Aplikace matematiky

Tamás Frey

Einige neue Methoden zur numerischen Berechnung von Eigenwerten

Aplikace matematiky, Vol. 10 (1965), No. 3, 206--212

Persistent URL: `http://dml.cz/dmlcz/102950`

Terms of use:

© Institute of Mathematics AS CR, 1965

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* `http://dml.cz`

EINIGE NEUE METHODEN ZUR NUMERISCHEN BERECHNUNG VON EIGENWERTEN

T. FREY

(zum Thema d)

1. In jedem Rechenzentrum ist es eine der wichtigsten Aufgaben, praktisch gut anwendbare Methoden für Eigenwertaufgaben zu suchen. In dieser Arbeit werden einige neue Störungsmethoden und eine sog. sukzessive Methode betrachtet.

2. Die klassische Störungsmethode ist bei gestörten Eigenwertaufgaben der Form

$$G(\lambda, x) \,.\, x = [A + \varepsilon \,.\, B(x, \lambda)] \,.\, x = \lambda x \tag{1}$$

gut anwendbar, und zwar folgendermassen: Mit Hilfe der Ausgangswerte der Eigenelemente berechnet man zuerst die Matrix D, die genau die angegebenen Ausgangs-Eigenelemente besitzt, ferner auch G. Mit Hilfe der Differenzen wertet man die Korrektionen mit nicht zu grosser Genauigkeit aus und erhält so neue Ausgangswerte für die Eigenelemente. Unser Algorithmus ist dann wieder anwendbar.

Die theoretische Begründung der Konvergenzgeschwindigkeit fehlt jedoch noch, so dass zuerst dieser Mangel behoben werden muss.

Es seien also A und B zwei lineare Operatoren, die den Hilbertraum H in sich abbilden. Es sei λ_0 ein einfaches Element des Eigenspektrums von A, mit rechts- bzw. linksseitigem Eigenelement x_0 bzw. $y_0(\|y_0\| = 1;\ (y_0, x_0) = 1;\ \|x_0\| \geqq 1)$. Es bezeichne $H_\perp$ den zu y_0 orthogonalen Unterraum von H und R den in $H_\perp$ eindeutig definierten inversen Operator zu $A - \lambda_0 E$, der $H_\perp$ in sich abbildet. Es sei ferner $\|R\|_{H_\perp}$ bzw. $\|B - A\| = \|\Delta\|$ durch ϱ bzw. δ bezeichnet.

Satz 1. *Ist*

$$\varrho\delta\|x_0\| \leqq \tfrac{1}{8}\,, \tag{2}$$

so besitzt auch B einen Eigenwert λ bzw. ein Eigenelement x in der Nähe von λ_0 bzw. x_0, und die Reihe

$$\lambda_0 + \sum_{k=1}^{\infty} (\lambda_k - \lambda_{k-1});\ \lambda_k - \lambda_{k-1} = (y_0, \Delta x_{k-1})\,, \tag{3}$$

bzw.

$$(4)\qquad \sum_{k=0}^{\infty} x_k;\; x_k = R\{(B - A)\, x_{k-1} - (\lambda_k - \lambda_{k-1}) \sum_{n=0}^{k-1} x_n - (\lambda_k - \lambda_0)\, x_{k-1}\}$$

strebt exponentiell mit dem Quotienten

$$(5)\qquad q = \tfrac{1}{2} - \sqrt{\tfrac{1}{4} - 2\varrho\delta\|x_0\|}$$

gegen λ bzw. x.

Beweis: Die Folgen (4) und (3) sind den Gleichungen

$$(6.0)\qquad Ax_0 - \lambda_0 x_0 = 0\,,$$

$$(6.1)\qquad Ax_1 - \lambda_0 x_1 - [Ax_0 - \lambda_0 x_0] + Bx_0 - \lambda_1 x_0 = 0\,,$$

$$\vdots$$

$$(6.n)\qquad Ax_n - \lambda_0 x_n - [Ax_{n-1} - \lambda_0 x_{n-1}] + Bx_{n-1} - \lambda_n(x_0 + x_1 + \ldots + x_{n-1}) + \\ + \lambda_{n-1}(x_0 + x_1 + \ldots + x_{n-2}) = 0$$

entsprechend definiert, wo eben die Wahl $\lambda_n = \lambda_{n-1} + (y_0, \Delta x_{n-1})$ (s. oben (3)) sichert, dass (6.n) für $x_n \in H_\perp$ eindeutig auflösbar sei. Addiert man nämlich (6.0), (6.1) bis (6.n), so erhält man

$$\{Ax_n - \lambda_0 x_n\} + B\,.\,(x_0 + x_1 + \ldots + x_{n-1}) - \lambda_n(x_0 + x_1 + \ldots + x_{n-1}) = 0\,,$$

woraus gleich unsere Aussagen folgen, falls nur Konvergenz in (3) bzw. (4) feststeht. Um auch letztere zu beweisen, setzen wir voraus, dass die Abschätzungen

$$(7)\qquad \|x_n\| \leqq \|x_0\|\,.\,q^n\,, \\ |\lambda_n - \lambda_{n-1}| \leqq \|x_0\|\,.\,\delta\,.\,q^{n-1}$$

für $n = 0, 1, 2, \ldots, N$ schon beweisen sind. Dann aber ist gemäss (3)

$$(8)\qquad |\lambda_{N+1} - \lambda_N| \leqq \|y_0\|\,.\,\|\Delta\|\,.\,\|x_N\| \leqq \delta\,.\,\|x_0\|\, q^N$$

bzw. gemäss (4) und (8)

$$(9)\qquad \|x_{N+1}\| \leqq \varrho\Big\{\delta\,.\,\|x_0\|\, q^N + \delta\,.\,\|x_0\|\, q^N\,.\,\frac{q^N - 1}{q - 1}\,.\,\|x_0\| + \\ + \delta\,.\,\|x_0\|\,\frac{q^{N1} - q}{q - 1}\,.\,\|x_0\|\,.\,q^N\Big\} \leqq \varrho\delta\,.\,\|x_0\|\,.\,q^N\Big\{1 + \frac{\|x_0\|}{1 - q} + \\ + q\,\frac{\|x_0\|}{1 - q} \leqq \|x_0\|\,.\,q^N\Big\{\varrho\delta\,.\,\Big[\|x_0\| + \frac{\|x_0\|\,(1 + q)}{1 - q}\Big]\Big\} = \\ = \|x_0\|\,.\,q^N\,.\,\frac{2\varrho\delta\|x_0\|}{1 - q} \leqq \|x_0\|\,.\,q^N\,.\,q = \|x_0\|\, q^{N+1}\,,$$

da

$$\lambda_{N+1} - \lambda_0 = \sum_{k-1}^{N-1} (\lambda_k - \lambda_{k-1}) \quad \text{und} \quad \frac{2\varrho\delta\|x_0\|}{1-q} \leqq q$$

gilt, wenn nur (2) und (5) feststeht. (7) gilt also auch für $N + 1$, womit der Beweis vollständig durchgeführt ist.

Es sei hier noch bemerkt, dass erstens die Formeln (3) und (4) mit denen der klassischen Störungsmethode völlig übereinstimmen (jedoch ganz anders hergeleitet sind), zweitens, dass die Abschätzungen unter (2), bzw. (5) noch ein wenig verschärfbar sind — wie aus (9) gleich entnommen werden kann.

3. Die obigen Überlegungen kann man leicht auch auf gestörte Probleme erweitern, was auch bei gestörten Matrizen-Eigenwertproblemen und sogar bei nichtlineraren selbstadjungierten Differential-Eigenwertproblemen wichtige Existenzsätze zur Folge hat.

Betrachten wir also das Eigenwertproblem

$$Bx + f(x, \lambda) = \lambda x \tag{10}$$

im Hilbertraum H, wo B ebenso wie oben ein zu A benachbarter linearer Operator ist und $f(x, \lambda) \in H$ in der Nähe von x_0, λ_0 klein ist und einer verallgemeinerten Lipschitz-Bedingung in x und λ mit genügend kleinen Konstanten genügt, d.h.

$$\begin{aligned} \|f(x_2, \lambda_2) - f(x_1, \lambda_1)\| &\leqq C_1 \cdot (\|x_2 - x_0\| + \|x_1 - x_0\|) |\lambda_2 - \lambda_1| + \\ &+ C_2 |\lambda_2 - \lambda_1|^2 + C_3(|\lambda_2 - \lambda_0| + |\lambda_1 - \lambda_0|) \cdot \|x_2 - x_1\| + C_4 \|x_2 - x_1\|^2 . \end{aligned} \tag{11}$$

Satz 2. *Ist*

$$C_1 + C_3 < \tfrac{1}{8}, \; \varrho\delta\|x_0\| < \tfrac{1}{12} \quad \text{und} \quad \|f(x_0, \lambda_0)\| < \delta\|x_0\| , \tag{12}$$

so besitzt auch (10) *einen Eigenwert* λ *bzw. ein Eigenelement* x *in der Nähe von* λ_0 *bzw.* x_0, *und die Reihe*

$$\begin{aligned} &\lambda_0 + \sum_{k=1}^{\infty} (\lambda_k - \lambda_{k-1}); \; \lambda_k - \lambda_{k-1} = \\ &= (y_0, \Delta x_{k-1}) + (y_0, [f(x_0 + \ldots + x_{k-1}, \lambda_{k-1}) - f(x_0 + \ldots + x_{k-2}, \lambda_{k-2})]) \end{aligned} \tag{13}$$

bzw.

$$\begin{aligned} &\sum_{k=0}^{\infty} x_k; \; x_k = R\{\Delta x_{k-1} - (\lambda_k - \lambda_{k-1}) \sum_{n=0}^{k-2} x_n - (\lambda_k - \lambda_0)\, x_{k-1} - \\ &- f(x_0 + \ldots + x_{k-1}, \lambda_{k-1}) + f(x_0 + \ldots + x_{k-2}, \lambda_{k-2})\} \end{aligned} \tag{14}$$

strebt exponentiell gegen λ *bzw.* x *mit einem Quotienten* $q \leqq \frac{1}{2}$.

Der Beweis kann ähnlicherweise geliefert werden wie im Falle des Satzes 1, nur müssen statt (7) die Abschätzungen

$$\|x_n\| \leqq \|x_0\| \cdot q^n \, ; \quad |\lambda_n - \lambda_{n-1}| \leqq \|x_0\| \, 2\delta q^{n-1} \tag{15}$$

in Betracht gezogen werden. Es ist dann leicht zu zeigen, dass mit $q = \frac{1}{2}$ die Abschätzungen (15) für alle n gelten. Man kann diese Abschätzung für q verfeinern, man bekommt dann aber sehr komplizierte Formeln für die Konstanten.

4. Die obigen Überlegungen kann man auch auf mehrfache Eigenwerte erweitern. Ist nämlich λ_0 ein r-facher Eigenwert, $\xi_1, \xi_2, \ldots, \xi_r$ bzw. $\eta_1, \eta_2, \ldots, \eta_r$ die entsprechenden unabhängigen rechts-, bzw. linksseitigen Eigenvektoren, $H_\perp$ der zu allen η_i orthogonale Unterraum von H, so soll

$$x_0 = \alpha_1\xi_1 + \alpha_2\xi_2 + \ldots + \alpha_r\xi_r$$

so gewählt werden, dass in (6.1) $(B - A)\, x_0 - (\lambda_1 - \lambda_0)\, x_0$ bei geeigenter Wahl von λ_1 zu $H_\perp$ [bzw. im Falle gestörter Probleme $(B - A)\, x_0 - (\lambda_1 - \lambda_0)\, x_0 + f(x_0, \lambda_0)$ zu $H_\perp$] gehört. Im weiteren soll (6.n) derart aufgelöst werden, dass die zu $H_\perp$ gehörende Lösung (4) mit einer $(n - 1)$ gliedrigen Linearkombination

$$\sum_{\substack{i=1 \\ j \neq i}}^{r} \alpha_j^{(n)} \xi_j$$

erweitert wird, und zwar so, dass in dem nächsten Schritt λ_{n+1} so gewählt wird, dass $(B - A)\, x_n^{(i)} - (\lambda_{n+1}^{(i)} - \lambda_n^{(i)}) \cdot (x_0 + x_1^{(i)} + \ldots + x_{n-1}^{(i)}) - (\lambda_{n+1}^{(i)} - \lambda_0)\, x_n^{(i)}$ (bzw. eine entsprechende Formel im Falle gestörter Probleme) zu $H_\perp$ gehört. So bekommt man bei jeder Wahl von i $(i = 1, 2, \ldots, r)$ eine entsprechende Reihe für $x^{(i)}$ bzw. $\lambda^{(i)}$.

Als Konvergenzbedingung bekommt man aber hier anstatt (2):

$$\varrho\delta\|x_0\| \leqq \frac{1}{4(r + 1)},$$

und für q die Formel

$$q = \frac{1}{2} - \sqrt{\frac{1}{4} - (r + 1)\, \varrho\delta\|x_0\|}\,.$$

5. Bei den bisher betrachteten Methoden brauchte man eine gute Abschätzung für alle Eigenwerte und Eigenelemente. Hier wollen wir eine andere Störungsmethode zeigen, bei der man nur eine grobe Abschätzung für einen Eigenwert braucht. Der Grundgedanke besteht darin, dass man anstatt des ursprünglichen homogenen Problems ein inhomogenes betrachtet, und die Störungsgrösse durch die abgeänderte Bedingungsgleichung berechnet. Bei selbstadjungierten Differentialgleichungssystemen ändert man nämlich eine homogene Randbedingung in eine inhomogene ab, bei Matrizenproblemen ändert man einen Zeilenvektor. Wir werden den letzteren

Fall betrachten. Es sei also λ_0 ein Annäherungswert eines Eigenwertes des verallgemeinerten Problems

$$(16)\qquad Ax - \lambda Bx = 0\,; \quad A = \begin{pmatrix} a_1^* \\ \vdots \\ a_{n-1}^* \\ a_n^* \end{pmatrix}, \quad B = \begin{pmatrix} b_1^* \\ \vdots \\ b_n^* \end{pmatrix}.$$

Es sei ferner $C = \begin{pmatrix} a_1^* \\ \vdots \\ a_{n-1}^* \\ c_n^* \end{pmatrix}$ mit $c_n^* \neq a_n^*$ eine entsprechende Matrix, und betrachten wir das Hilfsgleichungssystem

$$(17)\qquad Cx - Bx = e_n\,; \quad a_n^* x - \lambda b_n^* x = 0$$

mit $e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. Ist λ ein Eigenwert von (16), so ist der Vektor, der (17) bei diesem λ genügt, ein entsprechender Eigenvektor. Es sei noch vorausgesetzt, dass c_n^* so gewählt ist, dass B und $(C - \lambda_0 B)$ nicht singulär sind. Man bezeichne ferner $(C - \lambda_0 B)^{-1}$ durch R.

Man sucht jetzt die den Gleichungen (17) genügenden Werte λ und x in der Form

$$(18)\qquad \lambda = \lambda_0 + \varepsilon\,; \quad x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \dots$$

Um nun dem ersten Teil von (17) bei jeder Wahl von ε zu genügen, bekommt man das System

$$(18.0)\qquad Cx_0 - \lambda_0 Bx_0 = e_n\,, \quad \text{d.h.} \quad x_0 = Re_n = RB(B^{-1}e_n)\,,$$

$$(18.1)\qquad Cx_1 - \lambda_0 Bx_1 - Bx_0 = 0\,, \quad \text{d.h.} \quad x_1 = RBx_0\,,$$

$$\vdots$$

$$(18.\text{n})\qquad C_{xn} - \lambda_0 Bx_n - Bx_{n-1} = 0\,, \quad \text{d.h.} \quad x_n = RBx_{n-1}\,.$$

Es ist also

$$(19)\qquad \begin{aligned} x &= x_0 + \varepsilon x_1 + \dots = RB[B^{-1}e_n + \varepsilon RB(B^{-1}e_n) + \varepsilon^2 (RB)^2\, B^{-1}e_n + \dots] = \\ &= RB\{E + \varepsilon RB + \varepsilon^2 R^2 B^2 + \dots\}\, B^{-1}e_n = [E - \varepsilon RB]^{-1}\, RB(B^{-1}e_n) = \\ &= [E - \varepsilon RB]^{-1}\, Re_n\,, \end{aligned}$$

falls nur ε genügend klein ist, da RB mit $[E - RB]^{-1}$ vertauschbar ist. Dann ergibt aber die zweite Gleichung in (17) die Bedingung

$$(20)\qquad \begin{gathered} a_n^* x - (\lambda_0 + \varepsilon)\, b_n^* x = 0\,, \quad \text{d.h.} \\ [a_n^* - (\lambda_0 + \varepsilon)\, b_n^*]\, [E - \varepsilon RB]^{-1}\, Re_n = 0 \end{gathered}$$

für ε. Da aber (17) gemäss $c_n^* x - (\lambda_0 + \varepsilon) b_n^* x = 1$ und $[E - RB]^{-1} = \{R[R^{-1} - B]\}^{-1} = [C - (\lambda_0 + \varepsilon) B]^{-1} (C - \lambda_0 B)$ ist, bekommt man anstatt (20) die Gleichung

$$(21) \qquad (a_n^* - c_n^*) [E - \varepsilon RB]^{-1} Re_n = 1 \quad \text{bzw.}$$
$$(a_n^* - c_n^*) [C - (\lambda_0 + \varepsilon B)]^{-1} e_n = 1 ,$$

woraus man ε mit Hilfe eines Schnittes der Reihe $\sum_{k=0} (\varepsilon RB)^k$ genügend genau auch numerisch berechnen kann.

Satz 3. *Wählt man c_n^* so, dass der zu λ_0 am nächsten liegende Eigenwert von $(C - \lambda B) x = 0$ weiter von λ_0 liegt, als der zu λ_0 am nächsten liegende Eigenwert von $(A - \lambda B) x = 0$, so besitzt (21) eine im absoluten Wert kleinste Lösung ε, für welche $\|\varepsilon RB\| < 1$ ist, d.h. für welche die Reihe (19) absolut konvergent ist.*

Beweis: Es bezeichne μ bzw. ν den zu λ_0 am nächsten liegenden Eigenwert von $(C - \lambda B) x = 0$ bzw. $(A - \lambda B) x = 0$. Daraus aber, dass $\mu - \lambda_0$ der absolut kleinste verallgemeinerte (nämlich in Hinsicht zu B) Eigenwert von $D = C - \lambda_0 B$ ist, und somit auch die Norm von RB nicht grösser als $|\lambda_0 - \mu|$ ist, folgt

$$(22) \qquad \|RB\| \leqq |\mu - \lambda_0| .$$

Da nun weiter $\varepsilon = \nu - \lambda_0$ mit dem zugehörigen Eigenvektor der Gleichung (21) genügt und $|\nu - \lambda_0| < |\mu - \lambda_0|$ ist, konvergiert die Reihe (19) mit diesem $\varepsilon = \nu - \lambda_0$ gleichmässig gegen den betrachteten Eigenvektor.

6. Eine sukzessive Methode für Matrizeneigenwertprobleme können wir folgenderweise bekommen: kennt man alle Eigenwerte und Eigenelemente der Matrizen B_k (es seien dies $\lambda_i^{(k)}$ resp. $s_i^{(k)}$ $(i = 1, 2, \ldots, n)$), kann man bei beliebiger Wahl von a^*, b^* und $j_1^{(k)}$ bzw. $j_2^{(k)} [j_1^{(k)} \neq j_2^{(k)};\ 1 \leqq j_1^{(k)} \leqq n;\ 1 \leqq j_2^{(k)} \leqq n]$ auch Eigenelemente von

$$(23) \qquad B_k + s_{j_1}^{(k)} \,.\, a^* + s_{j_2}^{(k)} \,.\, b^* = B_{k+1}$$

angeben — und zwar mit Hilfe der Lösung einer quadratischen Gleichung, eines linearen Gleichungssystems mit zwei Unbekannten und $(n-2)$ unabhängiger linearer Gleichungen.

Wählt man nun a^*, b^*, j_1 und j_2 so, dass

$$(24) \qquad \|B_{k+1} - A\|_\Delta^2 = \min$$

gilt, wo A die zu betrachtende Matrix ist und $\|\ \|_\Delta^2$ die absolute Quadratsumme der Matrizenelemente bezeichnet (sog. Innenprodukt-Norm, d.h. Absolutquadratsumme der Eigenwerte), so bekommt man a^* und b^* mit Hilfe der Lösung von n unabhängigen linearen Gleichungssystemen mit je zwei Unbekannten.

Nach jedem Schritt setzt man in dieser Weise die Differenz (in Innenproduktnorm) zwischen B_{k+1} und A herab; nach endlich vielen Schritten kann man sich also A genügend annähern.

Über die Konvergenzgeschwindigkeit kann man im allgemeinen nicht viel sagen. Stellt man nämlich A als $B_k + \Delta$ dar, so soll man Δ durch eine Kombination $s_{j_1}^{(k)} . a^* + s_{j_2}^{(k)} . b^*$ möglichst gut in der Norm approximieren. Wären nun alle Spaltenvektoren von Δ orthogonal zur Ebene der Vektoren s_{j_1} und s_{j_2}, so könnte man mit B_{k+1} nicht besser als mit B_k approximieren. Da aber die Eigenvektoren von B_k ein vollständiges System darstellen, so kann das nicht für alle Kombinationen j_1 und j_2 der Fall sein. Je paralleller jedoch die Spaltenvektoren von $A - B_k$ zueinander und je orthogonaler sie zu den Eigenvektoren von B_k verlaufen, desto wenig besser approximiert B_{k+1} als B_k das betrachtete A.

In der Praxis erwies sich die oben angegebene sukzessive Methode für eine grobe Approximation von A als ganz gut; verfügt man aber schon über eine Approximation, bei welcher die Störungsmethode anwendbar ist, dann ist letztere vorzuziehen.

T. Frey, Magyar Tudományos Akadémia Számítás-technikai Központja, Uri u. 53, Budapest I, Ungarn.