Prof. Dr. Dirk Lebiedz
Dipl.-Math. oec. Klaus Stolle
Institut für Numerische Mathematik
Universität Ulm

Angewandte Numerik 1
SoSe 2020
04.05.2020

# Lösungsvorschlag Übungsblatt 02

**Besprechung** in den Tutorien in der Woche vom 11.05.2020 bis 15.05.2020

Für dieses Übungsblatt gibt es 22 Theorie- und 12 Matlab-Punkte.
Die 70-Prozent-Grenzen liegen aktuell (inklusive Blatt 02) bei 41,3 Theorie- und 8,4 Matlabpunkten.

**Aufgabe 4** *(Relative Konditionszahl)* (3T+1T+2T+3T+3T Punkte)

Für $p, q \in \mathbb{R}$ mit $p^2 \geq -q$ (also $p^2 + q \geq 0$) sollen die Lösungen $x_1 \leq x_2$ der quadratischen Gleichung

$$x^2 + 2px - q = 0$$

berechnet werden. Diese Lösungen $x_1$ und $x_2$ hängen von den Parametern $p$ und $q$ ab. Wir wollen in dieser Aufgabe nur die Berechnung der größeren Nullstelle $x_2$ betrachten:

a) Berechnen Sie die Verstärkungsfaktoren $\phi_p(p, q)$ und $\phi_q(p, q)$. Überlegen Sie sich dazu zunächst, wie Sie die in die Verstärkungsfaktoren eingehende Funktion $f$ definieren müssen.

b) Geben Sie die relative Konditionszahl $\kappa_{rel}$ in Abhängigleit von $p$ und $q$ an.

c) Erklären Sie Ihren Kommilitonen und Ihrem Tutor, für welche $p$ und welche $q$ das Problem der Berechnung der größeren der beiden Nullstellen der quadratischen Gleichung gut konditioniert ist. Für welche $p$ und welche $q$ ist das Problem schlecht konditioniert?

d) Geben Sie ein konkretes gut konditioniertes und ein konkretes schlecht konditioniertes Beispiel an. Wie lauten die Verstärkungsfaktoren $\phi_p(p, q)$ und $\phi_q(p, q)$ und die relative Konditionszahl $\kappa_{rel}$ für diese Beispiele?

e) Skizzieren Sie beide Beispiele. Wie können Sie sich die Bedeutung der Verstärkungsfaktoren und der Konditionszahl anhand Ihrer Skizzen veranschaulichen?

**Lösungsvorschlag:**

a) Mit der Mitternachtsformel ist $f$ gegeben durch $f(p, q) = -p \pm \sqrt{p^2 + q}$.
Da $p^2 + q \geq 0$ gilt und nur die größere Nullstelle $x_2$ betrachtet werden soll, genügt es $f(p, q) = -p + \sqrt{p^2 + q}$ zu betrachten. Nach (Skript) sind die Verstärkungsfaktoren definiert durch

$$\phi_p(p, q) = \frac{|p|}{|f(p, q)|}\left|\frac{\partial}{\partial p} f(p, q)\right| \quad \text{und} \quad \phi_q(p, q) = \frac{|q|}{|f(p, q)|}\left|\frac{\partial}{\partial q} f(p, q)\right|.$$

Mit

$$\frac{\partial}{\partial p} f(p, q) = -1 + \frac{1}{2}\frac{2p}{\sqrt{p^2 + q}},$$
$$\frac{\partial}{\partial q} f(p, q) = \frac{1}{2}\frac{1}{\sqrt{p^2 + q}}$$

folgt

$$\phi_p(p,q) = \frac{|p|}{|-p+\sqrt{p^2+q}|}\left|-1+\frac{p}{\sqrt{p^2+q}}\right|$$

$$= \frac{|p|\left|\frac{-\sqrt{p^2+q}+p}{\sqrt{p^2+q}}\right|}{|-p+\sqrt{p^2+q}|}$$

$$= \frac{|p|}{\sqrt{p^2+q}}$$

$$\phi_q(p,q) = \frac{|p|}{|-p+\sqrt{p^2+q}|}\left|\frac{1}{2}\frac{1}{\sqrt{p^2+q}}\right|$$

$$\stackrel{(*)}{=} \frac{|-p+\sqrt{p^2+q}||-p-\sqrt{p^2+q}|}{2|-p+\sqrt{p^2+q}||\sqrt{p^2+q}|}$$

$$= \frac{|-(p+\sqrt{p^2+q})|}{2|\sqrt{p^2+q}|} = \frac{|p+\sqrt{p^2+q}|}{2|\sqrt{p^2+q}|}$$

$$(*) : |q| = |-q| = |-q+p^2-p^2| = |-p+\sqrt{p^2+q}| \cdot |-p-\sqrt{p^2+q}|$$

b)

$$\kappa_{rel}(f) = \kappa_{rel}^{\infty}(f) = \max\left(\frac{|p|}{\sqrt{p^2+q}}, \frac{|p+\sqrt{p^2+q}|}{2|\sqrt{p^2+q}|}\right)$$

c) Das Problem ist gut konditioniert, falls $\phi_p$ und $\phi_q$ klein sind, also insbesondere, wenn $q > 0$ ist. Dann sind $\phi_p$ und $\phi_q$ sogar $< 1$. Schlecht konditioniert ist das Problem, falls $\phi_p$ und $\phi_q$ groß sind, also insbesondere, wenn $p^2 + q \approx 0$.

d) gut konditioniert:

$$p = 1, \quad q = 25 \Rightarrow \phi_p(p,q) = 0.2041, \quad \phi_q(p,q) = 0.5103 \quad \Rightarrow \quad \kappa_{rel} = 0.5103$$

schlecht konditioniert:

$$p = 1, \quad q = -0.9999 \Rightarrow \phi_p(p,q) = 100, \quad \phi_q(p,q) = 0.50.5 \quad \Rightarrow \quad \kappa_{rel} = 100$$

e) Ist das Problem gut konditioniert, schneidet die Parabel die x-Achse relativ steil, wodurch eine kleine Änderungen bzw. Fehler in $p$ und $q$ auch nur kleine Auswirkungen auf die Schnittpunkte haben. Ist das Problem schlecht konditioniert, schneidet die Parabel die x-Achse relativ flach, wodurch sich kleine Änderunge bzw. Fehler stark auswirken (vgl. Abb. 1).

**Aufgabe 5** *(Matrixnormen (Spaltensummen- und Zeilensummen-Norm))* (3T+3T+4T Punkte)

In dieser Aufgabe wollen wir die beiden Formeln für die Zeilensummen- und die Spaltensummen-Norm aus Bemerkung 2.20 auf Folie 32 herleiten.

Wir betrachten die beiden Vektornormen

i) $\|\cdot\|_1 : \mathbb{R}^n \to \mathbb{R},\ v \mapsto \|v\|_1$ mit $\|v\|_1 := \sum_{k=1}^n |v_k| \quad \forall\, v = (v_1, \ldots, v_n)^\top \in \mathbb{R}^n$ (Summennorm) und

ii) $\|\cdot\|_\infty : \mathbb{R}^n \to \mathbb{R},\ v \mapsto \|v\|_\infty$ mit $\|v\|_\infty := \max_{k=1,\ldots,n} |v_k| \quad \forall\, v = (v_1, \ldots, v_n)^\top \in \mathbb{R}^n$ (Maximumnorm).

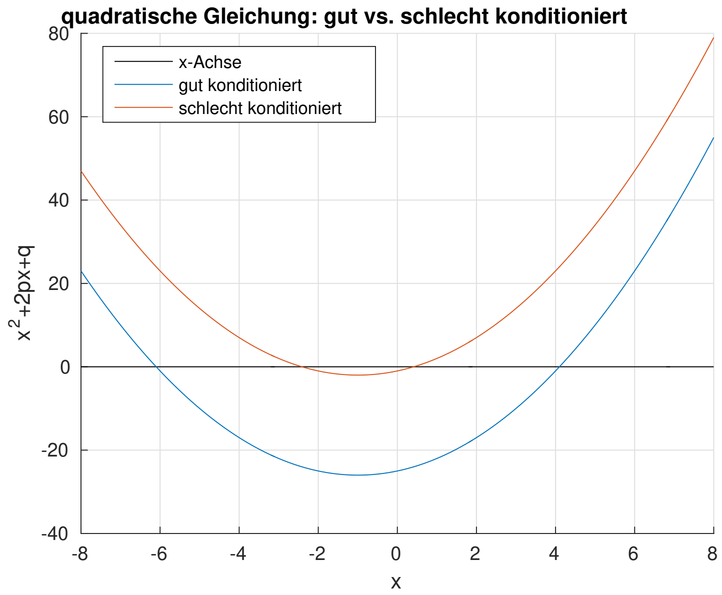


Abbildung 1: Gut und schlecht konditionierte Lösung der quadrateischen Gleichung.

Für die Vektorräume $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$, ausgestattet mit der $p$-Norm für $1 \leq p \leq \infty$, und die Matrix $B \in \mathbb{R}^{m\times n}$ definiert man die *von der Vektornorm induzierte Matrixnorm* als

$$\|B\|_p \; := \; \|B\|_{X\to Y} \; = \; \sup_{x\neq 0} \frac{\|B\,x\|_{Y,p}}{\|x\|_{X,p}} \; = \; \sup_{\|x\|_{X,p}=1} \|B\,x\|_{Y,p}.$$

a) Zeigen Sie: Die *Maximumnorm induziert die Zeilensummen-Norm*, also

$$\|B\|_\infty \; = \; \max_{i=1,\ldots,m} \sum_{j=1}^{n} |b_{i,j}|.$$

b) Zeigen Sie: Die *Summennorm induziert die Spaltensummen-Norm*, also

$$\|B\|_1 \; = \; \max_{j=1,\ldots,n} \sum_{i=1}^{m} |b_{i,j}|.$$

c) Sei $A \in \mathbb{R}^{n\times n}$, $n \in \mathbb{N}$ eine invertierbare Matrix und $\|\cdot\|_p$ eine durch eine Vektornorm induzierte Matrixnorm. Dann bezeichnet

$$\kappa_{\|\cdot\|_p}(A) := \|A\|_p\,\|A^{-1}\|_p$$

die *Konditionszahl* der Matrix $A$.

Seien nun eine invertierbare Matrix $A \in \mathbb{R}^{n\times n}$ und zwei Vektoren $v_1 \in \mathbb{R}^n$, $v_2 \in \mathbb{R}^n$ gegeben mit: $\|v_1\|_1 = 4$, $\|A\,v_1\|_1 = 20$, $\|v_2\|_1 = 3$ und $\|A\,v_2\|_1 = 1$.

Geben Sie eine größtmögliche untere Schranke $c$ für $\kappa_{\|\cdot\|_1}(A)$ an.

**Lösungsvorschlag:**

a)

$$\begin{aligned}
||B||_\infty &= \max_{||x||_\infty=1} ||Bx||_\infty = \max_{||x||_\infty=1} \left\{ \max_{i=1,\ldots,m} \left| \sum_{j=1}^{n} b_{ij} x_j \right| \right\} \\
&= \max_{i=1,\ldots,m} \left\{ \max_{||x||_\infty=1} \left| \sum_{j=1}^{n} b_{ij} x_j \right| \right\} \overset{(*)}{=} \max_{i=1,\ldots,m} \sum_{j=1}^{n} |b_{ij}|
\end{aligned}$$

($*$): Die Summe nimmt hier genau dann das Maximum an, wenn $x_j = sign(b_{ij})$ gilt. Dann gilt:

$$|b_{ij}| = b_{ij}x_j > b_{ij}\hat{x}_j \quad \forall \hat{x}_j \in (-1,1) \quad \forall j \in 1,...,n$$

b)

$$\begin{aligned}||B||_1 &= \max_{||x||_1=1} ||Bx||_1 = \max_{||x||_1=1} \sum_{i=1}^{n} \left| \sum_{j=1}^{m} b_{ij}x_j \right| = \max_{||x||_1=1} \sum_{i=1}^{n} \sum_{j=1}^{m} |b_{ij}||x_j| sign(b_{ij}x_j) \\ &= \max_{||x||_1=1} \sum_{j=1}^{m} |x_j| \sum_{i=1}^{n} |b_{ij}| sign(b_{ij}x_j) = \max_{||x||_1=1} \sum_{j=1}^{m} |x_j| \sum_{i=1}^{n} |b_{ij}| \overset{(**)}{=} \max_{j=1,...,n} \sum_{i=1}^{m} |b_{ij}|\end{aligned}$$

($**$): Die hintere Summe wird genau dann maximal, wenn $x_j = 1$ und $x_k = 0 \forall k \neq j$ für $b_{ij} \geq b_{ik} \forall k \neq j$.

c) Es gilt:

$$||A||_1||v||_1 \geq ||Av||_1 \quad \Longleftrightarrow \quad ||A||_1 \geq \frac{||Av||_1}{||v||_1} \quad \forall v \neq 0$$

$$\Rightarrow ||A||_1 \geq \max\left(\frac{||Av_1||_1}{||v_1||_1}, \frac{||Av_2||_1}{||v_2||_1}\right) = 5$$

Außerdem gilt:

$$b_1 := Av_1 \quad \text{und} \quad b_2 := Av_2 \quad \Longleftrightarrow \quad v_1 = A^{-1}b_1 \quad \text{und} \quad v_2 = A^{-1}b_2$$

$$\Rightarrow ||A^{-1}||_1 \geq \max\left(\frac{||A^{-1}b_1||_1}{||b_1||_1}, \frac{||A^{-1}b_2||_1}{||b_2||_1}\right) = 3$$

$$\Rightarrow \kappa_{||\cdot||_1}(A) = ||A||_1||A^{-1}||_1 \geq 5 \cdot 3 = 15$$

**Aufgabe 6** *(Programmier-Aufgabe: Matrixnormen)* (8M+4M Punkte)

a) Schreiben Sie eine Funktion `norm = mynorm(A, flag)`, welche eine Matrix $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$ und einen String `flag` als Parameter bekommt und je nach Wert von `flag` eine bestimmte Norm von $A$ zurückliefert. Ihre Funktion sollte folgenden Werten von `flag` die nachfolgenden Normen zuordnen:

| `flag` | Math. Bez. | Berechnung |
|---|---|---|
| `'one'` | $\|\cdot\|_1$ | $\|A\| = \max_{j=1,...,n} \left(\sum_{i=1}^m \|a_{i,j}\|\right)$ |
| `'infty'` | $\|\cdot\|_\infty$ | $\|A\| = \max_{i=1,...,m} \left(\sum_{j=1}^n \|a_{i,j}\|\right)$ |
| `'frobenius'` | $\|\cdot\|_F$ | $\|A\| = \left(\sum_{i=1}^m \sum_{j=1}^n \|a_{i,j}\|^2\right)^{1/2}$ |

**Hinweise:** Es ist möglich die Teilaufgabe a) ohne Schleife zu lösen. Zum Vergleich zweier Strings eignet sich der Befehl `strcmp`.

**Die Verwendung der Matlab-Funktion `norm()` ist für diese Teilaufgabe nicht zulässig.**

b) Testen Sie Ihre in a) implementierte Funktion. Erstellen Sie dazu mehrere zufällige Matrizen unterschiedlicher Dimensionen. Hierzu können Sie den Matlab-Befehl `rand` verwenden. Transformieren Sie die mit `rand` erhaltenen Werte auf verschiedene Wertebereiche.

Vergleichen Sie die Ergebnisse der von Ihnen implementierten Funktion mit den Werten, die Sie von der Matlab-Funktion `norm()` erhalten.

**Lösungsvorschlag:**

a) Siehe Moodle-Kurs im Abschnitt „Lösungsvorschläge“ unter „Blatt 02 Aufgabe 6 Matlab-Lösung“.

b) Siehe Moodle-Kurs im Abschnitt „Lösungsvorschläge“ unter „Blatt 02 Aufgabe 6 Matlab-Lösung“.

**Hinweise:**
Die Lösungen der Theorieaufgaben und Ihren Text zu den Programmieraufgaben können Sie in LaTeX erstellen oder handschriftlich aufschreiben und einscannen. Die Programmieraufgaben sind in Matlab zu lösen. Der Source Code muss strukturiert und dokumentiert sein.
Falls Sie die Aufgaben im Team lösen, geben Sie bitte auf allen Lösungen alle an der Aufgabe beteiligten Teammitglieder an. Jedes Teammitglied muss eine Lösung abgeben.
Speichern Sie Ihre Lösungen und Ihre Ergebnisse in einem Directory mit dem Namen **Blatt02_*Vorname*_*Nachname*** und verpacken Sie dieses in eine `.zip`-Datei. Laden Sie **spätestens 48 Stunden vor Ihrem Tutorium** diese `.zip`-Datei in Moodle hoch.