Projekt 4: Rückwärtsstabilität der Cholesky-Zerlegung

1. Wir betrachten eine Rundungsfehleranalyse für die Cholesky-Zerlegung auf einem Computer mit Rechengenauigkeit `eps`. Zur Vereinfachung unserer Überlegungen nehmen wir an, daß die Grundrechenoperationen sowie das Wurzelziehen folgendem Modell genügen: Seien $\{+^*, -^*, *^*, /^*, \sqrt{}^*\}$ die Realisierungen der 4 Grundrechenarten und des Wurzelziehens auf dem Computer. Für zwei Gleitkommazahlen x, y gibt es ein $\delta = \delta(x, y)$ mit $|\delta| \leq \texttt{eps}$, so daß

$$x \texttt{op}^* y = (x \,\texttt{op}\, y)(1 + \delta) \qquad \texttt{op} \in \{+, -, *, /, \sqrt{}\}.$$

(Dies besagt, daß die Grundrechenarten und das Wurzelziehen rückwärtsstabil realisiert werden.)

a) Die Zahlen $\delta_i \in \mathbb{R}$, $i = 1, \ldots, n$ mögen $|\delta_i| \leq \texttt{eps}$ erfüllen und $\rho_i \in \{-1, 1\}$. Zeigen Sie: Es gibt ein $\theta_n \in \mathbb{R}$ mit

$$|\theta_n| \leq \gamma_n := \frac{n\texttt{eps}}{1 - n\texttt{eps}}$$

so daß

$$\prod_{i=1}^{n}(1 + \delta_i)^{\rho_i} = 1 + \theta_n.$$

Hier (und im Folgenden) ist die implizite Annahme, daß $n\texttt{eps} < \texttt{1}$ ist.

b) Seien $(x_i)_{i=1}^n$ und $(y_i)_{i=1}^n$ zwei Folgen von Gleitkommazahlen. Wir bezeichnen mit $\texttt{skalar}(x, y)$ die Funktion, die das Skalarprodukt $x^\top y$ durch $\texttt{skalar}(x, y) := s_n$ mit

$$s_1 := x_1 *^* y_1, \quad s_{i+1} := s_i +^* (x_{i+1} *^* y_{i+1}), \qquad i = 1, \ldots, n - 1$$

bestimmt. Zeigen Sie folgendes Rückwärtsstabilitätsresultat: Das erhaltene Ergebnis $s_n = \texttt{skalar}(x, y)$ erfüllt für geeignet gewählte θ_n, θ'_n, $\theta_{n-1}, \ldots, \theta_2$:

$$s_n = x_1 y_1(1 + \theta_n) + x_2 y_2(1 + \theta'_n) + x_3 y_3(1 + \theta_{n-1}) + \cdots + x_n y_n(1 + \theta_2);$$

hier erfüllen alle θ_i, θ'_i die Abschätzungen $|\theta_i| \leq \gamma_i$ und $|\theta'_i| \leq \gamma_i$.

c) Schließen Sie auf folgendes Vorwärtsstabilitätsresultat ($|x|$ und $|y|$ sind Vektoren, bei denen die Betragsfunktion komponentenweise angewendet wurde):

$$|x^\top y - \texttt{skalar}(x, y)| \leq \gamma_n |x|^\top |y|$$

2. Für symmetrische, positive definite Matrizen $A \in \mathbb{R}^{n \times n}$ wird normalweise die *Cholesky-Zerlegung* berechnet, d.h. eine Matrix $L \in \mathcal{L}_n$, deren Diagonaleinträge positiv sind, so daß

$$LL^\top = A$$

a) Formulieren Sie einen Algorithmus (analog zur LU-Zerlegung nach Crout), der den Choleskyfaktor L bestimmt.

b) Überlegen Sie sich, daß der berechnete Faktor $\widehat{L}$ folgendes erfüllt:

$$\begin{aligned} \left| a_{ij} - \sum_{k=1}^{i} \widehat{l}_{ik}\widehat{l}_{jk} \right| &\leq \max\{\gamma_2, \gamma_{i-1}\} \sum_{k=1}^{i} |\widehat{l}_{ik}||\widehat{l}_{jk}|, \qquad i > j, \\ \left| a_{ii} - \sum_{k=1}^{i} \widehat{l}_{ik}^2 \right| &\leq \max\{\gamma_2, \gamma_{i-1}\} \sum_{k=1}^{i} |\widehat{l}_{ik}|^2 \qquad \forall i. \end{aligned}$$

Schließen Sie auf folgendes Rückwärtsstabilitätsresultat: Falls die Cholesky-Zerlegung einer symmetrischen, positiv definiten Matrix A nicht abbricht, dann ist der erhaltene Choleskyfaktoren $\widehat{L}$ der exakte Choleskyfaktor einer Matrix $A + \Delta A$, wobei $\Delta A \in \mathbb{R}^{n\times n}$ folgendes erfüllt:

$$|\Delta A| \leq \gamma_n |\widehat{L}|\,|\widehat{L}^\top|.$$

Hier agiert wieder die Betragsfunktion in der Definition von $|\widehat{L}|$ und $|\widehat{L}^\top|$ komponentenweise; die Ungleichung $\leq$ ist ebenfalls komponentenweise zu verstehen.

c) Zeigen Sie folgende Variante des obigen Stabilitätsresultats:

$$|\Delta A| \leq \frac{\gamma_n}{1-\gamma_n} dd^\top, \qquad d = (a_{11}, a_{22}, a_{33}, \ldots, a_{nn})^\top.$$

Hinweis: Seien $\widehat{l_i}$, $i = 1, \ldots, n$, die *Zeilen* von $\widehat{L}$. Betrachten Sie zuerst die Diagonalelemente von $A + \Delta A$, und zeigen Sie eine Abschätzung für die $\|\widehat{l_i}\|_2$, $i = 1, \ldots, n$.

3. Ziel ist zu sehen, daß das obige Rückwärtsstabilitätsresultat qualitativ richtig ist. Hierzu wird in Matlab die Cholesky-Zerlegung (`help chol`) in *einfacher Genauigkeit* durchgeführt (*Hinweis:* `help single`, `help eps`, `help cast`) und der Fehler $A - LL'$ in doppelter Genauigkeit bestimmt. Die Matrizen $A \in \mathbb{R}^{n\times n}$ in doppelter Genauigkeit seien mittels $\mathtt{L} = \mathtt{tril}(\mathtt{hilb}(\mathtt{n}))$ und $A = \mathtt{L} * \mathtt{L}'$ definiert.

a) Schreiben Sie ein Matlabprogramm, welches A bestimmt, anschließend mittels `single` auf einfache Genauigkeit konvertiert und dann die Cholesky -Zerlegung (ebenfalls in einfacher Genauigkeit) bestimmt. Konvertieren Sie anschließend den erhaltenen Choleskyfaktor L in "Double" und berechnen Sie den Fehler $A - L * L'$ in doppelter Genauigkeit (hierzu müssen Sie natürlich A in den Typ "double" umwandeln). Plotten Sie semilogarithmisch für $n = 1, \ldots, 100$ die Schranke $\gamma_n \mathtt{norm}(|\widehat{L}|\,|\widehat{L}'|)$ sowie den Fehler $\mathtt{norm}(A - L * L')$. Diskutieren Sie Ihre Ergebnisse.