

Aplikace matematiky

Miroslav Šisler

Bemerkungen zur Optimierung eines zweiparametrischen Iterationsverfahrens

Aplikace matematiky, Vol. 21 (1976), No. 3, 213--220

Persistent URL: <http://dml.cz/dmlcz/103639>

Terms of use:

© Institute of Mathematics AS CR, 1976

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://project.dml.cz>

BEMERKUNGEN ZUR OPTIMIERUNG EINES ZWEIPARAMETRIGEN
ITERATIONSVERFAHRENS

MIROSLAV ŠISLER

(Eingegangen am 12. November 1975)

Die Arbeit knüpft unmittelbar an die Resultate der Arbeiten [2] und [3] desselben Verfassers an. Es handelt sich wieder um ein zweiparametriges Iterationsverfahren für die Lösung eines linearen Gleichungssystems mit n Unbekannten von der Form

$$(1) \quad \mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{b}.$$

Das Iterationsverfahren ist dabei folgenderweise definiert:

$$(2) \quad \mathbf{x}_{v+1} = \mathbf{T}(\alpha, \beta) \mathbf{x}_v + \mathbf{P}(\alpha, \beta) \mathbf{b}, \quad v = 0, 1, 2, \dots$$

$$(3) \quad \mathbf{T}(\alpha, \beta) = (\alpha \mathbf{E} + \beta \mathbf{L})^{-1} [(\alpha - 1) \mathbf{E} + (\beta + 1) \mathbf{L} + \mathbf{U}],$$

$$\mathbf{P}(\alpha, \beta) = (\alpha \mathbf{E} + \beta \mathbf{L})^{-1};$$

hier bezeichnet α, β reelle, von Null verschiedene Parameter und $\mathbf{B} = \mathbf{L} + \mathbf{U}$, wobei \mathbf{B} im allgemeinen eine unsymmetrische Blockmatrix ist, deren Diagonalblöcke Nullmatrizen sind und \mathbf{L} bzw. \mathbf{U} die untere bzw. obere verallgemeinerte Dreiecksmatrix ist.

In der Arbeit setzt man voraus, dass \mathbf{B} eine nichtsinguläre 2-zyklische Matrix ist. Man setzt nämlich voraus, dass die Matrix \mathbf{B} einen gewissen Vektorraum \mathbf{R} der Dimension n in \mathbf{R} abbildet, dass der Vektorraum \mathbf{R} durch die direkte Summe der Unterräume v_1, \dots, v_m ($m \leq n$) gebildet wird und dass die Matrizen \mathbf{L}, \mathbf{U} die Bedingungen

$$\mathbf{L}v_i \subset v_{i+1}, \quad \mathbf{U}v_i \subset v_{i-1}, \quad i = 1, \dots, m$$

erfüllen, wobei $v_i = 0$ für $i < 1$ und $i > m$ ist.

Es gilt nun der folgende Satz:

Satz 1. *Es sei λ ein Eigenwert der Matrix $\mathbf{T}(\alpha, \beta)$. Falls $\lambda \neq (\beta + 1)/\beta$ ist und falls die Zahl μ die Beziehung*

$$(4) \quad (1 - \alpha + \lambda\alpha)^2 = \mu^2(1 + \beta - \lambda\beta)$$

erfüllt, ist μ der Eigenwert der Matrix \mathbf{B} . Falls im Gegenteil μ ein Eigenwert der Matrix \mathbf{B} ist, dann ist jede Wurzel λ der quadratischen Gleichung (4) ein Eigenwert der Matrix $\mathbf{T}(\alpha, \beta)$.

Der Satz folgt sofort vom Satz 1 der Arbeit [2] für $p = 2, h = k = 1$.

In der Arbeit wird man ferner voraussetzen, dass die Eigenwerte μ_i^2 der Matrix \mathbf{B}^2 die Ungleichungen $0 < \mu_i^2 < 1, i = 1, \dots, n$ erfüllen.

Man bezeichne M bzw. m eine solche positive Zahl, dass $M^2 = \max_{i=1, \dots, n} \mu_i^2$ bzw. $m^2 = \min_{i=1, \dots, n} \mu_i^2$ gilt.

Es ist aus der Arbeit [2] bekannt, dass für $\beta = -1$ die durch die Formeln (2),(3) definierte Methode dem gewöhnlichen Oberrelaxationsverfahren entspricht. Dabei folgt aus dem Satz 2 der Arbeit [3], dass der Spektralradius $\varrho(\mathbf{T}(\alpha, -1))$ der Matrix $\mathbf{T}(\alpha, -1)$ für $\alpha = \frac{1}{2}[1 + \sqrt{(1 - M^2)}]$ seinen Minimalwert $[1 - \sqrt{(1 - M^2)}] / [1 + \sqrt{(1 - M^2)}]$ annimmt.

Man bezeichne jetzt mit Ω die Menge solcher Punkte $[\alpha, \beta], \alpha > 0$ für die

$$(5) \quad \varrho(\mathbf{T}(\alpha, \beta)) \leq [1 - \sqrt{(1 - M^2)}] / [1 + \sqrt{(1 - M^2)}]$$

gilt, d. h. die untersuchte Methode konvergiert zumindestens so schnell, wie das Oberrelaxationsverfahren. Es gilt dann der folgende Satz:

Satz 2. *Es sei*

$$b_1(\alpha) = [(1 + \sqrt{(1 - M^2)})^2 \sqrt{(1 - M^2)} + 4\alpha^2 \sqrt{(1 - M^2)} - 4\alpha(1 + \sqrt{(1 - M^2)})] / 2M^2(1 + \sqrt{(1 - M^2)}),$$

$$b_2(\alpha) = [(1 + \sqrt{(1 - M^2)})^2 (1 - M^2) + 4\alpha^2 - 4\alpha(1 + \sqrt{(1 - M^2)})] / 2M^2(1 + \sqrt{(1 - M^2)}),$$

$$b_3(\alpha) = (1 - \alpha)^2 / m^2 - \alpha^2(1 - \sqrt{(1 - M^2)})^2 / m^2 (1 + \sqrt{(1 - M^2)}) - 1.$$

Dann ist $[\alpha, \beta] \in \Omega$ genau dann, wenn gleichzeitig

$$\beta \leq b_1(\alpha), \quad \beta \leq b_2(\alpha), \quad \beta \geq b_3(\alpha), \quad \alpha > 0$$

gilt, wobei die Gleichheit in der Beziehung (5) nur für die Grenzpunkte der Menge Ω eintritt.

Der Satz folgt sofort aus dem Satz 3 der Arbeit [3].

Der folgende Satz 3 beschreibt die konkrete Form des Gebietes Ω in der Abhängigkeit von der Zahl m . Bevor wir den Satz formulieren werden, führen wir zuerst folgende Bezeichnungen ein:

$$\alpha_1 = \frac{1}{2}(1 + \sqrt{(1 - M^2)}), \quad \alpha_3 = \frac{1}{2}(1 + \sqrt{(1 - M^2)})^4 \sqrt{(1 - M^2)},$$

$$\alpha_5 = \frac{1}{2}(1 + \sqrt{(1 - M^2)}) \sqrt{(1 - M^2)}.$$

Satz 3. I. Es sei $m^2 \leq 1 - \sqrt{(1 - M^2)}$. Dann enthält die Menge Ω den einzigen Punkt $[\alpha_1, -1]$.

II. Es sei

$$(6) \quad 1 - \sqrt{(1 - M^2)} < m^2 \leq 2(1 - \sqrt{(1 - M^2)})(1 - \sqrt[4]{(1 - M^2)^3})/[2 - \sqrt{(1 - M^2)} - \sqrt[4]{(1 - M^2)^3}].$$

Dann ist Ω die Menge der Punkte, für die gleichzeitig die Ungleichungen

$$\alpha_2 \leq \alpha \leq \alpha_1, \quad b_3(\alpha) \leq \beta \leq b_1(\alpha)$$

gelten, wobei α_2 die positive Wurzel der Gleichung $b_1(\alpha) = b_3(\alpha)$ ist. (Eine solche Wurzel existiert und liegt im Intervall $\langle \alpha_3, \alpha_1 \rangle$; es ist $\alpha_2 = \alpha_3$, solange in (6) die Gleichheit gilt.)

III. Es sei

$$(7) \quad 2(1 - \sqrt{(1 - M^2)})(1 - \sqrt[4]{(1 - M^2)^3})/[2 - \sqrt{(1 - M^2)} - \sqrt[4]{(1 - M^2)^3}] < m^2 \leq M^2.$$

Dann ist die Menge Ω durch die folgende Ungleichungen definiert:

$$\text{für } \alpha_4 \leq \alpha \leq \alpha_3 \text{ ist } b_3(\alpha) \leq \beta \leq b_2(\alpha),$$

$$\text{für } \alpha_3 \leq \alpha \leq \alpha_1 \text{ ist } b_3(\alpha) \leq \beta \leq b_1(\alpha).$$

Dabei ist α_4 die positive Wurzel der Gleichung $b_2(\alpha) = b_3(\alpha)$. (Eine solche Wurzel existiert und liegt im Intervall $\langle \alpha_5, \alpha_3 \rangle$; für $m^2 = M^2$ gilt $\alpha_4 = \alpha_5$.)

Satz 3 folgt sofort aus dem Satz 3 der Arbeit [3].

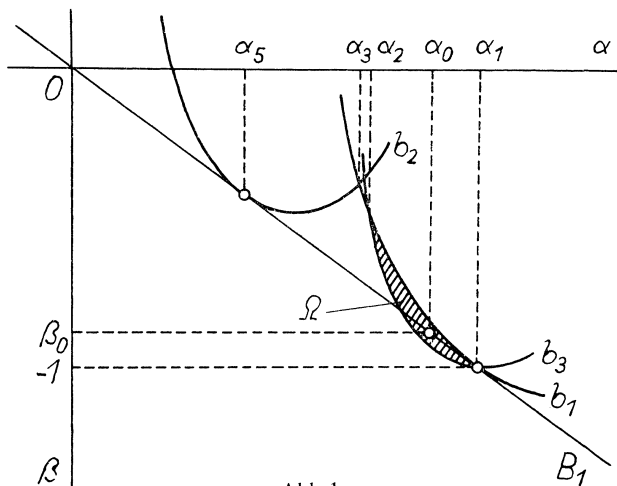
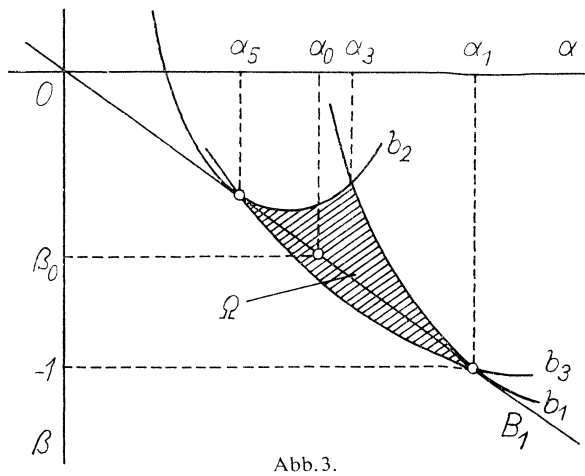
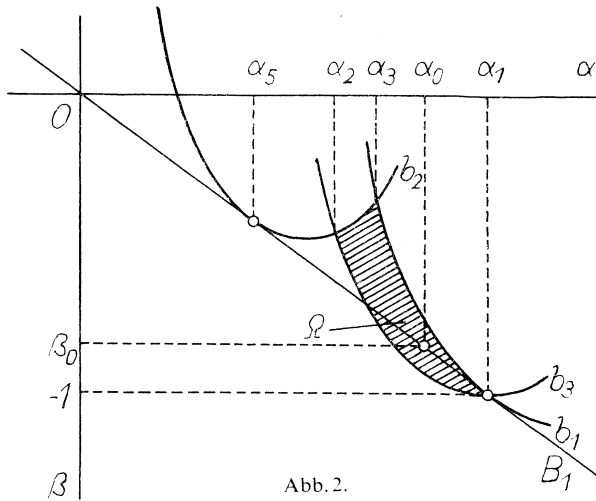


Abb. 1.

Bemerkung 1. Die Gestalt der Menge Ω für verschiedene m ist von Abb. 1, 2 und 3 sichtbar. Für den Fall II des Satzes 3 mit der scharfen Ungleichung in (6) gilt die Abb. 1. Für den Fall III des Satzes 3 mit der scharfen Ungleichung in (7) gilt die Abb. 2. Die Abb. 3 gilt für den Fall $m^2 = M^2$. Man bemerkt noch, dass die Gerade $B_1(\alpha) = -2\alpha(1 - \sqrt{(1 - M^2)})/M^2$ eine Tangente der Graphen der Funktionen $b_1(\alpha)$ bzw. $b_2(\alpha)$ mit den Berührungspunkten $\alpha = \alpha_1$ bzw. $\alpha = \alpha_5$ ist (siehe [3]).

Aus dem Satz 3 folgt, dass das optimale Parameterpaar $[\alpha_{opt}; \beta_{opt}]$ für welches $\varrho(\mathbf{T}(\alpha_{opt}, \beta_{opt})) = \min_{\alpha, \beta \in \Omega} \varrho(\mathbf{T}(\alpha, \beta))$ ist, im Inneren des Gebietes Ω liegt. Im folgenden Satz 4 wird ein gewisses Paar der Parameter $[\alpha_0, \beta_0]$ gefunden, das eine praktisch



anwendbare Approximation des Punktes $[\alpha_{\text{opt}}, \beta_{\text{opt}}]$ bietet. Durch das Symbol B_1 bezeichnet man jetzt die Menge der Punkte des Graphen der Funktion $\beta = B_1(\alpha)$.

Satz 4. Es sei $1 - \sqrt{(1 - M^2)} < m^2 \leq M^2$ und

$$(8) \quad \alpha_0 = (1 + \sqrt{(1 - M^2)})(1 - m^2)/[1 + \sqrt{(1 - M^2)} - m^2],$$

$$\beta_0 = -2(1 - m^2)/[1 + \sqrt{(1 - M^2)} - m^2].$$

Dann ist $[\alpha_0, \beta_0]$ der innere Punkt des Gebietes Ω und es gilt

$$(9) \quad \varrho(\mathbf{T}(\alpha_{\text{opt}}, \beta_{\text{opt}})) \leq \varrho(\mathbf{T}(\alpha_0, \beta_0)) = \min_{[\alpha, \beta] \in B_1} \varrho(\mathbf{T}(\alpha, \beta)) =$$

$$= \sqrt{[m^2(M^2 - m^2)](1 + \sqrt{(1 - M^2)})^2 (1 - m^2)} <$$

$$< [1 - \sqrt{(1 - M^2)}]/[1 + \sqrt{(1 - M^2)}],$$

wobei für $m^2 = M^2$ offenbar die Gleichheiten

$$\varrho(\mathbf{T}(\alpha_{\text{opt}}, \beta_{\text{opt}})) = \varrho(\mathbf{T}(\alpha_0, \beta_0)) = 0$$

gelten.

Beweis. Die Graphen der Funktionen $B_1(\alpha) = -2\alpha(1 - \sqrt{(1 - M^2)})/M^2$ und $b_3(\alpha)$ schneiden sich in den Punkten $[\alpha_1, -1]$ und $[\alpha_6, \beta_6]$, wo

$$\alpha_6 = (1 + \sqrt{(1 - M^2)})(1 - m^2)/2\sqrt{(1 - M^2)},$$

$$\beta_6 = \beta_1(\alpha_6) = -(1 - m^2)/\sqrt{(1 - M^2)}$$

und $\alpha_5 \leq \alpha_6 < \alpha_1$ gilt (für $m^2 = M^2$ ist dann $\alpha_6 = \alpha_5$, $\beta_6 = -\sqrt{(1 - M^2)}$, wobei für $\alpha_6 < \alpha < \alpha_1$ $b_3(\alpha) < B_1(\alpha)$ gilt). Man kann leicht beweisen, dass $\beta_0 = B_1(\alpha_0)$ (d. h. $[\alpha_0, \beta_0] \in B_1$) und $\alpha_6 < \alpha_0 < \alpha_1$ gilt, so dass $[\alpha_0, \beta_0] \in \text{Int } \Omega$ ist.

Man untersuche nun die die Zahl $\varrho(\mathbf{T}(\alpha, \beta))$ auf der Geraden B_1 . Es ist aus den Arbeiten [2] und [3] bekannt, dass $\varrho^2(\mathbf{T}(\alpha, \beta)) = \max_{i=1, \dots, n} |\lambda_i|^2$ ist, wo λ_i die den Eigenwerten μ_i^2 , $i = 1, \dots, n$ der Matrix \mathbf{B}^2 entsprechende Wurzeln der quadratischen Gleichung (4) sind. Es ist weiter aus der Arbeit [3] bekannt, dass der gegebenen Zahl μ_i^2 zwei komplex adjungierte Wurzeln der Gleichung (4) entsprechen, solange $\alpha > 0$ und

$$-2\alpha(1 + \sqrt{(1 - \mu_i^2)})/\mu_i^2 \leq \beta \leq -2\alpha(1 - \sqrt{(1 - \mu_i^2)})/\mu_i^2$$

ist; das Quadrat ihrer Absolutwerte gleicht dabei der Zahl

$$(10) \quad L(\mu_i) = [(1 - \alpha)^2 - \mu_i^2(1 + \beta)]/\alpha^2.$$

Falls man nun $B_2(\alpha) = -2\alpha(1 + \sqrt{(1 - M^2)})/M^2$ bezeichnet, kann man leicht fest-

stellen, dass für $\alpha_0 \leq \alpha \leq \alpha_1$ die Beziehungen

$$\begin{aligned} -2\alpha(1 + \sqrt{(1 - \mu_i^2)})/\mu_i^2 &\leq -2\alpha(1 + \sqrt{(1 - M^2)})/M^2 = B_2(\alpha) < \\ &< b_3(\alpha) \leq B_1(\alpha) = -2\alpha(1 - \sqrt{(1 - M^2)})/M^2 \leq \\ &\leq -2\alpha(1 - \sqrt{(1 - \mu_i^2)})/\mu_i^2 \end{aligned}$$

gelten. Davon folgt, dass für alle Punkte des Gebietes Ω , für die $b_3(\alpha) \leq \beta \leq B_1(\alpha)$ ist (dieses Gebiet bezeichnet man mit Ω') und für alle Zahlen μ_i^2 , $i = 1, \dots, n$ die Gleichung (4) nur komplex adjungierte Wurzeln besitzt, deren Quadrat im Absolutwert der Zahl (10) gleich ist.

Für $[\alpha, \beta] \in \Omega'$ gilt jetzt

$$(11) \quad L(\mu_i) \leq L(m)$$

angesichts dessen, dass für $[\alpha, \beta] \in \Omega'$ die Ungleichung $\beta > -1$ gilt.

Es gilt ferner

$$(12) \quad \frac{\partial}{\partial \beta} L(m) = -m^2/\alpha^2 < 0.$$

Aus (11) und (12) folgt

$$(13) \quad \min_{[\alpha, \beta] \in \Omega'} \varrho^2(\mathbf{T}(\alpha, \beta)) = \min_{[\alpha, \beta] \in B_1} L(m).$$

In Hinsicht darauf, dass für $[\alpha, \beta] \in B_1$ $\beta = B_1(\alpha) = -2\alpha(1 - \sqrt{(1 - M^2)})/M^2$ ist, bekommt man nach der Einsetzung in (10) den Ausdruck

$$l(\alpha) = [(1 - \alpha)^2 - m^2 + 2\alpha m^2(1 - \sqrt{(1 - M^2)})/M^2]/\alpha^2.$$

Aus der Beziehung

$$\frac{d}{d\alpha} l(\alpha) = 2\{\alpha[1 - m^2(1 - \sqrt{(1 - M^2)})/M^2] - (1 - m^2)\}/\alpha^3 = 0$$

bekommt man sofort $\alpha = (1 - \sqrt{(1 - M^2)})(1 - m^2)/[1 - \sqrt{(1 - M^2)} - m^2] = \alpha_0$ und $\beta = -2(1 - m^2)/[1 + \sqrt{(1 - M^2)} - m^2] = \beta_0$ angesichts $\beta = B_1(\alpha)$.

Es ist also

$$(14) \quad \begin{aligned} \min_{[\alpha, \beta] \in B_1} L(m) &= [(1 - \alpha_0)^2 - m^2(1 + \beta_0)]/\alpha_0^2 = \\ &= m^2(M^2 - m^2)/(1 - \sqrt{(1 - M^2)})^2(1 - m^2). \end{aligned}$$

Nachdem $\Omega' \subset \Omega$ ist und (13), (14) gilt, bekommt man sofort die Behauptung (9)

des Satzes 4 (die Ungleichung

$$\begin{aligned} & m^2(M^2 - m^2)/(1 - \sqrt{(1 - M^2)})^2 (1 - m^2) < \\ & < [1 - \sqrt{(1 - M^2)}]^2/[1 + \sqrt{(1 - M^2)}]^2 \end{aligned}$$

kann man dann leicht beglaubigen).

Bemerkung 2. In dem Satz 4 haben wir vorausgesetzt, dass $1 - \sqrt{(1 - M^2)} < m^2 \leq M^2$ ist. Solange $m^2 \leq 1 - \sqrt{(1 - M^2)}$ ist, folgt von dem Satz 3, dass das Gebiet Ω nur den Punkt $[\frac{1}{2}(1 + \sqrt{(1 - M^2)}), -1]$ enthält, wobei

$$\varrho(\mathbf{T}(\frac{1}{2}(1 + \sqrt{(1 - M^2)}), -1)) = [1 - \sqrt{(1 - M^2)}]/[1 + \sqrt{(1 - M^2)}]$$

ist; die Konvergenzgeschwindigkeit des untersuchten Verfahrens gleicht in diesem Fall der Konvergenzgeschwindigkeit des Oberrelaxationsverfahrens (siehe Satz 2 aus [3]).

Solange $m^2 = M^2$ ist, ist $[\alpha_0, \beta_0] = [\sqrt{(1 - M^2)}, -2\sqrt{(1 - M^2)}/(1 + \sqrt{(1 - M^2)})]$ und es gilt $[\alpha_0, \beta_0] = [\alpha_{\text{opt}}, \beta_{\text{opt}}]$, $\varrho(\mathbf{T}(\alpha_0, \beta_0)) = 0$ (vergl. Satz 5 aus [3]).

Falls $1 - \sqrt{(1 - M^2)} < m^2 < M^2$ ist, gilt für das im Satz 4 gesuchte Paar $[\alpha_0, \beta_0]$ die Ungleichung

$$\varrho(\mathbf{T}(\alpha_0, \beta_0)) < [1 - \sqrt{(1 - M^2)}]/[1 + \sqrt{(1 - M^2)}]$$

und das untersuchte Verfahren gibt eine schnellere Konvergenz als das Oberrelaxationsverfahren. Den Punkt $[\alpha_0, \beta_0]$ kann man also für eine gewisse praktisch anwendbare Approximation des Punktes $[\alpha_{\text{opt}}, \beta_{\text{opt}}]$ halten.

Zum Schluss bemerken wir noch, dass die Ableitung der expliziten Formel für $[\alpha_{\text{opt}}, \beta_{\text{opt}}]$ im Falle $1 - \sqrt{(1 - M^2)} < m^2 < M^2$ sehr schwierig ist, da die Aufgabe zur Lösung komplizierter algebraischer Gleichungen höheren als zweiten Grades führt. Bezeichnet man nämlich mit $L_1(\mu_i)$, $L_2(\mu_i)$ die Quadrate der reellen Wurzeln der Gleichung (4), die der Zahl μ_i^2 entsprechen, d. h.

$$L_1(\mu_i) = [-2(1 - \alpha)\alpha - \mu_i^2\beta + \sqrt{(\mu_i^4\beta^2 + 4\alpha\beta\mu_i^2 + 4\alpha^2\mu_i^2)}]^2/\alpha^4,$$

$$L_2(\mu_i) = [-2(1 - \alpha)\alpha - \mu_i^2\beta - \sqrt{(\mu_i^4\beta^2 + 4\alpha\beta\mu_i^2 + 4\alpha^2\mu_i^2)}]^2/\alpha^4,$$

kann man aus dem Verlauf der Funktionen $L_1(\mu_i)$, $L_2(\mu_i)$, $L(\mu_i)$ in der Menge Ω beweisen, dass die Zahl $\varrho(\mathbf{T}(\alpha, \beta))$ ihren Minimalwert an der Kurve C annimmt, die folgendermassen definiert ist:

$$[\alpha, \beta] \in C \Leftrightarrow \max(L_1(M), L_2(M)) = L(M).$$

Aus dem Satz über die gebundene Extrempunkte folgt dann, dass man den Punkt $[\alpha_{\text{opt}}, \beta_{\text{opt}}]$ als einen im Gebiet Ω liegenden Schnittpunkt der Kurve C mit dem

Graphen der Funktion

$$\beta = C(\alpha) = \{M^2(1 - \alpha - m^2) - \alpha^2 m^4\} / m^2 \{M^2(1 - \alpha - m^2) + \alpha m^2\}$$

bekommen kann; es gilt dabei, dass $C(\alpha) \geq B_1(\alpha)$ ist und dass die Gerade B_1 die Tangente der Kurve C mit dem Berührungspunkt $[\alpha_0, \beta_0]$ ist.

Literaturverzeichnis

- [1] Young, D. M.: Iterative Solution of large linear Systems. Academic Press, 1971.
- [2] Šisler, M.: Über ein zweiparametriges Iterationsverfahren. Aplikace matematiky, 18, 1973, 325—332.
- [3] Šisler, M.: Über die Optimierung eines zweiparametrigen Iterationsverfahrens. Aplikace matematiky, 20, 1975, 126—142.

Souhrn

POZNÁMKY K OPTIMALISACI JEDNÉ DVOJPARAMETRICKÉ ITERAČNÍ METODY

MIROSLAV ŠISLER

Práce se zabývá jistou iterační metodou pro řešení soustavy lineárních rovnic tvaru $\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{b}$ s 2-cyklickou maticí. Iterační metoda je dána předpisem $\mathbf{x}_{v+1} = \mathbf{T}(\alpha, \beta) \mathbf{x}_v + \mathbf{P}(\alpha, \beta) \mathbf{b}$, kde $\mathbf{T}(\alpha, \beta)$, $\mathbf{P}(\alpha, \beta)$ jsou jisté matice závislé na matici \mathbf{B} a dvou číselných parametrech α, β . Vzhledem k tomu, že uvažovaná metoda je zobecněním řady běžných iteračních metod, včetně metody superrelaxační, jsou v práci udány explicitní formule pro jisté hodnoty parametrů α, β , pro něž metoda konverguje rychleji než metoda superrelaxační.

Anschrift des Verfassers: RNDr. Miroslav Šisler, CSc., Matematický ústav ČSAV, Žitná 25, 115 67 Praha 1.