Institut für Analysis und Numerik 01.05.2012
Universität Magdeburg
Thomas Richter thomas.richter@ovgu.de
Henry von Wahl henry.vonwahl@ovgu.de

Blatt 3 zur Vorlesung Numerische Mathematik
Sommersemester 2020

Abgabe bis 8. Mai, 23:59

Thema der nächsten 2 Wochen sind Methoden zur Berechnung von Eigenwerten einer Matrix $A \in \mathbb{R}^{n\times n}$. Zunächst geht es um die generelle Einordnung des Problems und seine Konditionierung. Im Anschluss untersuchen wir Methoden zur Approximation von Eigenwerten. Zunächst handelt es sich um Verfahren, die einzelne Eigenwerte berechnen, im Anschluss betrachten wir Zerlegungsmethoden, mit denen alle Eigenwerte einer Matrix gleichzeitig berechnet werden können.

Das Eigenwertproblem Man lese Kapitel 5, *Berechnung von Eigenwerten* bis einschließlich Abschnitt *5.2 Direkte Methode*. Satz 5.9 und sein Beweis ist zentral. Man beantworte kurz die folgenden Fragen:

1. Was ist der Rayleigh-Quotient?

Zur Matrix $A \in \mathbb{R}^{n\times n}$ und Vektor $w \in \mathbb{R}^n$ ist durch

$$\lambda(w) = \frac{(Aw, w)}{\|w\|^2}$$

der Rayleigh-Quotient gegeben. Ist w ein Eigenvektor, so ist $\lambda(w)$ zugehöriger Eigenwert. Wird der Rayleigh-Quotient im Rahmen der Potenzmethode oder der inversen Iteration verwendet und ist die Matrix A symmetrisch, so halbiert sich die Konvergenzrate.

2. Wovon hängt die Konditionierung des Eigenwertproblems ab?

Die Eigenwerte hängen von der Kondition derjenigen Matrix ab, die sich als Basis aus Eigenvektoren ergibt. Das Problem kann beliebig schlecht konditioniert sein. Existiert aber eine orthonormalsystem aus Eigenvektoren, so ist die Konditionierung sehr gut (Konditionszahl 1).

3. Man gebe mit den Gerschgorin-Kreisen eine möglichst scharfe Abschätzung für die Eigenwerte der folgenden Matrix an

$$A = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 8 & 7 \\ 4 & 2 & -2 \end{pmatrix}$$

Bei zeilenweiser Berechnung ergeben sich die Gebiete ($B_r(m)$ ist Kreis von Radius r um Punkt m)

$$E_1 = B_1(4), \quad E_2 = B_7(8), \quad E_3 = B_6(-2).$$

Da die Gebiete sich alle überlappen, liegen die drei Eigenwerte in

$$E_1 \cup E_3$$

Spaltenweise ergibt sich

$$F_1 = B_4(4), \quad F_2 = B_2(8), \quad F_3 = B_8(-2).$$

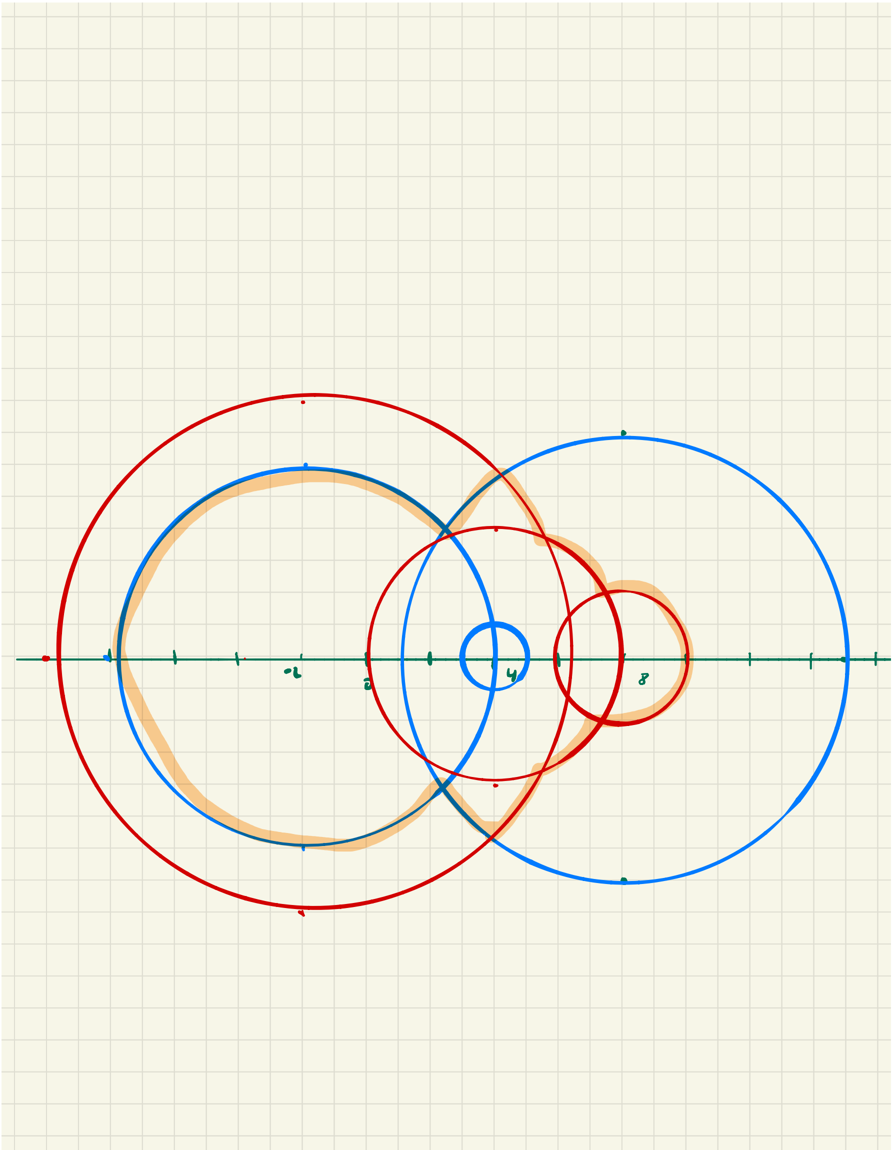
Auch hier überlappen sich alle Kreise und die Eigenwerte liegen in

$$F_1 \cup F_2 \cup F_3$$

Insgesamt können wir also nur sagen, dass die Eigenwerte in liegen.

$$(E_1 \cup E_3) \cap (F_1 \cup F_2 \cup F_3)$$

also im Schitt der roten und blauen Vereinigungsmengen:



4. Warum ist es nicht sinnvoll, Eigenwerte numerisch über das charakteristische Polynom zu berechnen, obwohl wir z.B. mit dem Newton-Verfahren eine sehr effiziente Methode zur Approximation einer Nullstelle kennen?

Das aufstellen des charakteristischen Polynoms erfordert sehr viele Additionen, bei denen es jeweils zur Auslöschung kommen kann. Das Nullstellenproblem ist bei Polynomen darüber hinaus schlecht konditioniert (wenn Nullstellen nahe beieinander liegen). Insgesamt ist dieser Weg daher sicher instabil.

Iterative Verfahren Man lese Abschnitt *5.3 Iterative Verfahren*. Zentrag ist Satz 5.11 über die Potenzmethode. Die Inverse Iteration ist dann nur noch eine Folgerung. Man beantworte die folgenden Fragen

1. Welche Eigenwerte können mit der Potenzmethode, welche mit der inversen Iteration (ohne Shift) berechnet werden?

   Potenzmethode: der betragsgrößte, wenn er separiert ist. Inverse Iteration der entsprechend (separierte) betragskleinste.

2. Warum ist es schlecht, wenn der Startvektor bei der Potenzmethode ein Eigenvektor ist?

   Falls $Ax^0 = \lambda x^0$ so bleiben wir immer im Eigenrau, d.H. es ist stets $Ax^k = \lambda x^k$ und nur λ kann identifiziert werden.

3. Was bedeutet *... mit nichttrivialer Komponente in Bezug auf* w_n in Satz 5.11?

   Das bedeutet, dass die Darstellung von x_0 in einer Basis aus Eigenvektoren eine Darstellung

$$x_0 = \sum_{i=1}^{n} \alpha_i w_i$$

   mit $\alpha_n \neq 0$ besitzt.

4. Wovon hängt die Konvergenzrate beider Verfahren ab?

   Jeweils vom "Abstand" von betragsgrößtem und zweitgrößtem Eigenwert, also $|\lambda_{n-1}|/|\lambda_n|$. Entsprechend die beiden kleinsten bei der inversen Iteration.

5. Was ist der wesentliche Aufwand bei der Durchführung der inversen Iteration?

   In jedem Schritt muss ein lineares Gleichungssystem gelöst werden, d.h. n^3. Falls kein Shift verwendet wird könnte einmal die LR-Zerlegung erstellt werden. Im Anspruch nur noch Vorwärts- und Rückwärtselimination, also n^2.

Übungsaufgaben

Aufgabe 3.1

Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix mit den Eigenwerten

$$|\lambda_1| \leqslant \cdots \leqslant |\lambda_{n-1}| < |\lambda_n|.$$

Man zeige, dass die Potenzmethode bei Verwendung des Rayleigh-Quotienten (vergleiche Bemerkung 5.12) quadratisch konvergiert.
Hierzu ist eine Modifikation der Abschätzung im Beweis zu Satz 5.11 notwendig.

Die Matrix A ist symmetrisch pos. def. D.h., es existiert ein orthogonalsystem von Eigenvektoren $(w_i, w_j) = \delta_{ij}$.
Es sei

$$x^{(0)} = \sum_k \alpha_k w_k.$$

Hiermit

$$x^{(i+1)} = \frac{Ax^{(i)}}{\|Ax^{(i)}\|} \frac{A\frac{Ax^{(i-1)}}{\|Ax^{(i-1)}\|}}{\|A\frac{Ax^{(i-1)}}{\|Ax^{(i-1)}\|}\|} = \frac{A^i x^{(0)}}{\|A^i x^{(0)}\|}.$$

Weiter:

$$A^i x^{(0)} = \sum_{k=1}^{n} \alpha_k \lambda_k^i w_k,$$

also:

$$\|A^i x^{(0)}\|_2^2 = \sum_{k,l=1}^{n} \alpha_k \alpha_l \lambda_k^i \lambda_l^i (w_k, w_l)_2 = \sum_{k=1}^{n} \alpha_k^2 \lambda_k^{2i}.$$

Dann gilt mit $\alpha = (\alpha_{n-1}/\alpha_n)^2$

$$\begin{aligned}
\lambda^{(i+1)} &= (Ax^{(i)}, x^{(i)})_2 = \frac{(Ax^{(i)}, x^{(i)})_2}{(x^{(i)}, x^{(i)})_2} \\
&= \frac{1}{\sum_{k=1}^{n} \alpha_k^2 \lambda_k^{2i}} \left(\sum_{k=1}^{n} \alpha_k \lambda_k^{i+1} w_k, \sum_{k=1}^{n} \alpha_k \lambda_k^{i} w_k \right) \\
&= \frac{\sum_{k=1}^{n} \alpha_k^2 \lambda_k^{2i+1}}{\sum_{k=1}^{n} \alpha_k^2 \lambda_k^{2i}} \\
&= \lambda_n \frac{1 + \sum_{k=2}^{n} \left(\frac{\alpha_k}{\alpha_n}\right)^2 \left(\frac{\lambda_k}{\lambda_n}\right)^{(2i+1)}}{1 + \sum_{k=2}^{n} \left(\frac{\alpha_k}{\alpha_n}\right)^2 \left(\frac{\lambda_k}{\lambda_n}\right)^{(2i)}} \\
&= \lambda_n \frac{1 + \alpha \left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{(2i+1)} + o(1)}{1 + \alpha \left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{2i} + o(1)}
\end{aligned}$$

$$= \lambda_n \left(1 + O\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^{2i} \right).$$

Der letzte Schritt folgt aus der Abschätzung

$$\frac{1 + \alpha x^{2i+1}}{1 + \alpha x^{2i}} = 1 + \alpha \frac{x^{2i+1} - x^{2i}}{1 + \alpha x^{2i}} = 1 + \alpha x^{2i} \frac{x - 1}{1 + \alpha x^{2i}} = 1 + O(|x|^{2i}).$$

Programmieraufgabe 3.2

a) Man implementiere die Potenzmethode. Dabei breche man die Iteration ab, wenn sich der approximierte Wert um weniger als 10^{-6} ändert, oder wenn 500 Schritte durchgeführt wurden.
Für die vorgegebenen Testmatrizen gebe man die Anzahl der benötigten Schritte und den gefundenen Eigenwert an.

Hinweis: Der Startvektor kann zufällig gewählt werden. Dies ist in der Vorlage beschrieben.

b) Man vergleiche die Konvergenz des Verfahrens bei Verwendung der Approximation

$$\lambda_n = \frac{\tilde{x}_k^{(i)}}{x_k^{(i-1)}}$$

sowie bei Verwendung des Rayleigh-Quotienten

$$\lambda_n = \frac{(Ax^{(i-1)}, x^{(i-1)})_2}{(x^{(i-1)}, x^{(i-1)})_2}.$$

Geben Sie für die Testmatrizen bei beiden Methoden die Anzahl der benötigten Schritte aus. Führen Sie jeweils mehrere Tests durch und geben Sie einen Durchschnittswert an. Die Ergebnisse variieren, da der Startvektor zufällig gewählt wird.
Was beobachten Sie im Fall des Rayleigh-Quotienten bei Anwendung auf eine nichtsymmetrische Matrix?

Abgabe per Mail an Henry von Wahl (henry.vonwahl@ovgu.de). Abgabe der Kurzfragen in 2er Gruppen, der Übungsaufgaben in den 5er Gruppen. Die theoretische Gruppenaufgabe ist freiwillig.