x^* = A^T b} \qquad (8.8)$$

(Normalgleichung)

d) $x^* \in \mathbb{R}^n$ löst also genau dann das lineare Ausgleichsproblem (8.5), wenn x^* die Normalgleichung erfüllt!

e) $A^T A \in \mathbb{R}^{(n,n)}$ symmetrisch, pos. semidefinit

$A^T A$ regulär $\Leftrightarrow$ Rang $(A) = n$

Numerische Lösung von (8.8) mittels Cholesky-Zerlegung von $A^T A$.

Aufwand:

Berechnung von $A^T A$: $\frac{1}{2} n^2 m$ wesentl. Op.

Cholesky-Zerlegung : $\frac{1}{6} n^3$ wesentl. Op.

Warnung: Die Normalgleichung ist häufig schlecht konditioniert und zwar schlechter als das ursprüngliche Ausgleichs=problem, d.h. die Lösung über (8.8) ist instabil! Vgl. dazu Aufgabe 26.

8.3 QR-Zerlegung

Eine Matrix $Q \in \mathbb{R}^{(m,m)}$ heißt orthogonal, falls gilt:

$$Q^T Q = I_m . \qquad (8.9)$$

Orthogonale Matrizen sind also insbesondere regulär mit $Q^{-1} = Q^T$. Wegen

$$\|Qx\|_2 = \left((Qx)^T Qx\right)^{1/2} = \left(x^T Q^T Q x\right)^{1/2} = (x^T x)^{1/2} = \|x\|_2$$

sind orthogonale Transformationen $y = Qx$ längentreu (längenerhaltend).

Die Grundidee besteht nun darin, das lin. Ausgleichsproblem $\min_{x \in \mathbb{R}^n} \|Ax - b\|_2$

durch orthogonale Transformationen in eine einfachere Form zu bringen, genauer:

$$\min_{x \in \mathbb{R}^n} \| Ax - b \|_2 \longrightarrow \min_{x \in \mathbb{R}^n} \| Q^T A x - Q^T b \|_2$$

mit $Q^T \in \mathbb{R}^{(m,m)}$ orthogonal.

Q^T wird dabei so bestimmt, dass $Q^T A$ eine obere Δ-Matrix wird:

$$Q^T A = \begin{bmatrix} R \\ 0 \end{bmatrix} \quad \text{oder} \quad A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} \qquad \underline{(8.10)}$$

(8.10) heißt demgemäß eine <u>QR-Zerlegung</u> der Matrix A.

Satz (8.11)

Jede Matrix $A \in \mathbb{R}^{(m,n)}$ mit $m \geq n$ und $\text{Rang}(A) = n$ besitzt eine QR-Zerlegung (8.10). Mit

$$Q^T b =: \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix}, \quad \tilde{b}_1 \in \mathbb{R}^n, \; \tilde{b}_2 \in \mathbb{R}^{m-n}$$

erhält man die Lösung x^* des Ausgleichsproblems aus $Rx = \tilde{b}_1$. Ferner ist

$$\| r \|_2 = \| Ax^* - b \|_2 = \| \tilde{b}_2 \|.$$

Beweis:

Existenz folgt später durch Angabe eines Algorithmus.

Weiter:

$$\|r\| = \|Ax - b\| = \|Q^T A x - Q^T b\|_2$$

$$= \left\| \begin{pmatrix} Rx - \tilde{b}_1 \\ -\tilde{b}_2 \end{pmatrix} \right\|_2$$

$$= \sqrt{\|Rx - \tilde{b}_1\|_2^2 + \|\tilde{b}_2\|^2}$$

Dieser Ausdruck wird minimal für $Rx = b_1$ ∎

Givens Rotationen *)

Dies sind Drehungen in der (k, l)-Ebene

$$\Omega_{k,l} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & 1 \ddots 1 & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow k \\ \\ \leftarrow l \\ \\ \\ \end{array} \in \mathbb{R}^{(m,m)} \qquad \underline{(8.12)}$$

wobei $c := \cos(\theta)$, $s := \sin(\theta)$.

---

*) James Wallace Givens (1953)
(1910–1993)

Der Winkel Θ wird so gewählt, dass das Element (ℓ,ν) in dem Produkt $\Omega_{k\ell} A$ verschwindet: (Spaltenindex ν)

$$\begin{bmatrix} \ddots & & & \\ & c & & s \\ & & \ddots & \\ & -s & & c \\ & & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ a_{k\nu} \\ \vdots \\ a_{\ell\nu} \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \tilde{a}_{k\nu} \\ \vdots \\ 0 \\ \vdots \end{bmatrix}$$

$$\tilde{a}_{k\nu} = \sqrt{a_{k\nu}^2 + a_{\ell\nu}^2}\,, \qquad \cos\Theta = \frac{a_{k\nu}}{\tilde{a}_{k\nu}}\,,$$

$$\sin\Theta = \frac{a_{\ell\nu}}{\tilde{a}_{k\nu}}\,.$$

<u>Eliminationsvorgang:</u>

$$\begin{bmatrix} * & \cdots & * \\ * & & * \\ \vdots & & \vdots \\ * & & * \\ * & \cdots & * \end{bmatrix} \xrightarrow{\Omega_{m-1,m}} \begin{bmatrix} * & \cdots & * \\ * & & * \\ \vdots & & \vdots \\ * & & * \\ 0 & \cdots & * \end{bmatrix} \xrightarrow{\Omega_{m-2,m-1}} \begin{bmatrix} * & \cdots & * \\ * & & * \\ \vdots & & \vdots \\ 0 & & * \\ 0 & \cdots & * \end{bmatrix} \cdots\cdots$$

$$\xrightarrow{\Omega_{1,2}} \begin{bmatrix} * & \cdots & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & & \vdots \\ 0 & * & & * \\ 0 & * & \cdots & * \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} * & \cdots & * \\ & * & \vdots \\ & & * \\ 0 & & * \\ & & \end{bmatrix}$$

$\Longrightarrow$

$$\begin{pmatrix} R \\ 0 \end{pmatrix} = \left(\Omega_{n,n+1} \cdots \Omega_{m-1,m}\right) \cdots \left(\Omega_{12} \cdots \Omega_{m-1,m}\right) \cdot A$$

Die $\Omega_{k,l}$ sind orthogonale Matrizen, daher ist auch das Produkt orthogonal.

Die Existenz einer QR-Zerlegung ist damit gezeigt!

Der Algorithmus wird auf die erweiterte Matrix (A,b) angewendet.

Aufwand: $2mn^2$ wesentl. Operationen

Householder Reflexionen *)

Eine Matrix $Q \in \mathbb{R}^{(m,m)}$ der Form

$$Q = I_m - 2\frac{vv^T}{v^Tv}, \quad v \neq 0 \qquad (8.13)$$

heißt Householder-Matrix.

Sie beschreibt geometrisch die Spiegelung des $\mathbb{R}^m$ an der Hyperebene $v^\perp$:

$$x = \alpha v \quad \Rightarrow \quad Qx = -x$$

$$x \perp v \quad \Rightarrow \quad Qx = x.$$

*) Alston Scott Householder (1904–1993)

Elementare Eigenschaften (8.14)

a) Q ist symmetrisch, also $Q^T = Q$

b) Q ist orthogonal, also $Q^T Q = Q Q^T = I_m$

c) Q ist involutorisch, also $Q^2 = I_m$

d) Q hat die Eigenwerte

$$\lambda_1 = +1 \quad (m-1)\text{-fach}$$
$$\lambda_2 = -1 \quad \text{einfach},$$

insbesondere: $\det Q = -1$, $\operatorname{cond}_2(Q) = 1$

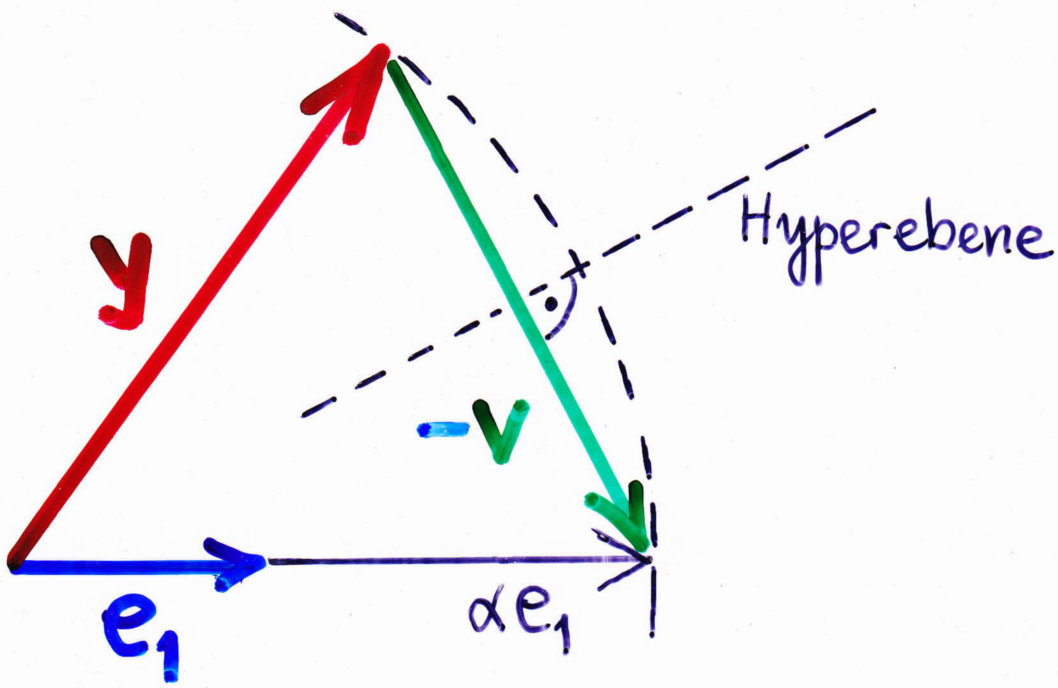
Bestimmung von Q:

Ein gegebener Vektor $y \in \mathbb{R}^m$ soll vermöge Q in ein Vielfaches von $e_1 = (1, 0, \dots, 0)^T$ abgebildet werden:

$$Qy = \alpha e_1$$

Längentreue $\Longrightarrow$ $\alpha = \pm \|y\|_2$

$$\Longrightarrow \quad \boxed{v = y - \alpha e_1, \quad \alpha = -\operatorname{sgn}(y_1)\|y\|_2} \tag{8.15}$$



Das Vorzeichen ist so gewählt, dass v auslöschungsfrei berechnet werden kann.

Mit (8.15), (8.13) findet man:

$$\boxed{Q = I_m - \frac{1}{\beta} v v^T, \quad \beta = \|y\|^2 + |y_1| \, \|y\|} \tag{8.16}$$

Zumeist wird anstelle von Q nur (v, β) gespeichert.

Produkt Qx : Für $x \in \mathbb{R}^m$ findet man

$$Qx = x - \frac{v^T x}{\beta} v. \tag{8.17}$$

Die Lösung des linearen Ausgleichsproblems erfolgt nun wie beim Givens-Verfahren:

$$(A|b) \xrightarrow{Q_1} \left[\begin{array}{cccc|c} \alpha_1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & * & \cdots & * & * \end{array}\right]$$

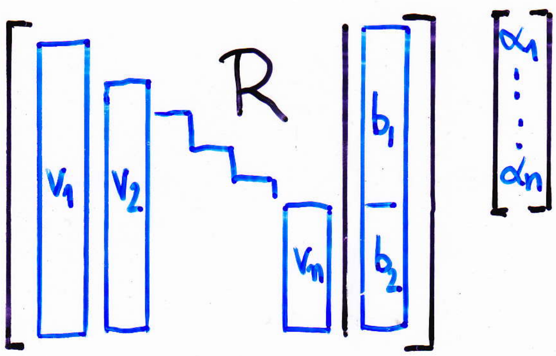
$$\xrightarrow{Q_2} \left[\begin{array}{ccccc|c} \alpha_1 & * & \cdots & \cdots & * & * \\ 0 & \alpha_2 & * & & \vdots & \vdots \\ \vdots & 0 & * & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & * & \cdots & * & * \end{array}\right]$$

$$\vdots$$

$$\xrightarrow{Q_n} \left[\begin{array}{ccc|c} \alpha_1 & \cdots & * & * \\ & R & \vdots & \vdots \\ & & \alpha_n & * \\ & 0 & & * \\ & & & \vdots \\ & & & * \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c} * \\ \vdots \\ * \end{array}}\right\} b_1 \\ \left.\vphantom{\begin{array}{c} * \\ \vdots \\ * \end{array}}\right\} b_2 \end{array}$$

$$\Longrightarrow \quad Rx = b_1, \qquad \|r\|_2 = \|b_2\|_2$$

<u>Speicherung:</u>



8.4. Verallgemeinerte Inverse

1. Für $A \in \mathbb{R}^{(m,n)}$, $m \geq n$, und Rang $A = n$ ist die Abbildung

$$b \in \mathbb{R}^m \longmapsto x^* \in \mathbb{R}^n \text{ s.t. } \|Ax^* - b\|_2 = \text{Min.!}$$

wohldefiniert und linear.

Normalgleichung $\Longrightarrow$ $x^* = (A^TA)^{-1}A^Tb$

Die zugehörige Abbildungsmatrix

$$\boxed{A^+ := (A^TA)^{-1}A^T} \qquad (8.18)$$

heißt Pseudoinverse von A.

Es gilt $A^+A = I_n$ und für $m=n$: $A^+ = A^{-1}$

2. Falls $m < n$ oder Rang $A = p < n$ ist $\|Ax-b\|_2 = \text{Min.!}$ zwar lösbar – aber nicht mehr eindeutig.

Der Lösungsraum ist affin-linear: $\bar{x} + \text{Kern}(A)$, $\bar{x}$: beliebige Lösung!

$\Longrightarrow$ Es gibt eine eindeutig bestimmte normminimale Lösung von $\|Ax-b\|_2 = \text{Min.}$

Diese werde mit x^* bezeichnet.

Die Abbildung $b \mapsto x^*$ ist dann linear, die zugehörige Abbildungsmatrix A^+ heißt wieder Pseudoinverse von A.

Wir entwickeln nun ein numerisches Verfahren zur Berechnung der Pseudoinversen von $A \in \mathbb{R}^{(m,n)}$; genauer zur Berechnung von

$$x = A^+ b \quad , \quad b \in \mathbb{R}^m \text{ vorgegeben}$$

Dazu bestimmen wir von A eine QR-Zerlegung mit Spaltenpivoting:

$$\boxed{AP = Q \begin{bmatrix} R & S \\ 0 & 0 \end{bmatrix}} \qquad (8.19)$$

$P = (e_{p_1} \dots e_{p_n}) \in \mathbb{R}^{(n,n)}$: Permutationsmatrix

$R \in \mathbb{R}^{(p,p)}$: reguläre obere Dreiecksmatrix

$Q \in \mathbb{R}^{(m,m)}$: orthogonale Matrix

$S \in \mathbb{R}^{(p,n-p)}$: kann ev. entfallen

I. Allg. wird jeweils eine Spalte maximaler Länge als Pivotspalte gewählt, so dass gilt:

$$r_{11} \geq r_{22} \geq \dots \geq r_{pp} > 0 .$$

Damit wird

$$\|Ax-b\|_2 = \|APy - b\|_2 \ ; \quad y := P^T x$$

$$= \left\| Q \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - b \right\|_2$$

$$= \left\| \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \tilde{b} \right\|_2 \ ; \quad \tilde{b} = Q^T b$$

$$= \left\| \begin{pmatrix} Ry_1 + Sy_2 - \tilde{b}_1 \\ -\tilde{b}_2 \end{pmatrix} \right\|_2$$

Damit ist der erste Teil des folgenden Satzes gezeigt:

Satz (8.20)

Es sei eine QR-Zerlegung von A gemäß (8.19) gegeben, $\tilde{b} := Q^T b = \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix}$ mit $\tilde{b}_1 \in \mathbb{R}^p$, $\tilde{b}_2 \in \mathbb{R}^{m-p}$, $y := P^T x = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ mit $y_1 \in \mathbb{R}^p$, $y_2 \in \mathbb{R}^{n-p}$.

a) x minimiert $\|Ax-b\|_2 \iff$

$$\iff Ry_1 + Sy_2 = \tilde{b}_1 . \quad \|Ax-b\|_2 = \|\tilde{b}_2\|_2 .$$

b) Mit $u := R^{-1}\tilde{b}_1$, $V := R^{-1}S$ gilt:

x normminimal $\iff$

$$\iff \begin{cases} (I + V^T V)\, y_2 = V^T u \\ y_1 = u - V y_2 \end{cases}$$

Beweis: zu b):

Nach obigem gilt:

$$\begin{aligned}\|x\|^2 = \|y\|^2 &= \|y_1\|^2 + \|y_2\|^2 \\ &= \|u - Vy_2\|^2 + \|y_2\|^2 \\ &= \|u\|^2 - 2\langle u, Vy_2\rangle + \langle Vy_2, Vy_2\rangle + \|y_2\|^2 \\ &= \|u\|^2 + y_2^T\{-2V^Tu + V^TVy_2 + y_2\} \\ &= \|u\|^2 + y_2^T\{(I+V^TV)\,y_2 - 2V^Tu\}\end{aligned}$$

Dies ist eine quadratische Form in y_2 mit positiv definiter Koeffizientenmatrix $(I+V^TV)$.

Diese besitzt ein eindeutig bestimmtes striktes globales Minimum, das durch

$$\nabla_{y_2}\|x\|^2 = 2\left[(I+V^TV)\,y_2 - V^Tu\right] = 0$$

bestimmt ist ∎

Bemerkungen:

a) Der Beweis zeigt zugleich, dass die Zuordnung $b \longmapsto x^*$ wohldefiniert und linear ist.

b) A^+ erfüllt die Penrose-Axiome *

- $(A^+A)^T = A^+A$
- $(AA^+)^T = AA^+$
- $A^+AA^+ = A^+$
- $AA^+A = A$

Algorithmus zu $x = A^+b$ (8.21)

1. QR-Zerlegung gemäß (8.19)
2. Löse $RV = S$ ((n-p) Rückwärts-Subst.)
3. Cholesky-Zerlegung $I + V^TV =: LDL^T$
4. $\begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{pmatrix} := Q^T b$
5. Löse $Ru = \tilde{b}_1$ (Rückwärts-Subst.)
6. Löse $LDL^T y_2 = V^T u$ (Vorwärts/Rückw.)
7. $y_1 := u - V y_2$
8. $x := Py$

---

*) Roger Penrose (1955)

8.5 Singulärwert-Zerlegung

Eine Zerlegung einer Matrix $A \in \mathbb{R}^{(m,n)}$ der Form

$$A = U S V^T \qquad \underline{(8.22)}$$

mit orthogonalen Matrizen $U \in \mathbb{R}^{(m,m)}$, $V \in \mathbb{R}^{(n,n)}$ und einer Diagonalmatrix

$$S = \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_p & \\ \hline & 0 & & 0 \end{array}\right) \in \mathbb{R}^{(m,n)}$$

mit $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p > 0$ heißt eine <u>Singulärwert-Zerlegung</u> von A, die σ_i heißen <u>singuläre Werte</u> von A.

MATLAB: $[U, S, V] = \mathrm{SVDS}(A, \ldots)$

<u>Bemerkungen (8.23)</u>

<u>a)</u> Die $\sigma_i \geq 0$ sind eindeutig bestimmt, denn

$$A^T A = V S^T U^T U S V^T = V (S^T S) V^T$$

$\Longrightarrow$ Die σ_i^2 sind die positiven Eigenwerte der symmetrischen und positiv semidefiniten Matrix $A^T A$.

b) Mit $V = (v_1, \dots, v_n)$, $U = (u_1, \dots, u_m)$ gilt:

$$A v_i = \sigma_i u_i \quad , \; i = 1, \dots, p$$
$$A v_i = 0 \quad , \; i = p+1, \dots, n$$

$$A^T u_i = \sigma_i v_i \quad , \; i = 1, \dots, p$$
$$A^T u_i = 0 \quad , \; i = p+1, \dots, m.$$

Weiter folgt:

$V = (v_1, \dots, v_n)$ ONB aus EVen von $A^T A$,

$U = (u_1, \dots, u_m)$ ONB aus EVen von $A A^T$.

c) $\|A\|_2 = \sigma_1$, $\quad \mathrm{cond}_2(A) = \dfrac{\sigma_1}{\sigma_p}$.

Satz (8.23)

Zu $A \in \mathbb{R}^{(m,n)}$ gibt es stets eine Singulär=wertzerlegung (8.22). Die Matrizen U, V sind jedoch i. Allg. nicht eindeutig bestimmt.

Beweis: Mit $\sigma_1 := \|A\|_2 = \max\limits_{\|x\|_2 = 1} \|Ax\|_2$

folgt: $\exists\, u_1 \in \mathbb{R}^m, v_1 \in \mathbb{R}^n, \|u_1\| = 1, \|v_1\| = 1$:

$$A v_1 = \sigma_1 u_1$$

Es ist $\sigma_1 > 0$ (wenn $A \neq 0$ vorausges. wird).

Man erweitere u_1 bzw. v_1 zu einer ONB $U = (u_1, .., u_m)$ von $\mathbb{R}^m$ bzw. $V = (v_1, .., v_n)$ von $\mathbb{R}^n$. Damit setze man

$$A_1 := U^T A V = \begin{pmatrix} u_1^T \\ \vdots \\ u_m^T \end{pmatrix} (Av_1, .., Av_n)$$

$$= \left(\begin{array}{c|c} \sigma_1 & w^T \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right) \qquad (\text{da } Av_1 = \sigma_1 u_1)$$

Wir zeigen, dass w verschwindet:

$$\sigma_1^2 = \|A\|_2^2 = \|A_1\|_2^2$$

$$\geq \frac{\left\|A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix}\right\|_2^2}{\left\|\begin{pmatrix} \sigma_1 \\ w \end{pmatrix}\right\|_2^2}$$

$$= \frac{(\sigma_1^2 + \|w\|^2)^2 + \|Bw\|^2}{\sigma_1^2 + \|w\|^2}$$

$$\geq \sigma_1^2 + \|w\|^2 \qquad \Longrightarrow \qquad \|w\| = 0$$

Damit ist gezeigt: $A_1 = U^T A V = \begin{pmatrix} \sigma_1 & 0 \dots 0 \\ 0 & \\ \vdots & B \\ 0 & \end{pmatrix}$

Die Behauptung folgt mit vollständ. Induktion.

▲

Bemerkung (8.24)

Mit Hilfe der Singulärwertzerlegung einer Matrix

$$A = USV^T, \quad S = \begin{pmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_p & \\ 0 & & & 0 \end{pmatrix}$$

lässt sich die Pseudoinverse folgendermaßen darstellen

$$A^+ = VS^+U^T, \quad S^+ := \begin{pmatrix} \sigma_1^{-1} & & & 0 \\ & \ddots & & \\ & & \sigma_p^{-1} & \\ 0 & & & 0 \end{pmatrix}.$$

Anwendung auf lineare Ausgleichsprobleme

$$\begin{aligned} \|r\|_2^2 &= \|Ax - b\|_2^2 \\ &= \|USV^Tx - b\|_2^2 \\ &= \|SV^Tx - U^Tb\|_2^2 \\ &= \|Sy - \tilde{b}\|_2^2, \quad y := V^Tx, \ \tilde{b} := U^Tb \\ &= \sum_{j=1}^{p} (\sigma_j y_j - \tilde{b}_j)^2 + \sum_{j=p+1}^{n} \tilde{b}_j^2 \end{aligned}$$

Damit folgt:

$$y_j = \tilde{b}_j / \sigma_j, \ j = 1, .., p \quad \& \quad y_{p+1}, .., y_n \text{ beliebig}$$

$$x = Vy = \sum_{j=1}^{n} y_j v_j$$

Die Lösung kleinster Norm erhält man, wenn man $y_{p+1} = \ldots = y_n = 0$ setzt.

Umformung:

$$x = \sum_{j=1}^{n} y_j v_j = \sum_{j=1}^{p} \frac{\tilde{b}_j}{\sigma_j} v_j = \sum_{j=1}^{p} \frac{u_j^T b}{\sigma_j} v_j = \left[\sum_{j=1}^{p} \frac{1}{\sigma_j} v_j u_j^T\right] b$$

und damit:

$$x = A^+ b,$$

$$A^+ = V S^+ U^T.$$