1. Berechnen Sie die LR-Zerlegung der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 2 \\ -1 & -6 & -2 \end{pmatrix}.$$

(10)

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}}_{L_1} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ -1 & 2 & 2 \\ -1 & -6 & -2 \end{pmatrix}}_{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & -4 & 1 \end{pmatrix}$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{L_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & -4 & 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}}_{=: R}$$

$\Rightarrow \quad L_2 L_1 A = R$

$$\Rightarrow \quad A = LR \quad \text{mit} \quad L = L_1^{-1} L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

2. Zeigen Sie, dass die LR-Zerlegung einer Matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ ohne Pivotierung

$$N \doteq \frac{1}{3} n^3$$

Multiplikationen benötigt. (5)

Im k-ten Schritt wird eine $k \times k$ Blockmatrix unter Beibehaltung der ersten Zeile auf eine Nullspalte und eine $(k-1)\times(k-1)$ Blockmatrix reduziert. Dazu sind

$$(k-1)\cdot k \doteq k^2$$

Multiplikationen nötig, insgesamt also asymptotisch

$$\sum_{k=2}^{n} k^2 \doteq \int_0^n k^2 \, dk = \frac{1}{3} n^3 .$$

3. Sei $b \in \mathbb{R}^m$ und $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ mit $m \geq n$ und maximalem Rang.

(a) Zeigen Sie, dass

$$\|b - Ax\|_2^2 =: f(x)$$

minimal wird genau dann wenn

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

(b) Zeigen Sie, dass

$$\|(A^T A)^{-1} A^T\|_2 = \frac{1}{\min\limits_{\|z\|_2=1} \|Az\|_2}\,.$$

(c) Schließen Sie daraus, dass für die Lösungsabbildung aus Teil (a),

$$b \mapsto (A^T A)^{-1} A^T b$$

gilt, dass die relative Konditionszahl bezüglich der 2-Norm der Abschätzung

$$\kappa_{\mathrm{rel}} \leq \frac{\|b\|_2}{\|Ax\|_2}\,\kappa_2(A)$$

genügt.

(d) Wie unterscheidet sich die relative Konditionszahl aus Teil (c) von der relativen Konditionszahl (ebenfalls bezüglich der 2-Norm) eines linearen Gleichungssystems mit quadratischer nichtsingulärer Matrix A? Wann ist das lineare Ausgleichsproblem wesentlich schlechter konditioniert?

Hinweis: Die Konditionszahl bezüglich der 2-Norm für eine nichtquadratische Matrix ist definiert

$$\kappa_2(A) = \frac{\max\limits_{\|z\|_2=1} \|Az\|_2}{\min\limits_{\|z\|_2=1} \|Az\|_2}\,.$$

(5+5+5+5)

(a) $$0 \equiv \delta f = \delta\left((b-Ax)^T (b-Ax)\right)$$

$$= (-A\delta x)^T (b-Ax) + (b-Ax)^T A\,\delta x$$

$$= -2\,\delta x^T A^T A x + 2\,\delta x^T A^T b$$

Dies verschwindet für alle $\delta x \in \mathbb{R}^n$ nur, wenn $-A^TAx + A^Tb = 0$

Da A maximalen Rang hat, diese notwendige Bedingung äquivalent zu

$$x = (A^TA)^{-1} A^T b.$$

Da $f \geq 0$ und $f \to \infty$ für $\|x\| \to \infty$ existiert ein Minimum. Damit ist die notwendige Bedingung auch hinreichend.

(b) Wir setzen $x = Az$. Da $\operatorname{Ker} A^T \perp \operatorname{Im} A$ und $\operatorname{Ker} A = \{0\}$

gilt
$$\max_{x \neq 0} \frac{\|(A^TA)^{-1}A^T x\|_2}{\|x\|_2} = \max_{z \neq 0} \frac{\|(A^TA) A^T A z\|_2}{\|Az\|_2}$$

$$= \max_{z \neq 0} \frac{\|z\|_2}{\|Az\|_2}$$

$$\Rightarrow \|(A^TA)^{-1}A^T\|_2 = \max_{\|z\|_2=1} \frac{1}{\|Az\|_2} = \frac{1}{\min\limits_{\|z\|_2=1} \|Az\|_2}$$

(c) Die Lösungsabbildung ist linear mit Matrix $(A^TA)^{-1}A^T$, also gilt

$$\kappa_{abs} = \|(A^TA)^{-1}A^T\|_2$$

$$\Rightarrow \kappa_{rel} = \kappa_{abs} \frac{\|b\|_2}{\|x\|_2} = \kappa_{abs} \underbrace{\frac{\|Ax\|_2}{\|x\|_2}}_{\leq \max\limits_{\|z\|_2=1} \|Az\|_2} \frac{\|b\|_2}{\|Ax\|_2} \leq \frac{\|b\|_2}{\|Ax\|_2} \kappa_2(A)$$

(d) Wenn b im Bild von A liegt, gilt $b = Ax$, also $\kappa_{rel} \leq \kappa_2(A)$.

Dies entspricht dem Fall einer invertierbaren quadratischen Matrix.

Je größer der Winkel zwischen b und $\operatorname{Im} A$, desto größer $\frac{\|b\|_2}{\|Ax\|_2}$.
Dieser Faktor geht gegen ∞, wenn $b \perp \operatorname{Im} A$.

4. Gegeben sei die Funktion $f\colon [0,\infty) \to \mathbb{R}$,

$$f(x) = 1 - \frac{x}{x+1}. \quad = \frac{x+1-x}{x+1} = \frac{1}{x+1}$$

(a) Die Funktion werde ohne Umformung in Fließkommaarithmetik berechnet. Zeigen Sie, dass diese Berechnung für große x numerisch instabil ist.

(b) Formen Sie die Funktion so um, dass sie uneingeschränkt stabil in Fließkommaarithmetik berechnet werden kann. Führen Sie den Beweis der numerischen Stabilität.

(5+5)

(a) $x \hat{+} 1 = (x+1)(1+\delta_1)$ mit $|\delta_1| \leq \varepsilon_M$

$$x \hat{/} (x \hat{+} 1) = x/(x+1) \, \frac{1+\delta_2}{1+\delta_1} \quad \text{mit } |\delta_2| \leq \varepsilon_M$$

$$\doteq \frac{x}{x+1} (1+\delta_2)(1-\delta_1)$$

$$\doteq \frac{x}{x+1} (1+\delta_3) \quad \text{mit } |\delta_3| \leq 2\varepsilon_M$$

$$\Rightarrow 1 \hat{-} x \hat{/} (x \hat{+} 1) \doteq \left(1 - \frac{x}{x+1}(1+\delta_3)\right)(1+\delta_4) \quad \text{mit } |\delta_4| \leq \varepsilon_M$$

$$\doteq \left(1 - \frac{x}{x+1}\right)(1+\delta_4) - \frac{x}{x+1}\delta_3$$

$$\Rightarrow |\hat{f} - f| \doteq |f| \, |\delta_4| + \left|\frac{x}{x+1}\right| |\delta_3|$$

$$\Rightarrow \frac{|\hat{f}-f|}{|f|} \dot{\leq} \varepsilon_M + |x| \, 2\varepsilon_M$$

Kann beliebig groß werden, der Stabilitätsindikator ist also unbeschränkt!

(b) $f(x) = \frac{1}{x+1} \Rightarrow \hat{f} = \frac{1}{x+1} \, \frac{1+\delta_2}{1+\delta_1}$ (genau wie oben)

$$\doteq \frac{1}{x+1} (1+\delta_3) \quad \text{(ebenso wie oben)}$$

$\Rightarrow \frac{|\hat{f}-f|}{|f|} \leq 2\varepsilon_M$, d.h. $\sigma = 2$. Die Berechnung ist stabil.

5. Zu festen $x_1, \ldots, x_m$ seien fehlerbehaftete Messwerte $y_1, \ldots, y_m$ gegeben. Es wird eine quadratische Abhängigkeit angenommen, also

$$y = c_0 + c_1 x + c_2 x^2 .$$

Es gilt, die Koeffizienten c_0, c_1 und c_2 so zu bestimmen, dass der quadratische Fehler

$$E = \sum_{i=1}^{m} (y_i - (c_0 + c_1 x_i + c_2 x_i^2))^2$$

minimal wird.

(a) Formulieren Sie diese Aufgabe als lineares Ausgleichsproblem in Matrixschreibweise.

(b) Das Ausgleichsproblem kann sowohl (i) über die QR-Zerlegung oder (ii) über die Cholesky-Zerlegung der Normalgleichung (vergl. auch Aufgabe 3) gelöst werden. Beurteilen Sie beide Lösungswege hinsichtlich der Anzahl der Rechenoperationen sowie der numerischen Stabilität.

(5+5)

(a) Sei $A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_m & x_m^2 \end{pmatrix}$ $\quad b = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ $\quad c = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix}$

Dann läßt sich das Problem schreiben als: Bestimme $c \in \mathbb{R}^3$ so, dass $\| Ac - b \|_2$ minimal wird.

(b) Berechnung von $A^T A$ und $A^T b$ braucht $3 \cdot 3 \cdot m + 1 \cdot 3 \cdot m = 12m$ Multiplikationen.

Bei QR sind es $4 \cdot 3 \cdot (m-1) + 4 \cdot 1 \cdot (m-1)$ Multikationen, also geringfügig mehr.

Alle weiteren Operationen sind $O(1)$.

Beide Verfahren sind stabil, allerdings kann die Lösung der Normalgleichung wesentlich schlechter konditioniert sein, da $\kappa(A^T A)$ oft wesentlich größer als $\kappa(A)$ ist.