Normalengleichungen

Eine Näherungslösung für ein überbestimmtes lineares Gleichungssystem $Ax - b$ kann man durch Minimierung der Norm des Residuums $r = Ax - b$ erhalten:

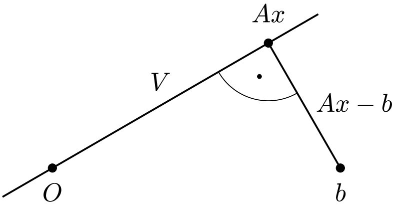
$$|Ax - b|^2 = \sum_{j=1}^{m} \left| \sum_{k=1}^{n} a_{j,k} x_k \; - b_j \right|^2 \to \min \; .$$

Man bestimmt eine beste Approximation $\sum_{k=1}^{n} x_k a_k$ zu b in dem von den Spalten $a_1, \ldots, a_n$ von A aufgespannten Unterraum $V = \operatorname{Bild} A$. Für eine reelle $m \times n$-Matrix A erfüllt jede Lösung x dieses sogenannten Ausgleichsproblems die Normalengleichungen

$$A^{\mathrm{t}} A \, x = A^{\mathrm{t}} b \, .$$

Dies bedeutet, dass das Residuum r (Fehler der Approximation) orthogonal zu dem von den Vektoren $a_1, \ldots, a_n$ aufgespannten Unterraum ist:

$$a_k^{\mathrm{t}} r = 0, \; \forall k \iff A^{\mathrm{t}} x - b \perp V \, .$$



Die Matrix $A^t A$ ist quadratisch und hat Dimension n. Sie ist genau dann invertierbar, wenn $\operatorname{Rang} A = n$, d.h. wenn die Spalten a_k von A linear unabhängig sind. In diesem Fall besitzt das Ausgleichsproblem eine eindeutige Lösung.

Die Normalengleichungen sind auch im singulären Fall lösbar. Das Residuum r ist auch dann eindeutig, die Lösung x jedoch nicht.

Alle Aussagen gelten, allgemeiner, ebenfalls für komplexe Matrizen; A^t ist dabei durch die adjungierte Matrix $A^* = \bar{A}^t$ zu ersetzen.

Beweis

Minimalität von $x \quad \Longleftrightarrow \quad \forall t \in \mathbb{R},\, y \in \mathbb{R}^n$:

$$|A(x+ty)-b|^2 \geq |Ax-b|^2, \quad \forall t \in \mathbb{R}, \forall y \in \mathbb{R}^n$$

Einsetzen der Definition der Norm,

$$\begin{aligned} |A(x+ty)-b|^2 &= ((Ax-b)^{\mathrm{t}} + t(Ay)^{\mathrm{t}})(Ax-b+tAy) \\ |Ax-b|^2 &= (Ax-b)^{\mathrm{t}}(Ax-b)\,, \end{aligned}$$

und Vereinfachung mit $r = Ax - b \quad \rightsquigarrow$

$$p = \underbrace{tr^{\mathrm{t}}Ay + ty^{\mathrm{t}}A^{\mathrm{t}}r}_{2ty^{\mathrm{t}}A^{\mathrm{t}}r} + t^2y^{\mathrm{t}}A^{\mathrm{t}}Ay \geq 0$$

$(r^{\mathrm{t}}Ay = y^{\mathrm{t}}A^{\mathrm{t}}r)$
p: nicht-negative Parabel in t
$p \geq 0$ genau dann, wenn $y^{\mathrm{t}}(A^{\mathrm{t}}r) = 0$
y beliebig $\quad \Longrightarrow \quad A^{\mathrm{t}}r = (0, \ldots, 0)^{\mathrm{t}}$

Reguläres und singuläres Ausgleichsproblem

(i) Rang A maximal:

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

Normalengleichungen

$$\underbrace{\begin{pmatrix} 5 & 1 \\ 1 & 5 \end{pmatrix}}_{A^t A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 3 \\ 3 \end{pmatrix}}_{A^t b}$$

eindeutige Lösung $x = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}^t$ mit Residuum

$$r = Ax - b = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

(ii) Rang A nicht maximal:

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 2 \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

linear abhängige Spalten $\quad \rightsquigarrow \quad$ singuläre Normalengleichungen

$$\begin{pmatrix} 5 & 10 \\ 10 & 20 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

Lösung

$$x = \begin{pmatrix} 3/5 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}$$

nicht eindeutig, aber eindeutiges Residuum

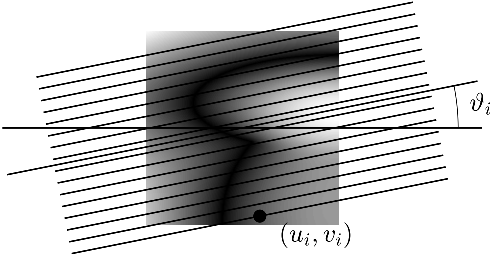
$$r = Ax - b = \begin{pmatrix} 2 & 4 \\ 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3/5 - 2t \\ t \end{pmatrix} - \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 6/5 \\ -12/5 \\ 0 \end{pmatrix}$$

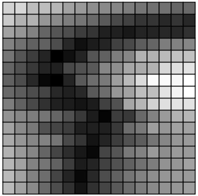
Beispiel

Computer-Tomographie:

Rekonstruktion einer Dichte $x(u, v)$ aus dem Intensitätsverlust von Röntgenstrahlen entlang von k Bündeln aus ℓ parallelen Geraden

$$\mathcal{R}_i : \ (u_i, v_i) + \mathbb{R}(\cos\vartheta_i, \sin\vartheta_i), \quad i = 1, \ldots, m = k\ell$$





Approximation von x durch eine stückweise konstante Funktion auf einem Raster von Quadraten $\mathcal{Q}_j$ und eine Näherung für die Linienintegrale

$$b_i = \int_{\mathbb{R}} x(u_i + t\cos\vartheta_i, v_i + t\sin\vartheta_i)\,\mathrm{d}t \approx \sum_{j=1}^{n} a_{i,j} x_j$$

mit x_j einer Approximation von $x(u,v)$ auf $\mathcal{Q}_j$ und

$$a_{i,j} = |\mathcal{R}_i \cap \mathcal{Q}_j|$$

der Länge des Durchschnitts der Geraden $\mathcal{R}_i$ mit dem Quadrat $\mathcal{Q}_j$

$m \gg n \;\rightsquigarrow\;$ Ausgleichsproblem zur Bestimmung von x aus den Daten b

Abbildung:
16×16 Raster, Winkel

$$\vartheta = 0, \pi/32, 2\pi/32, \ldots$$

mit 16 parallelen Scan-Richtungen im Abstand der Rasterquadratbreite
$\rightsquigarrow \quad 32 \cdot 16 = 512$ Gleichungen für $16 \cdot 16 = 256$ Unbekannte