Rechenaufwand der LR-Zerlegung:
$A \to A^{(1)}$: $n-1$ Divisionen, $(n-1)^2$ Multiplikationen und Additionen
$A \to L, R$: Also insgesamt $\sum_{j=1}^{n-1}(j^2+j) = \frac{n^3}{3} - \frac{n}{3}$ Multiplikationen und Divisionen
Hauptarbeit des Algorithmus liegt somit in der Berechnung der LR-Zerlegung.

Beachte:

$$\begin{aligned}\sum_{j=1}^{n-1} j^2 + \sum_{j=1}^{n-1} j &= \frac{(n-1)n(2n-1)}{6} + \frac{n(n-1)}{2} \\ &= \frac{n^3}{3} - \frac{n^2}{6} - \frac{2n^2}{6} + \frac{n}{6} + \frac{n^2}{2} - \frac{n}{2}.\end{aligned}$$

Speicherplatz: Da Elemente mit Werten 0 und 1 nicht notwendigerweise gespeichert werden müssen, lässt sich das Gaußsche Eliminationsverfahren bei Speicherung der Permutationsmatrix mit $n(n+2)$ Speicherplätzen realisieren. Die relevanten Einträge der Frobenius-Matrizen können im Array der Matrix A bzw. $A^{(k)}$ gespeichert werden. Die Projektionsmatrix P kann durch weitere n Speicherplätze repräsentiert werden.

Spaltenpivotwahl: Selbst wenn $a_{11} \neq 0$ bzw. im k-ten Schritt $a_{kk}^{(k-1)} \neq 0$ gilt, kann eine Zeilenvertauschung sinnvoll sein. Bei der Spaltenpivotwahl wählt man als Pivotelement im k-ten Schritt das Element $a_{jk}^{(k-1)}$ mit

$$|a_{jk}^{(k-1)}| = \max_{k \leq i \leq n} |a_{i,k}^{(k-1)}|.$$

Dies führt zu

$$|l_{ij}| \leq 1 \text{ für alle } i, j$$

und somit zu einer besseren Stabilität des Verfahrens (zur Stabilität: siehe unten).

Zwei Beispiele, welche die den Rest des Kapitels motivieren:

Beispiel 3. *(zur Kondition des Problems) Betrachte das Gleichungssystem*

$$\begin{pmatrix} 1 & 1 \\ 1 & 1-\epsilon \end{pmatrix} x = \begin{pmatrix} 4 \\ 4-\epsilon \end{pmatrix}.$$

Die Lösung ist offenbar

$$x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Ersetzen wir die rechte Seite durch

$$\bar{b} = \begin{pmatrix} 4+\epsilon \\ 4-2\epsilon \end{pmatrix},$$

wobei $0 < \epsilon << 1$ *sehr klein sein kann, so erhalten wir die Lösung*

$$\bar{x} = \begin{pmatrix} 1+\epsilon \\ 3 \end{pmatrix}.$$

Das Beispiel macht deutlich, dass "kleine" Störungen der Eingabedaten zu "großen" Änderungen in der Lösung führen können. Aber wie klein ist "klein" und wie groß ist "groß"? Um den Einfluss von diesen Störungen auf die Lösung messen zu können, beschäfftigen wir uns weiter unten mit Normen.

Wichtige Frage: Wie wirken sich Störungen der Eingabegrößen (hier A und b) auf die Lösung unabhängig vom gewählten Algorithmus aus? (Kondition des Problems)

Beispiel 4. *(zur Stabilität der Gauß-Elimination) Wir lösen das Gleichungssystem*

$$\begin{pmatrix} 5\cdot 10^{-3} & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1 \end{pmatrix}$$

in zweistelliger Gleitpunktrechnung, wobei wir als Pivotelement

a) *das Element $a_{11} = 5\cdot 10^{-3}$ wählen. Nach einem Schritt des Gaußschen Eliminationsverfahrens erhalten wir das System*

$$\begin{pmatrix} 5\cdot 10^{-3} & 1 \\ 0 & -200 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ -99 \end{pmatrix}$$

mit Lösung

$$x = \begin{pmatrix} 0 \\ 0.5 \end{pmatrix}.$$

b) *das Element $a_{21} = 1$ wählen. Wir erhalten nun nach Vertauschung der Zeilen und der Gauß-Elimination*

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}$$

mit Lösung

$$x = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

Beachte, dass für die exakte Lösung des Gleichungssystem gilt:

$$x = \begin{pmatrix} \frac{100}{199} \\ \frac{99}{199} \end{pmatrix} \approx \begin{pmatrix} 0.503 \\ 0.497 \end{pmatrix}.$$

Erklärung: Falls $|l_{21}|$ "groß" ist (hier $2\cdot 10^2$), gilt gemäß der Gauß-Elimination

$$\begin{aligned} a_{22}^{(1)} &= a_{22} - l_{21}a_{12} && \approx -l_{21}a_{12} \\ b_2^{(1)} &= b_2 - l_{21}b_1 && \approx -l_{21}b_1 \end{aligned}$$

und somit auch

$$x_2 = \frac{b_2^{(1)}}{a_{22}^{(1)}} \approx \frac{b_1}{a_{12}}.$$

Bei der Berechnung von x_1 kommt es jedoch zur Stellenauslöschung (vgl. Ü):

$$x_1 = \frac{b_1 - a_{12}x_2}{a_{11}}.$$

Der Ausweg hier ist ein Zeilentausch, d.h. die Anwendung der Spaltenpivotwahl. Wir können bei dieser Wahl $|l_{21}| \leq 1$ *bzw. allgemeiner* $|l_{ij}| \leq 1$ *für alle* i, j *garantieren. Tatsächlich kann aber auch bei der Gauß-Elimination mit Spaltenpivotwahl die ungünstige oben beschriebene Situation auftreten.*

Wichtige Frage: Wie wirken sich Rundungsfehler, welche während der Durchführung eines bestimmten Algorithmus entstehen, auf die Berechnung der Lösung aus? (Stabilität des Algorithmus)

1.2 Einschub: Gleitpunktrechnung, Matrixnormen

1.2.1 Gleitpunktrechnung

Die Menge der im Computer darstellbaren reellen Zahlen ist offenbar endlich. Bei der heute üblichen normalisierten Gleitpunktdarstellung (engl. floating point representation) wird eine Zahl dargestellt als

$$z = a \cdot d^e,$$

wobei die Basis d eine Zweierpotenz ist (in der Regel 2,8,16) und der Exponent e eine ganze Zahl mit

$$e_{min} \leq e \leq e_{max}.$$

Die so genannte *Mantisse* a ist entweder 0 oder eine Zahl mit $d^{-1} \leq |a| < 1$ der Form

$$a = v \sum_{i=1}^{l} a_i d^{-i},$$

wobei $a_i \in \{0, \ldots, d-1\}$, $a_1 \neq 0$ und v das Vorzeichen und l die Mantissenlänge bezeichnet. Die Mantissenlänge ist für die relative Genauigkeit der Darstellung verantwortlich. Jede Zahl $x \neq 0$ mit

$$d^{e_{min}-1} \leq |x| \leq d^{e_{max}}(1 - d^{-l})$$

lässt sich nach Rundung durch eine Gleitpunktzahl $\mathrm{rd}(x)$ darstellen: Sei

$$x = v\,\underbrace{(0.a_1 a_2 \ldots a_l a_{l+1} \ldots)}_{=a}\,d^e$$

mit $a_1 \neq 0$ und $a_i \in \{0, \ldots, d-1\}$. Wir definieren

$$\mathrm{rd}(x) = v \cdot a' d^e$$

mit

$$a' := \begin{cases} 0.a_1 a_2 \ldots a_l, & \text{falls } 0 \leq a_{l+1} < \frac{d}{2} \\ 0.a_1 a_2 \ldots a_l + d^{-l}, & \text{falls } a_{l+1} \geq \frac{d}{2}. \end{cases}$$

Offensichtlich gilt

$$|x - \mathrm{rd}(x)| \leq | x - g|$$

für alle anderen durch den Computer darstellbaren Zahlen (Maschinenzahlen) g.
Für den relativen Fehler von $\mathrm{rd}(x)$ gilt:

$$\begin{aligned} \frac{|x - \mathrm{rd}(x)|}{|x|} &\leq \frac{d}{2} \frac{d^{-(l+1)}}{|a|} \\ &\leq \frac{d}{2} d^{-l} \\ &= \frac{d^{(1-l)}}{2}. \end{aligned} \tag{1.7}$$

Wir bezeichnen die Zahl $eps := \frac{d^{(1-l)}}{2}$ als die (relative) Maschinengenauigkeit.
Gleichung (1.7) ist äquivalent zu

$$\mathrm{rd}(x) = x(1+\epsilon) \qquad \text{mit } |\epsilon| \leq eps. \tag{1.8}$$

Falls $|x|$ kleiner als die betragsmäßig kleinste Maschinenzahl $d^{e_{min}-1}$ ist, spricht man von Exponentenunterlauf (engl. underflow), im Fall $|x| > d^{e_{max}}(1-d^{-l})$ von Exponentenüberlauf (engl. overflow).

Das Resultat einer arithmetischen Operation $x \pm y,\ x \cdot y, x/y$ muss keine Maschinenzahl sein, selbst wenn es x und y sind. Wir definieren die so genannte Gleitpunktoperationen für zwei Maschinenzahlen x und y durch

$$\begin{aligned} x \hat{+} y &:= \mathrm{rd}(x+y) \\ x \hat{-} y &:= \mathrm{rd}(x-y) \\ x \hat{\cdot} y &:= \mathrm{rd}(x \cdot y) \\ x \hat{/} y &:= \mathrm{rd}(x/y). \end{aligned}$$

Offenbar gilt mit (1.8) ebenso

$$\begin{aligned} x \hat{+} y &= (x+y)(1+\epsilon_1) \\ x \hat{-} y &= (x-y)(1+\epsilon_2) \\ x \hat{\cdot} y &= (x \cdot y)(1+\delta_1) \\ x \hat{/} y &= (x/y)(1+\delta_2) \qquad |\epsilon_i|, |\delta_i| \leq eps. \end{aligned}$$

Bemerkung 1.
(i) *Die Gleitpunkt-Realisierung von $\circ \in \{+,-,\cdot,/\}$ ist im Allgemeinen nicht assoziativ, d.h. es kommt auf die Reihenfolge der auszuführenden Operationen an.*
(ii) *Für zwei Maschinenzahlen x, y gilt:*

$$x \hat{+} y = x, \text{ falls } |y| \leq \frac{eps}{d}|x|.$$

1.2.2 Matrixnormen

Ziel: Wir wollen Fehler und Abweichungen von Vektoren und Matrizen "messen", d.h. die "Größe" eines Vektors oder einer Matrix durch eine Zahl beschreiben.

Definition 1. *Eine Abbildung $\|.\| : V \to \mathbb{R}$, V ein Vektorraum, heißt eine Norm auf V, wenn gilt:*
(i) $\|v\| \geq 0$ *und* $(\|v\| = 0 \Leftrightarrow v = 0)$, *(positive Definitheit)*

(ii) $\|\alpha v\| = |\alpha|\|v\|$, *(Homogenität)*

(iii) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$ *(Dreiecksungleichung)*

für alle Vektoren $v, v_i \in V$ und $\alpha \in \mathbb{R}$.

Beispiel 5. *Wichtige Beispiele im $\mathbb{R}^n$ sind*
(i) $\|x\|_1 = \sum_{i=1}^n |x_i|$

(ii) $\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2}$

(iii) $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|$

Jede Norm auf $\mathbb{R}^{n\times n}$ heißt Matrixnorm. Von besonderem Interesse sind Matrixnormen, die zu einer gegebenen Vektornorm passen, d.h. es gilt

$$\|Ax\| \leq \|A\|\|x\| \tag{1.9}$$

für alle $x \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n\times n}$. Solche Normen sind hilfreich zur Herleitung von Abschätzungen.

Definition 2. *Sei $\|.\|$ eine beliebige Norm auf $\mathbb{R}^n$. Dann definieren wir die zugehörige Matrixnorm auf dem Raum der quadratischen $(n \times n)$-Matrizen durch*

$$\|A\| := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} \text{ für } A \in \mathbb{R}^{n\times n}.$$

Offenbar gilt für Matrixnormen der Definition 2 die Ungleichung (1.9), wobei $\|A\|$ die kleinste Zahl mit dieser Eigenschaft ist. Des Weiteren ist durch diese Matrixnorm tatsächlich eine Norm im Sinne von Definition 1 gegeben, d.h. es gelten die Eigenschaften (i)-(iii). Zusätzlich gilt

$$\begin{aligned}\|I\| &= 1\\ \|A \cdot B\| &\leq \|A\| \cdot \| B\|.\end{aligned}$$

Diese Abschätzung wird die *Submultiplikativität* der Matrixnorm genannt. Eine wichtige Beobachtung ist, dass die Matrixnorm aus Definition 2 von der speziellen Wahl der Norm auf $\mathbb{R}^n$ abhängt:

Satz 3. *Sei A eine quadratische $(n \times n)$-Matrix. Es gilt:*

(i) $\|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^n |a_{ij}|$ *(Spaltensummennorm)*

(ii) $\|A\|_2 = \sqrt{\text{größter EW von } A^TA}$ *(Spektralnorm)*

(iii) $\|A\|_\infty = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$ *(Zeilensummennorm)*

Ende des Einschubs.

1.3 Kondition linearer Gleichungssysteme

Wir wollen nun Normen benutzen, um bei einem linearen Gleichungssystem

$$Ax = b$$

den Einfluss von Abweichungen (Störungen) der Eingabegrößen A und b auf die Lösung x abzuschätzen.

Störungen der rechten Seite: Sei $\bar{x}$ die Lösung des Systems

$$Ax = \bar{b},$$

so gilt

$$x - \bar{x} \;=\; A^{-1}b - A^{-1}\bar{b} \;=\; A^{-1}(b - \bar{b})$$

und somit die Abschätzung der absoluten Abweichung

$$\underbrace{\|x - \bar{x}\|}_{\substack{\text{absolute Abweichung von}\\ \bar{x} \text{ zu } x \text{ gemessen in der}\\ \text{Norm } \|.\|.}} \;=\; \|A^{-1}(b - \bar{b})\| \;\leq\; \|A^{-1}\|\|b - \bar{b}\|. \tag{1.10}$$

Eine weitere aussagekräftige Größe ist die relative Abweichung von $\bar{x}$ zu x. Mit Abschätzung (1.10) folgt:

$$\underbrace{\frac{\|x-\bar{x}\|}{\|x\|}}_{\substack{\text{relative Abweichung von}\\ \bar{x} \text{ zu } x \text{ gemessen in der}\\ \text{Norm } \|.\|.}} \leq \frac{\|b\|\|A^{-1}\|}{\|x\|} \underbrace{\frac{\|b-\bar{b}\|}{\|b\|}}_{\substack{\text{relative Störung der}\\ \text{rechten Seite.}}}.$$

Mit $\|b\| = \|Ax\| \leq \|A\|\|x\|$ gilt:

$$\frac{\|x-\bar{x}\|}{\|x\|} \leq \|A\|\|A^{-1}\| \frac{\|b-\bar{b}\|}{\|b\|}. \tag{1.11}$$

Definition 3. *Wir nennen*

$$\operatorname{cond}(A) := \|A\|\|A^{-1}\|$$

die Konditionszahl der Matrix A.

Bemerkung 2. *Ungleichung* (1.11) *macht deutlich: Die Konditionszahl von A ist ein Maß der Sensitivität des relativen Fehlers gegenüber relativen Störungen der rechten Seite b. Diese Sensitivität scheint umso geringer desto kleiner* $\operatorname{cond}(A)$ *ist. Jedoch ist die Konditionszahl der Matrix A nur eine obere Schranke dieser Sensitivität und es gilt:*

$$\begin{aligned} 1 = \|I\| &= \|AA^{-1}\| \\ \leq \|A\|\|A^{-1}\| &= \operatorname{cond}(A). \end{aligned}$$

Für reelles $A = a$ ist die Konditionszahl minimal gleich 1.

Eigenschaften der Konditionszahl:

$$\operatorname{cond}(A) = \operatorname{cond}(\alpha A), \ \alpha \in \mathbb{R}\backslash\{0\}$$

$$\operatorname{cond}(A) = \frac{\max_{\|y\|=1} \|Ay\|}{\min_{\|z\|=1} \|Az\|} \tag{1.12}$$

Mit Gleichung (1.12) lässt sich die Kondition auch für nicht quadratische Matrizen formulieren.

Beispiel 6. *Wir betrachten das Gleichungssystem aus Beispiel 3 mit*

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1-\epsilon \end{pmatrix}.$$

Offenbar ist die Inverse gegeben durch

$$A^{-1} = -\frac{1}{\epsilon} \begin{pmatrix} 1-\epsilon & -1 \\ -1 & 1 \end{pmatrix}.$$

Für die Zeilensummennorm finden wir daher $\|A\|_\infty = 2$ bzw. $\|A^{-1}\|_\infty = \frac{2}{\epsilon}$ und somit

$$\operatorname{cond}_\infty(A) = \frac{4}{\epsilon}.$$

Für $b = (4,\ 4-\epsilon)^T$ und der Lösung $x = (3,\ 1)^T$ gilt zudem

$$\frac{\|b\|_\infty \|A^{-1}\|_\infty}{\|x\|_\infty} = \frac{8}{3\epsilon},$$

was die schlechte Konditionierung des Gleichungssystems in Beispiel 3 erklärt.

Störungen der Eingabegrößen A und b:

Satz 4. *Sei A eine invertierbare Matrix und*

$$Ax = b, \qquad \bar{A}\bar{x} = \bar{b}.$$

Seien weiter die relativen Abweichungen der Matrix $\bar{A}$ zu A und der rechten Seite $\bar{b}$ zu b beschränkt:

$$\frac{\|A-\bar{A}\|}{\|A\|} \leq \epsilon_A, \qquad \frac{\|b-\bar{b}\|}{\|b\|} \leq \epsilon_b.$$

Dann gilt die Abschätzung:

$$\frac{\|x-\bar{x}\|}{\|x\|} \leq \frac{\operatorname{cond}(A)}{1-\epsilon_A \cdot \operatorname{cond}(A)}(\epsilon_A+\epsilon_b)$$

falls $\epsilon_A \cdot \operatorname{cond}(A) < 1$.

Bemerkung 3. *Mit $\epsilon_A, \epsilon_b \leq \epsilon$ erhalten wir*

$$\frac{\|x-\bar{x}\|}{\|x\|} \leq 2\epsilon \cdot \operatorname{cond}(A) + \mathcal{O}(\epsilon^2).$$

Dabei bezeichnet $\mathcal{O}(\epsilon^2)$ eine Funktion, die selbst bei Division durch ϵ^2 im Grenzfall $\epsilon \to 0$ beschränkt bleibt.

Beweis: Offenbar gilt

$$\begin{aligned} b-\bar{b} &= Ax - \bar{A}\bar{x} \\ &= A(x-\bar{x}) + (A-\bar{A})\bar{x}. \end{aligned}$$

Nach Multiplikation mit A^{-1} erhalten wir entsprechend umgeformt

$$x-\bar{x} = A^{-1}\Big(b-\bar{b}-(A-\bar{A})\bar{x}\Big)$$

und somit die Abschätzung

$$\|x-\bar{x}\| \leq \|A^{-1}\|\Big(\underbrace{\|A-\bar{A}\|}_{\leq\epsilon_A\|A\|}\overbrace{\|\bar{x}\|}^{\leq\|x\|+\|x-\bar{x}\|}+\underbrace{\|b-\bar{b}\|}_{\epsilon_b\|b\|}\Big).$$

Mit $\|b\| \leq \|A\|\|x\|$ erhalten wir nach algebraischen Umformungen:

$$(1-\epsilon_A \cdot \operatorname{cond}(A))\|x-\bar{x}\| \leq \operatorname{cond}(A)\|x\|(\epsilon_A+\epsilon_b).$$

□

Bemerkung 4. *Nach Satz 4 mißt die Konditionszahl die relative Störempfindlichkeit der Lösung x von $Ax = b$ gegenüber relativen Abweichungen der Matrix A und der rechten Seite b. Sie ist aber nur eine obere Schranke dieser Störempfindlichkeit. Trotzdem ist die Abschätzung des Satzes optimal im folgenden Sinn: Für vorgegebene Matrix A lassen sich $\bar{A}$ und b finden, so dass Gleichheit gilt. Da wir aber nicht immer an beliebigen rechten Seiten interessiert sind und auch nicht beliebige Störungen zulassen, ist die Abschätzung des Satzes oft zu pessimistisch.*

Beispiel 7. *(Lubich) Betrachte das Gleichungssystem*

$$\begin{pmatrix} 1 & 1 \\ 0 & 10^{-8} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Es gilt

$$\begin{aligned} \mathrm{cond}_\infty(A) &= \|A\|_\infty \|A^{-1}\|_\infty \\ &= 2 \cdot 10^8. \qquad \text{(sehr groß)} \end{aligned}$$

Gestörtes System:

$$\begin{pmatrix} 1+\epsilon_1 & 1+\epsilon_2 \\ 0 & 10^{-8}(1+\epsilon_3) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} (1+\epsilon_4)b_1 \\ (1+\epsilon_5)b_2 \end{pmatrix}}_{=:\bar{b}},$$

wobei $0 \leq |\epsilon_i| \leq eps$ *mit der Maschinengenauigkeit eps. Wir untersuchen jetzt die Abhängigkeit der einzelnen Komponenten von den* ϵ_i*. Sei dazu* x *die Lösung des Ausgangssystems und* $\bar{x}$ *die des gestörten Systems.*

2.Komponente: Offenbar gilt:

$$\bar{x}_2 \;=\; \underbrace{10^8 b_2}_{=x_2} \frac{1+\epsilon_5}{1+\epsilon_3} \;=\; x_2\big(1 + \frac{\epsilon_5 - \epsilon_3}{1+\epsilon_3}\big)$$

und somit die Gleichheit

$$|x_2 - \bar{x}_2| = |x_2| \frac{|\epsilon_5 - \epsilon_3|}{|1+\epsilon_3|}.$$

Umgeformt:

$$\frac{|x_2 - \bar{x}_2|}{|x_2|} = \frac{|\epsilon_5 - \epsilon_3|}{|1+\epsilon_3|} \;\leq\; 2eps + \mathcal{O}(eps^2)$$

1.Komponente: Für $\bar{x}_1$ *finden wir*

$$\begin{aligned} \bar{x}_1 &= \big[(1+\epsilon_4)b_1 - (1+\epsilon_2)\bar{x}_2\big]/(1+\epsilon_1) \\ &= \big[\underbrace{b_1 - x_2}_{=x_1} + \epsilon_4 b_1 - \epsilon_2 \bar{x}_2 - x_2 \frac{\epsilon_5 - \epsilon_3}{1+\epsilon_3}\big]/(1+\epsilon_1) \\ &= x_1 + \big[\epsilon_4 b_1 - \epsilon_1 x_1 - \epsilon_2 \bar{x}_2 - x_2 \frac{\epsilon_5 - \epsilon_3}{1+\epsilon_3}\big]/(1+\epsilon_1). \end{aligned}$$

Mit $b_1 = x_1 + x_2$ *und* $\bar{x}_2 = x_2(1 + \frac{\epsilon_5 - \epsilon_3}{1+\epsilon_3})$ *erhalten wir die Darstellung*

$$\frac{|x_1 - \bar{x}_1|}{|x_i|} = \frac{1}{|x_i|}\big[(\epsilon_4 - \epsilon_1)x_1 + (\epsilon_4 - \epsilon_2 - \frac{\epsilon_5 - \epsilon_3}{1+\epsilon_3}(1+\epsilon_2))x_2\big]/(1+\epsilon_1)$$

und somit auch die Abschätzung

$$\frac{|x_1 - \bar{x}_1|}{|x_i|} \leq \big(2\frac{|x_1|}{|x_i|} + 4\frac{|x_2|}{|x_i|}\big)eps + \mathcal{O}(eps^2).$$

Insgesamt

$$\begin{aligned} \frac{|x_1 - \bar{x}_1|}{\|x\|_\infty} &\leq 6eps + \mathcal{O}(eps^2) \\ \frac{|x_2 - \bar{x}_2|}{\|x\|_\infty} &\leq 2eps + \mathcal{O}(eps^2). \end{aligned}$$

Wir betrachten allgemeiner die Situation:

$$\begin{aligned}\bar{a}_{ij} &:= a_{ij}(1+\epsilon_{ij}), && |\epsilon_{ij}| \le eps,\\ \bar{b}_i &:= b_i(1+\epsilon_i), && |\epsilon_i| \le eps.\end{aligned}$$

Offenbar gilt

$$\begin{aligned}\|A-\bar{A}\|_\infty &\le \|A\|_\infty eps,\\ \|b-\bar{b}\|_\infty &\le \|b\|_\infty eps.\end{aligned}$$

Betrachte nun alternativ das Gleichungssystem

$$DAx = Db \quad \text{mit } D = \operatorname{diag}(d_1,\dots,d_n),\ d_i \neq 0,$$

und das gestörte System

$$D\bar{A}x = D\bar{b}.$$

Seien x und $\bar{x}$ wieder die Lösungen der entsprechenden Systeme. Die Multiplikation des Systems mit einer invertierbaren Matrix von links ändert natürlich den Zusammenhang von x und $\bar{x}$ nicht, liefert aber bessere Abschätzung: Da

$$\begin{aligned}\|DA-D\bar{A}\|_\infty &\le \|DA\|_\infty eps\\ \|Db-D\bar{b}\|_\infty &\le \|Db\|_\infty eps\end{aligned}$$

folgt mit Satz 4

$$\begin{aligned}\frac{\|x-\bar{x}\|_\infty}{\|x\|_\infty} &\le \frac{\operatorname{cond}_\infty(DA)}{1-eps\cdot\operatorname{cond}_\infty(DA)}2eps\\ &= 2eps\cdot\operatorname{cond}_\infty(DA) + \mathcal{O}(eps^2).\end{aligned}$$

Wähle für obiges Beispiel konkret

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 10^8 \end{pmatrix}.$$

Damit

$$DA = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

und $\operatorname{cond}_\infty(DA) = 4$.

Beispiel 8. *(Deuflhard) Die Lösung des Gleichungssystems $Ax = b$ mit einer Diagonalmatrix*

$$A = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix}$$

ist offensichtlich ein gut konditioniertes Problem, da die Gleichungen entkoppelt sind (zwei unabhängige skalare Gleichungen). Andererseits ist aber

$$\operatorname{cond}_\infty(A) = \|A\|_\infty\|A^{-1}\|_\infty = \frac{1}{|\epsilon|}.$$

Die Konditionszahl gemessen in der Maximumsnorm $\|.\|_\infty$ wird daher beliebig groß für kleine $0 < |\epsilon| << 1$. Sie ist ein Maß der Sensitivität der Lösung gegenüber beliebigen Störungen, auch Störungen außerhalb der Hauptdiagonalen.

Matrizen mit kleiner Kondition:

(i) I, $\mathrm{cond}(\alpha I) = 1$

(ii) Orthogonale Matrizen $U^T U = I$. Denn es gilt:

$$\begin{aligned}\|Ux\|_2^2 &= x^T U^T \cdot Ux \\ &= x^T x \;=\; \|x\|_2^2\end{aligned}$$

und somit für die zugehörige Matrixnorm

$$\|U\|_2 = 1.$$

Da die Inverse $U^{-1} = U^T$ offenbar wieder orthogonal ist, gilt insgesamt

$$\mathrm{cond}_2(U) = 1.$$

(iii) Spline-Interpolation (später) führt auf Matrizen

$$A = \frac{1}{h}\begin{pmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{pmatrix}.$$

Da $\mathrm{cond}(A) = \mathrm{cond}(hA)$ gilt, gehen wir ohne Einschränkung von $h = 1$ aus. Es gilt weiterhin $\|A\|_\infty = 6$. Zur Bestimmung der Inversen von A schreiben wir

$$A = 4(I + N) \text{ mit } N \;=\; \begin{pmatrix} 0 & \frac{1}{4} & & & \\ \frac{1}{4} & 0 & \frac{1}{4} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{4} & 0 & \frac{1}{4} \\ & & & \frac{1}{4} & 0 \end{pmatrix}.$$

Nach dem Satz über die Neumann-Reihe gilt:

$$(I + N)^{-1} = \sum_{i=0}^{\infty} (-N)^i.$$

Denn: $(I + N) \cdot \sum_{i=0}^{\infty} (-N)^i \;=\; \sum_{i=0}^{\infty} (-N)^i - \sum_{i=0}^{\infty} (-N)^{i+1} \;=\; I$ und $\|N\|_\infty = \frac{1}{2} < 1$.
Damit folgt:

$$\begin{aligned}\|A^{-1}\|_\infty &= \frac{1}{4}\|(I + N)^{-1}\|_\infty \\ &\leq \frac{1}{4}(\|I\|_\infty + \|N\|_\infty + \|N\|_\infty^2 + \ldots) \\ &= \frac{1}{4}\sum_{i=0}^{\infty}\left(\frac{1}{2}\right)^i \;=\; \frac{1}{2}\end{aligned}$$

und für die Konditionszahl von A

$$\mathrm{cond}_\infty(A) \leq 3.$$

Die Matrix ist also unabhängig von h und n gut konditioniert.

Matrizen mit großer Kondition:

(i) Hilbertmatrizen $A = (\frac{1}{i+j-1})_{i,j=1,\ldots,n}$ also

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & & & \\ \frac{1}{3} & \frac{1}{4} & & & & \\ \frac{1}{4} & & & & & \\ \vdots & & & & & \\ & & & & & \end{pmatrix}$$

Es gilt:

n	cond(A)
1	1
2	27
3	740
4	2300
$\vdots$	
10	$3.5 \cdot 10^{13}$

(ii) Zu Beispiel 8:

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$$

mit $\max_j |a_j| >> \min_k |a_k|$. Dann gilt:

$$\mathrm{cond}_2 = \frac{\max_j |a_j|}{\min_k |a_k|} \;>>\; 1.$$