Iterative Lösung Linearer Gleichungssysteme

E. Olszewski, H. Röck, M. Watzl

31. Jänner 2002

Vorwort

C.F.Gauß in einem Brief vom 26.12.1823 an Gerling:

$<<$ Ich empfehle Ihnen diesen Modus zur Nachahmung. Schwerlich werden Sie je wieder direct elimineren, wenigstens nicht wenn Sie mehr als 2 Unbekannte haben. Das indirecte Verfahren lässt sich halb im Schlafe ausführen, oder man kann während desselben an andere Dinge denken.$>>$

Was ist iteratives Lösen linearer Gleichungssysteme?

- Iteratives Lösen linearer Gleichungssysteme ist ein Approximationsverfahren zum bestimmen der Lösung eines Gleichungssystems.
- Bisher uns bekannte Verfahren zum Lösen linearer Gleichungssysteme: Direkte Verfahren (z.B: Gaußsches Eliminationsverfahren).
- Bei den direkten Verfahren kommt man nach endlich vielen Schritten zur ”genauen” Lösung des Gleichungssystems.
- Bei den iterativen Verfahren nähert man sich in jedem Schritt der Lösung.

Historische Bemerkungen zu Iterationsverafhren

- Iterationsverfahren sind ca. 180 Jahre alt. Das erste Iterationsverfahren für lineare Gleichungssysteme stammt von Carl Friedrich Gauß.

- Während seinen Berechnungen stieß Gauß immer wieder auf Gleichungssysteme die zu groß waren, als daß er sie mit der direkten Methode der Gauß-Elimination berechnen konnte. Darum entwickelte er um 1820 das erste Iterationsverfahren: Das Gauß-Seidel Verfahren (auch bekannt als Einzelschrittverfahren).

- Ein sehr ähnliches Verfahren entwickelte Carl Gustav Jacobi 1845. Das sogenannte Jacobi Verfahren (auch bekannt als Gesamtschrittverfahren).

Historische Bemerkungen zu Iterationsverafhren (cont'd)

- Durch den Einzug von elektronischen Rechnern stieg die Anzahl der Gleichungen um eine weitere Größenordnung, und die oben genannten Verfahren erwiesen sich als zu langsam. Nach 100 Jahren Stillstand in der Entwicklung der iterativen Verfahren wurde das Gauß-Seidel Verfahren von Southwell etwas verbessert (Relaxationsverfahren).

- 1950 gelang Young ein Durchbruch mit einer wesentlichen Verbesserung (Beschleunigung) des Gauß-Seidel Verfahrens. Dieses Verfahren wird SOR-Methode genannt (Successive Overrelaxation Method). In der Folgezeit entstanden noch viele weitere noch effektivere Verfahren. z.B.: Verfahren der konjugierten Gradienten, Mehrgitterverfahren,etc.

Warum wurden Iterationsverfahren entwickelt?

- In Zeiten ohne Rechner war es so gut wie unmöglich, mit dem direkten Verfahren größere Gleichungssysteme zu lösen, da diese enorm viel Rechenaufwand erfordern.

- Es war also notwendig ein Verfahren zu entwickeln, mit dem man schneller an die Lösung kam.

- Direkte Verfahren sind zwar für jede Dimension anwendbar, der benötigte Rechenaufwand ($(O)n^3$) steigt aber mit der Dimension so stark an, dass man zur Lösung von 10^3, 10^4 oder noch größeren Gleichungen so manchen Rechner ins schwitzen bringt.

- Ein wichtiges Beispiel für das Auftreten großer Gleichungssysteme ist die Diskretisierung partieller Differentialgleichungen. In diesem Fall sind die Matrizen schwach besetzt (d.h sie enthalten viele Nullen), und eingnen sich besonders gut zur iterativen Lösung.

Vorteile von Iterationsverfahren gegenüber direkter Verfahren

- Je größer die Dimension des Gleichungssystems, desto schneller ist das Iterationsverfahren im Vergleich zu den direkten Verfahren.

- Schon nach wenigen Schritten erhält man eine ziemlich genaue Lösung.

- Durch die Rundungsfehler im Rechner liefert das Iterationsverfahren sogar meist genauere Ergebnisse als das direkte Verfahren. Darum wurde auch die Methode der Nachiteration entwickelt (auf das Ergebnis des direkten Verfahrens wird ein iteratives Verfahren angewandt).

- Bei den direkten Verfahren zur Lösung eines linearen Gleichungssystems $Ax = b$ wird die Koeffizientenmatrix A im Laufe der Rechnung verändert. Ist die Dimension von A sehr gross, so ist A im allgemeinen nicht im Hauptspeicher darstellbar.

- $\Rightarrow$ Direkte Verfahren oft nicht mehr möglich bzw. sehr aufwendige Datentransfers nötig.
  Iterative Verfahren verändern nicht die Koeffizientenmatrix, sondern berechnen sukkzessiv Näherungsvektoren für die Lösung und benötigen daher weniger Speicherplatz.

Nachteile von Iterationsverfahren gegenüber direkter Verfahren

- Es stellt sich natürlich die Frage ob iterative Verfahren für alle Gleichungssysteme eine Lösung bringen. (d.h: ob sie konvergieren).
- Dies festzustellen ist keine leichte Aufgabe und erfordert Rechenaufwand.
- Falls das Vefahren nicht konvergiert kann es natürlich nicht angewandt werden.In diesen Fällen muss man dann doch wieder auf das altbewährte direkte Verfahren zurückgreifen, dass ja nach endlich vielen Schritten terminiert.

Allgemeines Iterationsverfahren

- allgemeine Form eines linearen Gleichungssystems

$$Ax = y \tag{1}$$

speziell für:

⋆ A eine reguläre (n,n) Matrix
⋆ $x, y \in V_n(\mathbb{R})$

- durch Zerlegung von $A = N - P$ herleiten in die Form

$$x = Cx + d \tag{2}$$

wobei gilt:

⋆ C eine reguläre (n,n) Matrix
⋆ $x, d \in V_n(\mathbb{R})$

Allgemeines Iterationsverfahren

- durch die Herleitung erhalten wir

$$\begin{aligned} C &= E - N^{-1}A \\ d &= N^{-1}y \end{aligned}$$

- aus obiger Gleichung erhält man folgende Iterationsvorschrift

$$x_{n+1} = Cx_n + d \quad (3)$$

nach dem BANACHschen Fixpunktsatz folgt die Konvergenz für jeden beliebigen Startvektor x_0, sofern $||C|| < 1$ gilt.

- Die verschiedenen Verfahren unterscheiden sich nur durch die Wahl von N

Zerlegung

- Zur Beschreibung der einzelnen Verfahren verwenden wir folgende Zerlegung

$$A = A_L + D + A_R$$

$$\text{mit } A_L = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ a_{21} & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{pmatrix}, \; D = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

$$\text{und } A_R = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

- Durch die Wahl von N lassen sich folgende Iterationsverfahren kennzeichnen:

$N =$	Verfahren
D	Gesamtschrittverfahren(Jacobiverfahren)
$D + A_L$	Einzelschrittverfahren(Gauß-Seidel-Verfahren)
$\frac{1}{\omega}D + A_L$	Relaxationsverfahren

Gesamtschrittverfahren

- wurde 1845 von Carl Gustav Jacobi entwickelt
- wird beschrieben durch die Gleichung:

$$x_{n+1} = -D^{-1}(A_L + A_R)x_n + D^{-1}y \tag{4}$$

wobei gilt:

⋆ Diagonalelemente von A sind ungleich 0 damit D^{-1} existiert

Algorithmus

- Gegeben sei das Gleichungssystem $Ax = y$ mit einer (p,p)-Matrix A und $y \in V(K)$. Die Matrix A erfülle eine der hinreichenden Bedingungen für die Konvergenz des Gesamtschrittverfahrens

1. **Schritt:** Bringe das Gleichungssystem auf die Form Gl.4 dh.

$$c_{ij} := \begin{cases} 0 & i = j \\ -\frac{a_{ij}}{a_{ii}} & i \neq j \end{cases} \tag{5}$$

$$b_i := \frac{y_i}{a_{ii}} \tag{6}$$

2. **Schritt:** Wähle einen Startvektor x_0, o.B.d.A $x_0 = 0$. Wähle eine Fehlerschranke $\varepsilon > 0$ sowie eine maximale Schrittzahl N.
3. **Schritt:** Berechne ausgehend vom Startvektor x_0, für $n = 1, 2, \ldots$

$$x_{i,n} = \sum_{j=1}^{p} c_{ij} x_{j,n-1} + b_i \tag{7}$$

4. **Schritt:** Es wird solange iteriert, bis eines der Abbruchkriterien erfüllt ist.

Einzelschrittverfahren

- wurde von Gauß um 1820 entwickelt
- lässt sich einfach vom Gesamtschrittverfahren ableiten
- beim Gesamtschrittverfahren wird zunächst $x_{1,n}$ berechnet, danach $x_{2,n}$ mit dem alten Wert von $x_{1,n-1}$ usw.
- obwohl bei der Berechnung von $x_{2,n}$ bereits der bessere Näherungswert von $x_{1,n}$ verwendet werden könnte
- nutzt man die verbesserten Werte systematisch, sobald sie zur Verfügung stehen, kommt man zum Einzelschrittverfahren.
- Entscheidende Verbesserung zum vorherigen Algorithmus. Anstelle von Gl.7 wird folgende verwendet:

$$x_{i,n} := \sum_{j=1}^{i-1} c_{ij} x_{j,n} + \sum_{j=i+1}^{p} c_{ij} x_{j,n-1} + b_i \tag{8}$$

Rechenbeispiel

- Lineares Gleichungssystem:

$$\begin{array}{rcrcrcr} 4x_1 & + & x_2 & + & 2x_3 & = & 7 \\ 2x_1 & + & 4x_2 & + & x_3 & = & 7 \\ & & x_2 & + & 2x_3 & = & 3 \end{array}$$

- In Matrizenform $Ax = b$:

$$\begin{pmatrix} 4 & 1 & 2 \\ 2 & 4 & 1 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 3 \end{pmatrix}.$$

- Matrizen, die beim Iterativen Vorgehen benötigt werden:

$$C = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}, \qquad d = \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}, \qquad x_0 = 0.$$

Berechnung mit Jacobi-Verfahren

- Erste Iteration: $x_1 = Cx_0 + d$

$$x_1 = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}$$

- Zweite Iteration: $x_2 = Cx_1 + d$

$$x_2 = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}$$

$$x_2 = \begin{pmatrix} -\frac{19}{16} \\ -\frac{10}{8} \\ -\frac{7}{8} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{9}{16} \\ \frac{1}{2} \\ \frac{5}{8} \end{pmatrix}$$

Berechnung mit Jacobi-Verfahren (cont'd)

- Dritte Iteration: $x_3 = Cx_2 + d$

$$x_3 = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{9}{16} \\ \frac{1}{2} \\ \frac{5}{8} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}$$

$$x_3 = \begin{pmatrix} -\frac{7}{16} \\ -\frac{7}{16} \\ -\frac{1}{4} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{21}{16} \\ \frac{21}{16} \\ \frac{5}{4} \end{pmatrix}$$

- Vierte Iteration: $x_4 = Cx_3 + d$

$$x_4 = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{21}{16} \\ \frac{21}{16} \\ \frac{5}{4} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}$$

$$x_4 = \begin{pmatrix} -\frac{61}{64} \\ -\frac{31}{32} \\ -\frac{21}{32} \end{pmatrix} + \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{51}{64} \\ \frac{25}{32} \\ \frac{27}{32} \end{pmatrix}$$

Berechnung mit Gauß-Seidelschem Verfahren

- Die Matrizen C, d und a_0 von zuvor:

$$C = \begin{pmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}, \qquad d = \begin{pmatrix} \frac{7}{4} \\ \frac{7}{4} \\ \frac{3}{2} \end{pmatrix}, \qquad x_0 = 0.$$

- Allgemeines Vorgehen:

$$\begin{array}{rclllllll} x_{1,n} & = & & & - & \frac{1}{4}x_{2,n-1} & - & \frac{1}{2}x_{3,n-1} & + \frac{7}{4} \\ x_{2,n} & = & - & \frac{1}{2}x_{1,n} & & & - & \frac{1}{4}x_{3,n-1} & + \frac{7}{4} \\ x_{3,n} & = & & & - & \frac{1}{2}x_{2,n} & & & + \frac{3}{2} \end{array}$$

Berechnung Gauß-Seidelsches Verfahren (cont'd)

- Erste Iteration:

$$
\begin{aligned}
x_{1,1} &= -\tfrac{1}{4}x_{2,0} - \tfrac{1}{2}x_{3,0} + \tfrac{7}{4} \\
x_{2,1} &= -\tfrac{1}{2}x_{1,1} - \tfrac{1}{4}x_{3,0} + \tfrac{7}{4} \\
x_{3,1} &= -\tfrac{1}{2}x_{2,1} + \tfrac{3}{2}
\end{aligned}
$$

$$x_{1,1} = \frac{7}{4}$$

$$
\begin{aligned}
x_{2,1} &= -\frac{1}{2}\cdot\frac{7}{4} + \frac{7}{4} \\
x_{2,1} &= \frac{7}{8}
\end{aligned}
$$

$$
\begin{aligned}
x_{3,1} &= -\frac{1}{2}\cdot\frac{7}{8} + \frac{3}{2} \\
x_{3,1} &= \frac{17}{16}
\end{aligned}
$$

Berechnung Gauß-Seidelsches Verfahren (cont'd)

- Zweite Iteration:

$$
\begin{aligned}
x_{1,2} &= -\tfrac{1}{4}x_{2,1} - \tfrac{1}{2}x_{3,1} + \tfrac{7}{4}\\
x_{2,2} &= -\tfrac{1}{2}x_{1,2} - \tfrac{1}{4}x_{3,1} + \tfrac{7}{4}\\
x_{3,2} &= -\tfrac{1}{2}x_{2,2} + \tfrac{3}{2}
\end{aligned}
$$

$$
\begin{aligned}
x_{1,2} &= -\frac{1}{4}\cdot\frac{7}{8} - \frac{1}{2}\cdot\frac{17}{16} + \frac{7}{4}\\
x_{1,2} &= 1
\end{aligned}
$$

$$
\begin{aligned}
x_{2,2} &= -\frac{1}{2}\cdot - \frac{1}{4}\cdot\frac{17}{16} + \frac{7}{4}\\
x_{2,2} &= \frac{63}{64}
\end{aligned}
$$

$$
\begin{aligned}
x_{3,2} &= -\frac{1}{2}\cdot\frac{63}{64} + \frac{3}{2}\\
x_{3,2} &= \frac{129}{128}
\end{aligned}
$$