

Verfahren zur Berechnung des Spektralradius nichtnegativer irreduzibler Matrizen

Von

L. Elsner, Erlangen

(Eingegangen am 26. Juni 1970)

Zusammenfassung — Summary

Verfahren zur Berechnung des Spektralradius nichtnegativer irreduzibler Matrizen. Es werden mehrere Iterationsverfahren zur Bestimmung des Spektralradius ρ einer nichtnegativen Matrix A angegeben. Ihre Konvergenz wird unter der Voraussetzung bewiesen, daß A irreduzibel ist. Die Verfahren sind, im Gegensatz zur Potenzmethode, unempfindlich gegen das Vorhandensein mehrerer Eigenwerte vom Betrage ρ .

Methods for Computing the Spectral Radius of Nonnegative Irreducible Matrices. Several iterative methods for determining the spectral radius ρ of a nonnegative matrix A are given. Convergence is proved under the assumption of irreducibility. Contrary to the power method, these methods are insensitive to the presence of other eigenvalues with modulus ρ .

1. Einleitung

Es sei \mathfrak{S} die Menge der nichtnegativen irreduziblen $N \times N$ -Matrizen, $A \in \mathfrak{S}$ und $\omega = \rho(A)$ der Spektralradius von A . Dann gibt es einen positiven Vektor $y > 0$ mit $Ay = \omega y$. Die Bestimmung von y ist bekanntlich äquivalent mit der Bestimmung einer Diagonalmatrix Y mit positiven Diagonalelementen, so daß $Y^{-1}AY$ konstante Zeilensummen hat. Bei den meisten numerischen Verfahren wird Y als unendliches Produkt

$$Y = D_1 D_2 D_3 \dots$$

von Diagonalmatrizen der Gestalt

$$D_i = \text{diag} (1, \dots, 1, d_i, 1, \dots, 1)$$

aufgebaut. Sei $\Delta_i = D_1 D_2 \dots D_i$ und $A_i = \Delta_i^{-1} A \Delta_i$. In dieser Darstellung erscheinen die Verfahren am einfachsten, bei denen die D_i nur aus A_{i-1} berechnet werden. Solche Verfahren kann man etwa Einschrittverfahren nennen, im Gegensatz zu k -Schrittverfahren, wo D_i eine Funktion von A_{i-1}, \dots, A_{i-k} ist. In dieser Terminologie ist die Potenzmethode ein $2N$ -Schrittverfahren. Die Verfahren in [1] und [3] stellen N -Schrittverfahren dar.

Einschrittverfahren wurden in [2] behandelt. Der dort angegebene Algorithmus konvergiert allerdings nicht für alle $A \in \mathfrak{S}$. In der vorliegenden Arbeit geben wir Einschrittverfahren an, die für alle nichtnegativen irreduziblen Matrizen konvergieren.

Nach einigen Vorbereitungen wird das Hauptergebnis in *Satz 3* formuliert. Er gibt ziemlich allgemeine Kriterien für die Auswahl der d_i an, die die Konvergenz für alle $A \in \mathfrak{S}$ garantieren. Anhand dieser Ergebnisse werden im vierten Abschnitt einige spezielle Algorithmen angegeben. Insbesondere wird gezeigt, wie das Verfahren in [2] modifiziert werden kann, so daß es für alle $A \in \mathfrak{S}$ konvergiert. Ebenso wird eine vereinfachte Variante von [3] angegeben. Schließlich wird über (nicht sehr umfangreiche) numerische Erfahrungen berichtet.

2. Hilfsbetrachtungen

Wir verwenden im folgenden die Bezeichnungsweise $x > 0$ bzw. $x \geq 0$ für positive bzw. nichtnegative Vektoren. Analoges gilt für Matrizen. Ist x ein Vektor, so sei $\|x\| = \max_i |x_i|$.

Es sei $m = \min \{a_{ik} \mid a_{ik} > 0, i \neq k\}$.

Satz 1. Sei $x > 0$ ein gegebener Vektor, $\varepsilon > 0$ und

$$Ax \leq \omega(1 + \varepsilon)x.$$

Sei

$$\gamma(\varepsilon) = 1 + \frac{(1 + \varepsilon)^{N-1} - 1}{m^{N-1} \min_{i,k} \left(\frac{y_i}{y_k} \right)} \omega^{N-1}. \quad (1)$$

Dann gilt

$$\min_i \frac{x_i}{y_i} \leq \frac{x_k}{y_k} \leq \gamma(\varepsilon) \min_i \frac{x_i}{y_i} \quad k = 1, \dots, N \quad (2)$$

und

$$Ax \geq \frac{\omega}{\gamma(\varepsilon)} x. \quad (3)$$

Beweis. Wir betrachten zunächst den Fall $\omega = 1$, $y = (1, \dots, 1)$, also A stochastisch.

Sei $x_j = \min_i x_i$. Zu gegebenem $k \neq j$ existiert wegen der Irreduzibilität von A eine Zahl s , $1 \leq s \leq N - 1$ mit $(A^s)_{j,k} = a_{jk}^{(s)} > 0$. Wegen $A^s x \leq (1 + \varepsilon)^s x$, $A^s y = y$ gilt

$$\sum_n a_{jn}^{(s)} \left(\frac{x_n}{x_j} - 1 \right) \leq (1 + \varepsilon)^s - 1.$$

Da x_j minimal ist, sind alle Summanden der linken Seite nichtnegativ, also

$$\frac{x_k}{x_j} - 1 \leq ((1 + \varepsilon)^s - 1) / a_{jk}^{(s)}$$

und daher

$$x_j \leq x_k \leq \left(1 + \frac{(1 + \varepsilon)^s - 1}{a_{jk}^{(s)}}\right) x_j.$$

Wegen $a_{jk}^{(s)} \geq m^s \geq m^{N-1}$ gilt nun (2). Es folgt $x_k/x_i \geq \gamma^{-1}$ für beliebige i, k und daher

$$\sum_k a_{ik} \frac{x_k}{x_i} \geq \gamma^{-1} \sum_k a_{ik} = \gamma^{-1},$$

also (3).

Der allgemeine Fall wird nun auf diesen Fall zurückgeführt. Ist Y die Diagonalmatrix mit den Diagonalelementen y_i , so ist $B = \omega^{-1} Y^{-1} A Y$ stochastisch und $BY^{-1}x \leq (1 + \varepsilon) Y^{-1}x$. Für B und $Y^{-1}x$ ist bereits alles bewiesen. Durch Rücktransformation folgen die Aussagen auch für den allgemeinen Fall.

Bemerkung. Ist $A > 0$, so ist stets $s = 1$ möglich. Daher gilt dann (2) und (3) schon mit

$$\gamma(\varepsilon) = 1 + \frac{\varepsilon\omega}{m} \max_{i,k} \frac{y_i}{y_k}.$$

Es sei im folgenden y normiert, $\|y\| = 1$.

x^n ($n = 1, 2, \dots$) sei eine Folge positiver Vektoren, $\|x^n\| = 1$, $x^n = (x_1^n, x_2^n, \dots, x_N^n)$, $R_i^n = \sum_k a_{ik} x_k^n/x_1^n$ $i = 1, \dots, N$ $n = 1, 2, \dots$
 $R_n = \max_i R_i^n$.

Satz 2. *Es sei $\lim R_n = \omega$. Dann gilt*

$$\lim R_i^n = \omega \quad i = 1, \dots, N, \quad (4)$$

$$\lim x^n = y. \quad (5)$$

Beweis. Es ist $Ax^n \leq R_n x^n = \omega(1 + \varepsilon_n)x^n$ und offenbar $\lim \varepsilon_n = 0$. Daher strebt $\gamma(\varepsilon_n)$ gegen 1. (3) zeigt nun die Behauptung (4).

Ist x ein Häufungspunkt von $\{x^n\}$, so zeigt (2), daß x ein Vielfaches von y ist. Wegen der Normierung ist $x = y$. $\{x^n\}$ besitzt also nur einen Häufungspunkt, nämlich y . Daher konvergiert $\{x^n\}$ gegen y . Q. e. d.

Bemerkung. Satz 2 findet sich im wesentlichen bereits in [2]. Die Herleitung über Satz 1, der quantitative Aussagen enthält, ist neu.

3. Eine Klasse von Verfahren

Wir konstruieren rekursiv eine Vektorfolge x^n : Sei $x^1 = (1, \dots, 1)$. Es sei $\nu = \nu(n)$, $\mu = \mu(n)$ ausgewählt, so daß

$$R_\nu^n \leq R_i^n \leq R_\mu^n \quad i = 1, \dots, N$$

gilt. Mit einer noch zu bestimmenden Größe $d = d_n$ sei $\tilde{x} = (\tilde{x}_k)$ mit

$$\tilde{x}_k = \begin{cases} x_k^n & k \neq \nu \\ d x_\nu^n & k = \nu \end{cases} \quad x^{n+1} = \tilde{x} / \|\tilde{x}\|.$$

Dann können wir den folgenden Konvergenzsatz beweisen.

Satz 3. d_n werde so gewählt, daß folgende Aussagen gelten:

- (a) $0 < d_n < 1$, falls $R_v^n \neq R_\mu^n$.
- (b) Existiert eine Teilfolge $n(i)$ mit $d_{n(i)} \rightarrow 1$, so folgt $\lim R_n = \omega$.
- (c) Es existiert eine Zahl α , $0 < \alpha < 1$, α unabhängig von n mit $R_v^{n+1} \leq \alpha R_\mu^n + (1 - \alpha) R_v^n$.

Dann ist $\lim R_n = \omega$ und es sind (4) und (5) erfüllt.

Beweis. Wegen Satz 2 genügt es, $\lim R_n = \omega$ zu zeigen. Allgemein gilt

$$R_i^{n+1} = R_i^n + (d_n - 1) a_{iv} \frac{x_v^n}{x_i^n}, \quad i \neq v, \quad (6)$$

$$R_v^{n+1} = a_{vv} + \frac{R_v^n - a_{vv}}{d_n}. \quad (7)$$

Aus (a) und (c) folgt nun sofort

$$R_n \geq R_{n+1} \geq \dots \quad (8)$$

Die R_i sind also monoton fallend und nach unten durch ω beschränkt.

Es sei $I = \{p \in \{1, \dots, N\}, p = v(n) \text{ nur für endlich viele } n\}$. Wir unterscheiden zwei Fälle.

Fall 1. $I \neq \emptyset$. Da $I \neq \{1, \dots, N\}$ ist, gibt es wegen der Irreduzibilität von A ein $k \in I$, $l \notin I$ mit $a_{kl} > 0$.

Wegen (8) und (2) ist für alle n die Größe $a_{kl}^n = a_{kl} x_l^n / x_k^n$ nach unten beschränkt. Es sei für alle $n \geq N_0: v(n) \neq k$. Dann ist wegen (6) für $n \geq N_0$

$$R_k^{n+1} = R_k^n + a_{i_v}^n (d_n - 1).$$

Ist $1 - d_n \geq \varepsilon > 0$ für alle n , so folgt weiter

$$R_k^{n+s} \leq R_k^n - \varepsilon \sum_{r=n}^{n+s-1} a_{k\nu(r)}^r = R_k^n - \varepsilon w_s.$$

Da immer wieder $\nu(r) = l$ ist, gilt $\lim w_s = \infty$, ein Widerspruch gegen $R_k^{n+s} > 0$. Es gibt daher eine Teilfolge $n(i)$ mit $d_{n(i)} \rightarrow 1$. Mit (b) folgt die Behauptung.

Fall 2. $I = \emptyset$. Sei p ein Index, so daß für abzählbar viele n gilt $\mu(n) = p$. Durch Induktion läßt sich die Existenz von Zahlenfolgen $m(i)$, $n(i)$ zeigen mit

$$\begin{aligned} m(i) &< n(i) < m(i+1), \\ \nu(m(i)) &= \mu(n(i)) = p, \\ m(i) &< n < n(i) \Rightarrow \nu(n) \neq p. \end{aligned}$$

Dann gilt offenbar

$$R_p^{m(i)+1} \geq R_p^{n(i)} = R_n(i). \quad (9)$$

Nach (c) folgt

$$\begin{aligned} 0 \leq R_n(i) - \omega &\leq R_p^{m(i)+1} - \omega \leq \alpha R_{\mu}^{m(i)} + (1 - \alpha) R_p^{m(i)} - \omega \\ &\leq \alpha (R_{\mu}^{m(i)} - \omega) = \alpha (R_m(i) - \omega). \end{aligned}$$

Zusammen mit (8) folgt daraus die Behauptung. Q. e. d.

4. Spezielle Verfahren

Satz 3 legt das folgende Verfahren nahe:

Verfahren 1. Zu vorgegebenem $0 < \alpha < 1$ wähle d_n so, daß

$$R_v^{n+1} = \alpha R_{\mu}^n + (1 - \alpha) R_v^n$$

ist. Das führt auf

$$d_n = \frac{R_v^n - a_{vv}}{\alpha R_{\mu}^n + (1 - \alpha) R_v^n - a_{vv}} = \frac{R_v^n - a_{vv}}{R_v^n - a_{vv} + \alpha (R_{\mu}^n - R_v^n)}. \quad (10)$$

Damit ist offenbar (a) und (c) von *Satz 3* erfüllt. Aus $d_n(i) \rightarrow 1$ folgt $R_{\mu}^{n(i)} - R_v^{n(i)} \rightarrow 0$, und wegen $R_v^n \leq \omega \leq R_{\mu}^n$ gilt $\lim R_n(i) = \omega$, also wegen (8) auch $\lim R_n = \omega$. Mithin ist (b) erfüllt. *Verfahren 1* konvergiert daher für alle nichtnegativen irreduziblen A .

Etwas komplizierter ist das folgende Verfahren, das im wesentlichen eine vereinfachte Version des in [3] vorgeschlagenen Algorithmus ist. Wie in [3] sei

$$l_i^n(t) = R_i^n + a_{iv}^n(t-1) \quad i \neq v(n), \quad (11)$$

$$h_v^n(t) = a_{vv} + (R_v^n - a_{vv})/t. \quad (12)$$

Es sei ξ eine Lösung von $l_{\mu}^n(\xi) = h_v^n(\xi)$. Dann genügt ξ der Gleichung

$$g(\xi) \equiv a_{\mu v}^n(\xi^2 - \xi) + (R_{\mu}^n - a_{vv})\xi + (a_{vv} - R_v^n) = 0. \quad (13)$$

Wegen $g(0) < 0$, $g(1) > 0$, $g(\xi_0^n) \leq 0$ für $\xi_0^n = \frac{a_{vv} - R_v^n}{a_{vv} - R_{\mu}^n}$ gibt es genau eine Lösung ξ in

$$0 < \xi < 1,$$

für die sogar

$$\frac{R_v^n - a_{vv}}{R_{\mu}^n - a_{vv}} \leq \xi < 1 \quad (14)$$

gilt. Der explizite Wert von ξ ist

$$\xi = \xi^n = \begin{cases} \frac{R_v^n - a_{vv}}{R_{\mu}^n - a_{vv}} & (a_{\mu v} = 0) \\ \frac{a_{\mu v}^n + a_{vv} - R_{\mu}^n + ((R_{\mu}^n - a_{\mu v}^n - a_{vv})^2 + 4 a_{\mu v}^n (R_v^n - a_{vv}))^{\frac{1}{2}}}{2 a_{\mu v}^n} & (a_{\mu v} > 0) \end{cases} \quad (15)$$

Verfahren 2. Sei $0 < \alpha < 1$, $\beta = \alpha^{-1}$ und

$$d_n = \alpha (\xi_n - 1) + 1 = \alpha \xi_n + (1 - \alpha) \cdot 1.$$

Wegen $0 < \xi_n < 1$ gilt $1 - \alpha < d_n < 1$.

Wir verifizieren die Voraussetzungen von *Satz 3*.

(a) ist trivial, (b) rechnet man unmittelbar nach.

(c): Für reelle Zahlen $0 < \eta \leq \zeta$ gilt, wie man leicht beweist,

$$\frac{\eta}{\alpha \zeta + (1 - \alpha) \eta} \leq \frac{(1 - \alpha) \zeta + \alpha \eta}{\zeta}. \quad (16)$$

Aus (14) folgt $1 + \beta (d_n - 1) \geq \frac{R_v^n - a_{vv}}{R_\mu^n - a_{vv}}$. Daher

$$d_n \geq \frac{(1 - \alpha) (R_\mu^n - a_{vv}) + \alpha (R_v^n - a_{vv})}{R_\mu^n - a_{vv}} \quad (17)$$

und wegen (16)

$$\geq \frac{R_v^n - a_{vv}}{\alpha (R_\mu^n - a_{vv}) + (1 - \alpha) (R_v^n - a_{vv})}.$$

Aus (10) und (7) folgt dann

$$R_v^{n+1} \leq \alpha R_\mu^n + (1 - \alpha) R_v^n, \quad \text{Q. e. d.}$$

Bemerkung. Unter der in der Praxis wohl stets erfüllten Voraussetzung, daß für alle n die Relation $R_i^n = R_n$ nur für ein i gilt, fällt für $\alpha = 1/2$ *Verfahren 2* mit dem in [3] angegebenen Algorithmus zusammen. Allerdings ist unser Verfahren einfacher zu programmieren.

In [2] wird der Fall $\alpha = 1$ betrachtet. Dann ist

$$R_v^{n+1} = R_\mu^{n+1}. \quad (18)$$

(a) und (b) sind erfüllt. Ist $a_{\mu v} > 0$, so folgt aus (18), (6) und (7)

$$R_v^{n+1} - R_v^n = \frac{R_v^n - a_{vv}}{d_n \alpha_{\mu v}^n} (R_\mu^n - R_v^{n+1}).$$

Wegen (17) für $\alpha = 1$ folgt

$$\leq \frac{R_\mu^n - a_{vv}}{\alpha_{\mu v}^n} (R_\mu^n - R_v^{n+1}) \leq K (R_\mu^n - R_v^{n+1}),$$

wobei K unabhängig von n wählbar ist. Es folgt

$$R_v^{n+1} \leq \frac{K}{1 + K} R_\mu^n + \frac{1}{1 + K} R_v^n.$$

Für $a_{\mu v} = 0$ ist wegen $R_v^{n+1} = R_\mu^{n+1} = R_\mu^n$ eine solche Ungleichung nicht mehr möglich. Wie in [2] gezeigt wird, konvergiert dieses Verfahren auch nicht für alle irreduziblen A . Es ist nun klar, wie dieser Algorithmus abzuändern ist, damit Konvergenz eintritt: Im Fall $a_{\mu v} = 0$ muß die

Ungleichung in (c) erzwungen werden. Das kann dadurch geschehen, indem man dann d_n wie in *Verfahren 1* oder *2* wählt. Wir notieren nur die erste Möglichkeit.

Verfahren 3. Wähle mit einem $0 < \alpha < 1$

$$d_n = \begin{cases} \xi_n & \text{falls } a_{\mu\nu} > 0 \\ \frac{R_\nu^n - a_{\nu\nu}}{\alpha R_\mu^n + (1 - \alpha) R_\nu^n - a_{\nu\nu}} & \text{falls } a_{\mu\nu} = 0. \end{cases}$$

Die Konvergenz ist nach *Satz 3* gesichert.

Bemerkung. Ist die folgende Bedingung erfüllt

$$\exists p \text{ mit } a_{ip} > 0 \text{ für alle } i = 1, \dots, N, i \neq p,$$

so konvergiert das *Verfahren 2* auch für $\alpha = 1$. Da die Bedingung (c) von *Satz 3* nicht erfüllt ist, muß der Konvergenzbeweis im Fall 2 ($I = \emptyset$) modifiziert werden. Das geschieht wie in [2]. Da $p \notin I$, existiert eine Folge $k(i)$ mit $\nu(k(i)) = p$. Daher ist für ein $j \neq p$

$$\begin{aligned} R_{k(i)+1} &= R_j^{k(i)+1} = R_j^{k(i)} + a_{jp}^{k(i)} (d_{k(i)} - 1) \\ &\leq R_{k(i)} + a_{jp}^{k(i)} (d_{k(i)} - 1). \end{aligned}$$

Daher

$$0 < 1 - d_{k(i)} \leq (R_{k(i)} - R_{k(i)+1}) / a_{jp}^{k(i)}.$$

Da die R_n konvergieren, geht die rechte Seite gegen Null, d. h. $d_{k(i)} \rightarrow 1$ und wegen (b) gilt $R_n \rightarrow \omega$. Q. e. d.

Diese hinreichende Bedingung ist schwächer als die in [2] angegebene.

5. Beispiele

Es wurden etliche numerische Beispiele durchgerechnet, um einen Vergleich zwischen den Verfahren zu ermöglichen und den Einfluß des Parameters α zu studieren. Dabei ergab sich folgendes Bild, wenn man die Anzahl der Iterationsschritte zur Erreichung einer gewissen Genauigkeit $R_\mu^n - R_\nu^n$ zugrundelegt.

Bei vollbesetzter Matrix war das *Verfahren 3* (in diesem Fall identisch mit dem Verfahren von [2]) eindeutig überlegen. Beim Auftreten von Nullen verschob sich das Bild: Hier erwiesen sich alle drei Verfahren als etwa gleich gut. Zudem zeigte sich, daß es dann jeweils am günstigsten war, α im Bereich 0.5–0.7 zu wählen. Legt man die Anzahl der Rechenschritte zugrunde, so schneidet *Verfahren 1* wegen seiner Einfachheit etwas günstiger ab. Alle Verfahren erwiesen sich als unempfindlich gegen das Vorhandensein von Eigenwerten mit Beträgen nahe bei $\rho(A)$, wie auch in [2] beobachtet wurde.

Als Beispiel für schwachbesetzte Matrizen, deren Eigenwerte zusätzlich alle den gleichen Betrag haben, führen wir

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.25 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

an. In der folgenden Tabelle sind die Anzahl der Schritte angegeben, um $R_\mu - R_\nu \leq 10^{-4}$ zu erreichen.

α	Verfahren 1	Verfahren 2	Verfahren 3
0.9	82	100	129
0.7	28	28	34
0.5	16	30	20
0.3	29	49	20
0.1	115	195	76

Das *Verfahren 2* mit $\alpha = 1$ ist hier zufällig nach 6 Schritten exakt am Ziel.

Literatur

- [1] BRAUER, A.: On the Characteristic Roots of Non-negative Matrices, in: Recent Advances in Matrix Theory (SCHNEIDER H., ed.). Madison: University of Wisconsin Press. 1964.
- [2] HALL, C. A., and T. A. PORCHING: Computing the maximal eigenvalue and eigenvector of a positive matrix. SIAM J. Num. Anal. **5**, 269–274 (1968).
- [3] HALL, C. A., and T. A. PORCHING: Computing the maximal eigenvalue and eigenvector of a nonnegative irreducible matrix. SIAM J. Num. Anal. **5**, 470–474 (1968).

Ludwig Elsner
Institut für Angewandte Mathematik I
der Universität Erlangen
Erwin-Rommel-Straße 60, D-852 Erlangen
Deutschland