Prof. Dr. Karsten Urban
Dr. Timo Tonn
Dipl.-Math. oec. Klaus Stolle
Institut für Numerische Mathematik
Universität Ulm

WiSe 2023/2024
29.10.2023

Angewandte Numerik

Übungsblatt 02

Besprechung in den Tutorien von Montag, 06.11.2023 bis Freitag, 10.11.2023

Für dieses Übungsblatt gibt es 27 Theorie- und 23 Matlab-Punkte.
Die 70-Prozent-Grenzen liegen aktuell (inklusive Blatt 02) bei 34,3 Theorie- und 25,2 Matlab-Punkten.

Aufgabe 4 *(Programmieraufgabe: Unterschied zwischen Kondition und Stabilität)*
(3T+2T+2T+3T+2T+1T+3M+4M+4T Punkte)

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

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

berechnet werden.

a) Betrachten wir zunächst nur die Berechnung der größeren Nullstelle x_2.

   i) Berechnen Sie die Konditionszahlen κ_p und κ_q. Überlegen Sie sich dazu zunächst, wie Sie die in die Konditionszahlen eingehende Funktion φ definieren müssen.

   ii) Für welche p und welche q ist das Problem der Berechnung der größeren der beiden Nullstellen der quadratischen Gleichung gut konditioniert? Für welche p und welche q ist das Problem schlecht konditioniert? Warum?

   iii) Geben Sie ein konkretes gut konditioniertes und ein konkretes schlecht konditioniertes Beispiel an. Welche Werte haben die Konditionszahlen κ_p und κ_q für diese Beispiele?

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

b) Betrachten wir nun den Fall $q > 0$. Begründen Sie, dass die quadratische Gleichung zwei reelle Lösungen hat und dass man bei Verwendung der üblichen Formel $x_{1/2} = -p \pm \sqrt{p^2 + q}$ auf Auslöschung achten muss. Unter welchen Bedingungen tritt Auslöschung auf?

   Hinweis: Zum Begriff der *Auslöschung* siehe Folien 18 und 22 im Foliensatz „Angewandte Numerik, Vorlesung 1, Teil 2“.

c) Zeigen Sie: $x_1 x_2 = -q$ und $x_1 + x_2 = -2p$.

d) Schreiben Sie eine MATLAB-Funktion `[x1, x2] = nst(p, q)`, die x_1 und x_2 mit der üblichen Formel aus Aufgabenteil b) berechnet, sowie eine MATLAB-Funktion `[x1, x2] = nstStabil(p, q)`, die x_1 und x_2 mit einem numerisch stabilen Algorithmus berechnet.

e) Schreiben Sie ein MATLAB-Skript, das die Nullstellen $x_{1/2}$ mit Ihren beiden MATLAB-Funktionen aus Aufgabenteil d) für $q = 1$ und viele verschiedene Werte p berechnet. Wählen Sie dazu $p = 10^t$, wobei $t = 0, 0.1, 0.2, \ldots, 12$. Berechnen Sie die relativen Fehler in x_1 und x_2 und plotten Sie im von Auslöschung betroffenen Fall die relativen Fehler doppelt logarithmisch über p (MATLAB-Befehl `loglog`).

   Führen Sie Ihre Berechnungen auch für negative Werte p (also $p = -10^t$) durch und plotten Sie wieder die relativen Fehler des von Auslöschung betroffenen Falls doppelt logarithmisch (jetzt über $|p|$).

   Interpretieren Sie Ihre Schaubilder.

f) Erklären Sie den Begriff der „Stabilität“ und grenzen Sie ihn vom Begriff der „Kondition“ ab. Was bedeutet „gut konditioniert“ und was „numerisch stabil“?

Aufgabe 5 *(Aufwand und $\mathcal{O}$-Notation)* (2T+2T+2T+2T+2T Punkte)

Berechnen Sie den Aufwand für folgende Operationen ($n \in \mathbb{N}$). Geben Sie dabei sowohl die genaue Anzahl der Gleitkommaoperationen als auch den Aufwand in der $\mathcal{O}$-Notation an.

a) Vektor-Vektor-Produkt (Euklidisches Skalarprodukt): $x^T y$, wobei $x, y \in \mathbb{R}^n$,

b) Matrix-Vektor-Produkt: $A \cdot x$, wobei $A \in \mathbb{R}^{n\times n}$ und $x \in \mathbb{R}^n$,

c) Matrix-Vektor-Produkt: $L \cdot x$, wobei $L \in \mathbb{R}^{n\times n}$ eine linke untere Dreiecksmatrix und $x \in \mathbb{R}^n$,

d) Matrix-Matrix-Produkt: $A \cdot B$, wobei $A, B \in \mathbb{R}^{n\times n}$ und

e) Lösen eines gestaffelten linearen Gleichungssystems $L\,x = b$, wobei $L \in \mathbb{R}^{n\times n}$ eine unipotente linke untere Dreiecksmatrix und $b \in \mathbb{R}^n$ sind.

Aufgabe 6 *(Rechenzeiten Vorwärtseinsetzen)* (3M+1M+6M+2M+2M+2M Punkte)

In der Vorlesung haben Sie zwei MATLAB-Funktionen `x = Rinvb(R,z)` und `x = Rinvb2(R,z)` zur Lösung eines gestaffelten linearen Gleichungssystems $R\,x = z$ gesehen, wobei $R \in \mathbb{R}^{n\times n}$ ($n \in \mathbb{N}$) eine reguläre rechte obere Dreiecksmatrix und $z \in \mathbb{R}^n$ ein Vektor sind.

a) Übertragen Sie die Ideen dieser beiden MATLAB-Funktionen auf die Lösung eines gestaffelten linearen Gleichungssystems $L\,x = z$, wobei $L \in \mathbb{R}^{n\times n}$ ($n \in \mathbb{N}$) eine reguläre linke untere Dreiecksmatrix und $z \in \mathbb{R}^n$ ein Vektor sind.

   Schreiben Sie zwei entsprechende MATLAB-Funktionen `x = Linvb(L,z)` und `x = Linvb2(L,z)`.

b) Testen Sie Ihre MATLAB-Funktionen an geeigneten Beispielen.

c) Nun wollen wir die Laufzeit Ihrer MATLAB-Funktionen in Abhängigkeit von der Dimension der Matrix `L` untersuchen. Schreiben Sie dazu ein MATLAB-Skript, das eine reguläre linke untere Dreiecksmatrix `L` und einen Vektor `z` erzeugt, anschließend das gestaffelte lineare Gleichungssystem $L\,x = z$ mit Ihren MATLAB-Funktionen `Linvb` und `Linvb2` löst und jeweils die dafür benötigte Zeit misst. Die Dimension n der Matrix soll dabei die Werte $2^6, 2^7, \ldots, 2^{16}$ durchlaufen.

   Plotten Sie die benötigten Zeiten über der Dimension der Matrix. Mit welchem der MATLAB-Plot-Befehle erhalten Sie ein aussagekräftiges Schaubild?

Passen Ihre numerischen Ergebnisse zur Aussage des Satzes über den Rechenaufwand für Dreiecksmatrizen (Folie 21)? Wie können Sie das überprüfen?

Hinweis: Die MATLAB-Anweisungen `tic` und `toc` könnten bei der Zeitmessung hilfreich sein.

Achtung: Je nach verwendeter Hardware ist die Laufzeit Ihres MATLAB-Skripts möglicherweise sehr lange. Verwenden Sie also in der Entwicklungs- und Testphase eine kleinere obere Grenze für die Dimension der Matrix. Die vorgegebene obere Grenze von 2^{16} dient nur als Richtlinie. Sie dürfen sie für das abzugebende Schaubild (nach unten oder gerne auch nach oben) an Ihre Hardware anpassen.

d) Vektorisieren Sie nun Ihre beiden MATLAB-Funktionen `Linvb` und `Linvb2`.

   Schreiben Sie dazu zwei MATLAB-Funktionen `x = Linvb_Vec(L,z)` und `x = Linvb2_Vec(L,z)`, in denen Sie jeweils die innere For-Schleife mit einer Anweisung, die auf ein einzelnes Vektor- und ein einzelnes Matrixelement zugreift, durch eine Anweisung ersetzen, die auf mehrere Elemente eines Vektors und mehrere Elemente der Matrix gleichzeitig zugreift.

   Testen Sie auch diese beiden MATLAB-Funktionen.

e) Integrieren Sie Ihre beiden MATLAB-Funktionen aus Aufgabenteil d) in Ihr Skript aus Aufgabenteil c). Plotten Sie also auch die von Ihren MATLAB-Funktionen `Linvb_Vec` und `Linvb2_Vec` benötigten Zeiten über der Dimension der Matrix.

f) Unterscheiden sich die Laufzeiten Ihrer vier MATLAB-Funktionen? Falls ja: Erklären Sie die Unterschiede.

Hinweise:

Teamarbeit ist bei der Bearbeitung der Übungsblätter erlaubt und erwünscht. Falls Sie die Aufgaben im Team lösen, muss jedes Teammitglied eine eigene Lösung votieren und abgeben. Geben Sie dann bitte auf Ihrer Lösung alle an der Aufgabe beteiligten Teammitglieder an.

Die Programmieraufgaben sind in MATLAB zu lösen. Der Source Code muss strukturiert und dokumentiert sein.

Speichern Sie Ihre Lösungen der Theorieaufgaben, Ihre MATLAB-Programme und Ihre Ergebnisse in einem Directory mit dem Namen **Blatt02_*Vorname_Nachname*** und verpacken Sie dieses in eine `.zip`-Datei. **Spätestens bis Sonntag, 05.11.2023, 23:59 Uhr** müssen Sie diese `.zip`-Datei in Moodle hochladen und für die Aufgabenteile, die Sie vorstellen können, votieren.

Jonas Kruger

Aufgabe 4

$$q^2 + p > 0 \qquad \varphi: \mathbb{R}^2 \to \mathbb{R}, \ (p,q) \mapsto -p + \sqrt{p^2+q}$$

a) i) $$k_p = \frac{|x_p|}{|\varphi(p,q)|} \cdot \left|\frac{\partial \varphi(p,q)}{\partial p}\right| = \frac{|p|}{|-p+\sqrt{p^2+q}|} \left| -1 + \frac{\cancel{2}p}{\cancel{2}\sqrt{p^2+q}} \right|$$

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

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

$$k_q = \frac{|x_q|}{|\varphi(p,q)|} \left|\frac{\partial \varphi(p,q)}{\partial q}\right| = \left| \frac{q}{-p+\sqrt{p^2+q}} \right| \left| \frac{1}{2\sqrt{p^2+q}} \right| = \left| \frac{q(p+\sqrt{p^2+q})}{\cancel{-p^2+p^2}+q} \cdot \frac{1}{2\sqrt{p^2+q}} \right|$$

$$= \left| \frac{\cancel{q}(p+\sqrt{p^2+q})}{\cancel{q}} \cdot \frac{1}{2\sqrt{p^2+q}} \right| = \boxed{\frac{|p+\sqrt{p^2+q}|}{2\sqrt{p^2+q}}}$$

ii)

k_p und k_q sind für $p \in \mathbb{R}$ und $q \geq 0$ gut konditioniert, da

$k_p: \ |p| \leq \sqrt{p^2+q} \qquad k_q: \ |p\sqrt{p^2+q}\,| \leq 2\sqrt{p^2+q}$.

Für $q < 0$ wird das Problem schlecht konditioniert, da

$k_p: \ |p| > \sqrt{p^2-q} \qquad k_q: \ |p\sqrt{p^2-q}\,| > 2\sqrt{p^2-q}$

iii) gut konditioniert $\qquad x^2 + 4x - 5 = 0 \quad \rightarrow \quad p = 2, \ q = 5$

$$k_p = \frac{2}{\sqrt{2^2+5}} = \frac{2}{\sqrt{9}} = \frac{2}{3} \qquad k_q = \frac{|2+\sqrt{2^2+5}\,|}{2\sqrt{2^2+5}} = \frac{2+3}{2\cdot 3} = \frac{5}{6}$$

$p^2+q > 0 \quad p^2 > -q \quad q > -p^2 \quad p > 0$

schlecht konditioniert : $\quad x^2 + 8x + 15{,}5 = 0$

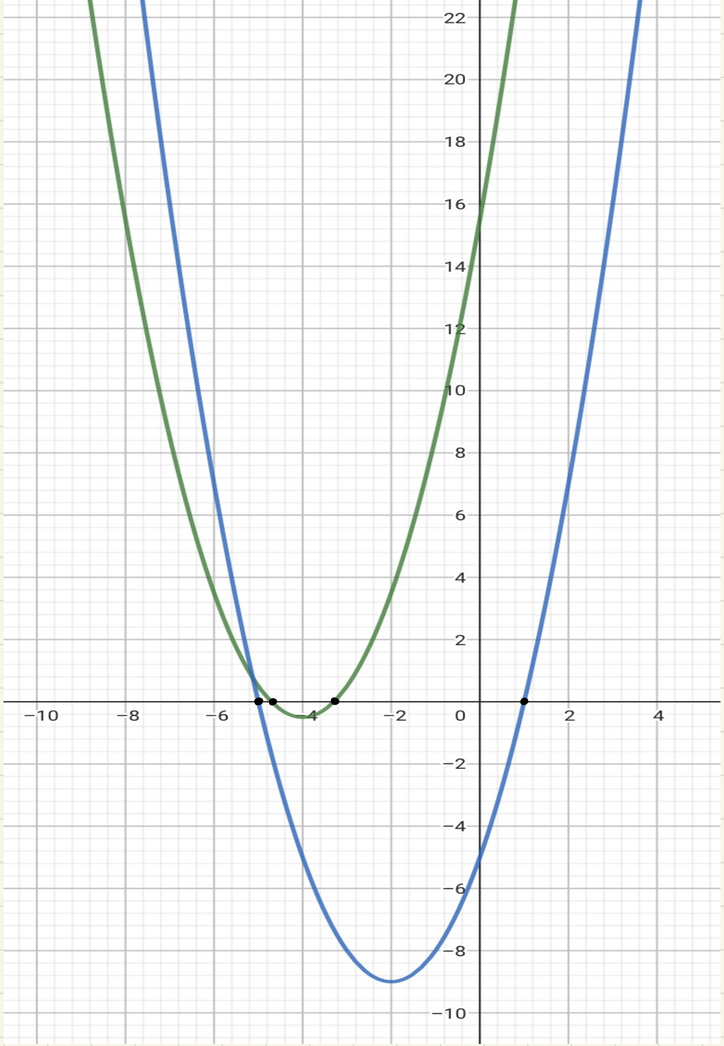
$$k_p = \frac{4}{\sqrt{4^2-15{,}5}} = \frac{4}{\sqrt{16-15{,}5}} = \frac{4}{\sqrt{0{,}5}} = 16$$

$$k_q = \frac{4+\sqrt{4^2-15{,}5}}{2\sqrt{4^2-15{,}5}} = \frac{4+1/4}{1/4} = 16+1 = 17$$

iv)

$x^2 + 8x + 15{,}5 = 0$

$x^2 + 4x - 5 = 0$



b) $$x_{1,2} = p \pm \sqrt{p^2 + q} \qquad p > 0$$

Diese Gleichung besitzt zwei reelle Lösungen, da nach der Bedingung: $p^2 + q > 0$ Wurzel immer zwei reelle Lösungen liefert. p ist auch eine reelle Zahl, wodurch beide Lösungen der Gleichung auch reelle Zahlen darstellen.

Auslöschung der Wurzel tritt im Fall $p^2 \approx -q$ auf.

c) Zeigen Sie: $x_1 x_2 = -q$ und $x_1 + x_2 = -2p$.

$$x_1 = -p - \sqrt{p^2 + q} \qquad x_2 = -p + \sqrt{p^2 + q}$$

$x_1 x_2 = -q$:

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

$x_1 + x_2 = -2p$

$$-p - \cancel{\sqrt{p^2+q}} - p + \cancel{\sqrt{p^2+q}} = -2p$$

e) Aufgaben in .zip

f) Erklären Sie den Begriff der „Stabilität“ und grenzen Sie ihn vom Begriff der „Kondition“ ab. Was bedeutet „gut konditioniert“ und was „numerisch stabil“?

Stabilität ist eine Eigenschaft eines Algorithmus und gibt an wie groß der relative

Stabilität ist eine Eigenschaft eines Algorithmus und gibt an wie groß der relative Fehler der Ausgabe von der Berechnung ist. Von einem Algorithmus mit guter Stabilität wird erwartet, dass die relativen Ausgabefehler nicht deutlich mehr verstärkt werden als die Konditionszahlen des Problems angeben.

"gut konditioniert" bedeutet, dass der relative Ausgabefehlers eines Problems nicht stark verstärkt wird.
"numerisch stabil" bedeutet, dass der Algorithmus zur Berechnung des Problems nicht Eingabefehler über die Konditionszahlen deutlich verstärkt.

Ein Problem kann gut konditioniert sein, aber mit einem numerisch instabilen Algorithmus immer noch zu größeren Rechenfehler führen.

5.

a) a) Vektor-Vektor-Produkt (Euklidisches Skalarprodukt): $x^T y$, wobei $x, y \in \mathbb{R}^n$,

$$x^T y = (x_1, x_2, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

n- Multiplikationen + (n-1) Additionen -> FLOP = $n + (n-1) = 2n-1 \in O(n)$

b) b) Matrix-Vektor-Produkt: $A \cdot x$, wobei $A \in \mathbb{R}^{n \times n}$ und $x \in \mathbb{R}^n$,

$$A \cdot x = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} x_1 + \dots + a_{1n} x_n \\ a_{21} x_1 + \dots + a_{2n} x_n \\ \vdots \\ a_{n1} x_1 + \dots + a_{nn} x_n \end{pmatrix}$$

n mal n- Multiplikationen + n-1 Additionen

-> FLOP = $n(n+n-1) = n(2n-1) = 2n^2 - n \in O(n^2)$

c)

c) Matrix-Vektor-Produkt: $L \cdot x$, wobei $L \in \mathbb{R}^{n \times n}$ eine linke untere Dreiecksmatrix und $x \in \mathbb{R}^n$

$$L \cdot x = \begin{pmatrix} a_{11} & & & 0 \\ \vdots & a_{22} & & \\ \vdots & & \ddots & \\ a_{1n} & \cdots & & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{array}{l} \begin{pmatrix} a_{11} x_1 \\ a_{12} x_1 + a_{22} x_2 \\ \vdots \\ a_{1n} x_1 + \dots + a_{nn} x_n \end{pmatrix} \end{array} \begin{array}{l} 1 \\ 3 \\ \vdots \\ 2n-1 \end{array}$$

$$\text{FLOP} = \sum_{k=1}^{n} (2k-1) = 2\sum_{k=1}^{n} k - \sum_{k=1}^{n} 1 = 2 \cdot \frac{n(n+1)}{2} - n = n^2 + n - n = n^2 \in O(n^2)$$

d) Matrix-Matrix-Produkt: $A \cdot B$, wobei $A, B \in \mathbb{R}^{n \times n}$

$$A \cdot B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + \dots + a_{1n}b_{n1} & a_{11}b_{12} + \dots + a_{1n}b_{n2} & \dots & a_{11}b_{1n} + \dots + a_{1n}b_{nn} \\ \vdots & & & \\ a_{n1}b_{11} + \dots + a_{nn}b_{n1} & & \cdots & a_{n1}b_{1n} + \dots + a_{nn}b_{nn} \end{pmatrix}$$

n-mal Matrix-mal-Vektor $\rightarrow \text{FLOP} = n(2n^2 - n) = 2n^3 - n^2 \in O(n^3)$

e) Lösen eines gestaffelten linearen Gleichungssystems $Lx = b$, wobei $L \in \mathbb{R}^{n\times n}$ eine unipotente linke untere Dreiecksmatrix und $b \in \mathbb{R}^n$ sind.

$$L x = b \qquad \begin{pmatrix} a_{11} & & 0 \\ \vdots & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1/a_{11} \\ (b_2 - (x_1 a_{21}))/a_{22} \\ \vdots \\ \end{pmatrix} \begin{array}{c} 1 \\ 3 \\ \vdots \\ 2n-1 \end{array}$$

$$\rightarrow \text{FLOP} = 1 + 3 + \ldots + (2n-1) = \sum_{k=1}^{n} (2k-1) = 2\sum_{k=1}^{n} k - \sum_{k=1}^{n} 1 = 2\,\frac{n(n+1)}{2} - n$$

$$= n^2 + n - n = n^2 \in O(n^2)$$

Aufgabe 6

c) Aussage passt für linvb, da quadr. Kurve. Bei Linvb2 fällt diese Kurve deutlich flacher aus. Mithilfe des plot-Befehls lässt sich das Ergebnis interpretieren.

f) erklärt in aufgabe6c.m