

Numerische Mathematik

Oliver Ernst

Professur Numerische Mathematik

Sommersemester 2015



TECHNISCHE UNIVERSITÄT
CHEMNITZ

① Einführung und Begriffe

- 1.1 Mathematische Modellbildung und numerische Simulation am Beispiel eines Wasserkreislaufs
- 1.2 Linearisierung und Iterationsverfahren am Beispiel des Newton-Verfahrens
- 1.3 Diskretisierung und Stabilität am Beispiel der Wärmeleitungsgleichung

② Gleitpunktarithmetik und Fehleranalyse

- 2.1 Gleitpunktzahlen
- 2.2 Rundung
- 2.3 Der IEEE-754 Standard
- 2.4 Korrekt gerundete Gleitpunktarithmetik
- 2.5 Numerische Stabilität und Fehleranalyse
- 2.6 Ein Beispiel

③ Direkte Verfahren zur Lösung linearer Gleichungssysteme

- 3.1 Vorbemerkungen
- 3.2 Störungstheorie
- 3.3 Das Lösen von Dreieckssystemen
- 3.4 Gauß-Elimination
- 3.5 Pivotisierung

- 3.6 Equilibrierung und Nachiteration
- 3.7 Stabilität bei der Gauß-Elimination

4 Direkte Verfahren für spezielle Systeme

- 4.1 Die Cholesky-Zerlegung
- 4.2 Bandmatrizen, Tridiagonalmatrizen
- 4.3 Schwach besetzte Matrizen
- 4.4 Vandermonde-Matrizen
- 4.5 Toeplitz-Matrizen

- 1 Einführung und Begriffe
- 2 Gleitpunktarithmetik und Fehleranalyse
- 3 Direkte Verfahren zur Lösung linearer Gleichungssysteme
- 4 Direkte Verfahren für spezielle Systeme

- 1 Einführung und Begriffe
- 2 Gleitpunktarithmetik und Fehleranalyse
- 3 Direkte Verfahren zur Lösung linearer Gleichungssysteme
- 4 Direkte Verfahren für spezielle Systeme
 - 4.1 Die Cholesky-Zerlegung
 - 4.2 Bandmatrizen, Tridiagonalmatrizen
 - 4.3 Schwach besetzte Matrizen
 - 4.4 Vandermonde-Matrizen
 - 4.5 Toeplitz-Matrizen

Satz 4.1

Es sei $A = [a_{i,j}] \in \mathbb{R}^{n \times n}$ [$\mathbb{C}^{n \times n}$] symmetrisch [Hermitesch].

Dann sind die folgenden Aussagen äquivalent:

- (a) A ist positiv definit, d.h. $\mathbf{x}^\top A \mathbf{x} > 0 \forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ [$\mathbf{x}^H A \mathbf{x} > 0 \forall \mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$].
- (b) Alle Eigenwerte von A sind positiv.
- (c) Alle Hauptuntermatrizen von A sind positiv definit.
- (d) Es gibt eine orthogonale Matrix $U \in \mathbb{R}^{n \times n}$ (d.h. $U^\top U = I_n$) [eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ (d.h. $U^H U = I_n$)] und positive Zahlen $\lambda_1, \lambda_2, \dots, \lambda_n$ mit

$$U^\top A U = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \quad [U^H A U = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)].$$

- (e) Für alle regulären Matrizen X ist $X^\top A X$ [$X^H A X$] positiv definit.

Direkte Verfahren für spezielle Systeme

Die Cholesky-Zerlegung

Satz 4.2

Es sei $A \in \mathbb{R}^{n \times n}$ [$\mathbb{C}^{n \times n}$] symmetrisch [Hermitesch]. A ist genau dann positiv definit, wenn es eine reguläre untere Dreiecksmatrix $L \in \mathbb{R}^{n \times n}$ [$\mathbb{C}^{n \times n}$] gibt mit

$$A = LL^T \quad [A = LL^H] \quad (\text{Cholesky-Zerlegung}).$$

Bemerkung 4.3

Statt der in Satz 4.2 beschriebenen Cholesky-Zerlegung wird oft die **rationale** (oder **wurzelfreie**) Cholesky-Zerlegung

$$A = L_1 D L_1^T \quad [A = L_1 D L_1^H]$$

verwendet. Hier ist L_1 eine normierte untere Δ -Matrix und D eine Diagonalmatrix mit positiven Hauptdiagonalelementen. Mit $\Delta := \text{diag}(L)$ gilt

$$L_1 = L \Delta^{-1} \quad \text{und} \quad D = \Delta^2.$$

Direkte Verfahren für spezielle Systeme

Die Cholesky-Zerlegung

Pseudocode: Cholesky-Zerlegung

for $j = 1 : n$ **do**

$$\ell_{j,j} := \left(a_{j,j} - \sum_{k=1}^{j-1} \ell_{j,k}^2 \right)^{1/2}$$

for $i = j + 1 : n$ **do**

$$\ell_{i,j} := \left(a_{i,j} - \sum_{k=1}^{j-1} \ell_{i,k} \ell_{j,k} \right) / \ell_{j,j}$$

end for

end for

Aufwand: $\frac{1}{3}n^3 + O(n^2)$ flops.

Bemerkung: Der untere Δ -Anteil von A kann mit L überschrieben werden. Der (echte) obere Δ -Anteil von A wird im Algorithmus nicht verwendet.

Direkte Verfahren für spezielle Systeme

Die Cholesky-Zerlegung

Lösung eines linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$, A s.p.d. [H.p.d.] durch Cholesky-Zerlegung:

- 1: Berechne Cholesky-Zerlegung $A = LL^T$
- 2: Löse $L\mathbf{y} = \mathbf{b}$
- 3: Löse $L^T\mathbf{x} = \mathbf{y}$

Dieses Verfahren ist rückwärtsstabil:

Satz 4.4

Die in Gleitpunktarithmetik mit Rundungseinheit u durch Cholesky-Zerlegung berechnete Lösung $\tilde{\mathbf{x}}$ des linearen Gleichungssystems $A\mathbf{x} = \mathbf{b}$ mit einer symmetrisch positiv-definiten Matrix A genügt einer Gleichung

$$(A + \Delta A)\tilde{\mathbf{x}} = \mathbf{b} \quad \text{mit} \quad \|\Delta A\|_\infty \leq 3n^2u\|A\|_\infty.$$

- 1 Einführung und Begriffe
- 2 Gleitpunktarithmetik und Fehleranalyse
- 3 Direkte Verfahren zur Lösung linearer Gleichungssysteme
- 4 Direkte Verfahren für spezielle Systeme
 - 4.1 Die Cholesky-Zerlegung
 - 4.2 Bandmatrizen, Tridiagonalmatrizen
 - 4.3 Schwach besetzte Matrizen
 - 4.4 Vandermonde-Matrizen
 - 4.5 Toeplitz-Matrizen

Direkte Verfahren für spezielle Systeme

Bandmatrizen, Tridiagonalmatrizen

Man nennt $A = [a_{i,j}] \in \mathbb{C}^{n \times n}$ eine **Bandmatrix** mit **unterer Bandbreite** b_L und **oberer Bandbreite** b_R , falls $a_{i,j} = 0$ für $i - j > b_L$ und für $j - i > b_R$ gilt.

Eine Bandmatrix T mit $b_L = b_R = 1$ heißt **Tridiagonalmatrix**:

$$T = \text{tridiag}(\mathbf{b}, \mathbf{a}, \mathbf{c}) = \begin{bmatrix} a_1 & c_1 & & & & \\ b_2 & a_2 & c_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & b_{n-1} & a_{n-1} & c_{n-1} \\ & & & & b_n & a_n \end{bmatrix}.$$

Direkte Verfahren für spezielle Systeme

Bandmatrizen, Tridiagonalmatrizen

Besitzt die Tridiagonalmatrix T eine LR-Zerlegung, dann hat diese die Form

$$T = \begin{bmatrix} 1 & & & & & \\ \ell_2 & 1 & & & & \\ & \ddots & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & 1 & \\ & & & & \ell_n & 1 \end{bmatrix} \begin{bmatrix} r_1 & c_1 & & & & \\ & r_2 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & r_{n-1} & c_{n-1} \\ & & & & & r_n \end{bmatrix}$$

mit

$$r_1 = a_1 \text{ sowie } \ell_j = b_j/r_{j-1}, \quad r_j = a_j - \ell_j c_{j-1}, \quad j = 2, 3, \dots, n.$$

Ist Gauß-Elimination ohne Pivotsuche für $T\mathbf{x} = \mathbf{d}$ durchführbar, so erfordert sie also einen Aufwand von nur $3(n-1)$ Gleitpunktoperationen.

- Analoge Eigenschaften besitzt die Gauß-Elimination bei allgemeinen Bandmatrizen mit unterer Bandbreite b_L und oberer Bandbreite b_R :
- Der Aufwand zur Berechnung der LR-Zerlegung beträgt (etwa) $(2b_L b_R + 1)n$ Flops. Wird nicht pivotisiert, so entsteht kein Fill-in, d.h. L besitzt untere Bandbreite b_L und R besitzt obere Bandbreite b_R .
- Bei Spaltenpivotsuche verändert sich die Struktur von L nicht, R allerdings besitzt i. Allg. obere Bandbreite $b_L + b_R$.
- Bei weniger strukturierten schwach besetzten Matrizen werden Umordnungsstrategien (z.B. reverse Cuthill-McKee) zur Bandbreitenminimierung eingesetzt (vgl. den nächsten Abschnitt).

- 1 Einführung und Begriffe
- 2 Gleitpunktarithmetik und Fehleranalyse
- 3 Direkte Verfahren zur Lösung linearer Gleichungssysteme
- 4 Direkte Verfahren für spezielle Systeme
 - 4.1 Die Cholesky-Zerlegung
 - 4.2 Bandmatrizen, Tridiagonalmatrizen
 - 4.3 Schwach besetzte Matrizen
 - 4.4 Vandermonde-Matrizen
 - 4.5 Toeplitz-Matrizen

- **Schwach besetzte Matrizen** sind (üblicherweise sehr große) Matrizen, bei denen viele Einträge den Wert Null besitzen,
- oder pragmatischer: Matrizen mit so vielen Nulleinträgen, dass es sinnvoll ist, Algorithmen einzusetzen, die weder diese Nulleinträge speichern noch auf ihnen operieren.
- Wieviel Fill-in bei der LR-Zerlegung schwach besetzter Matrizen auftritt hängt entscheidend von der Nummerierung der Zeilen und Spalten ab: Für

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

liefert Gauß-Elimination (mit Spaltenpivotsuche)

Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen

$$PA = \begin{bmatrix} 1.0 & & & & \\ 1.0 & 1.0 & & & \\ 1.0 & & 1.0 & & \\ 1.0 & 1.0 & 1.0 & 1.0 & \\ 1.0 & 1.0 & 1.0 & 0.5 & 1.0 \end{bmatrix} \begin{bmatrix} 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ & -1.0 & & -1.0 & -1.0 \\ & & -1.0 & -1.0 & -1.0 \\ & & & 2.0 & 1.0 \\ & & & & 1.5 \end{bmatrix},$$

also eine (nahezu) voll besetzte LR-Zerlegung.
(P entspricht der Permutation (1 3 2 4 5).)

Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen

Vertauscht man in A die erste und letzte Zeile sowie die erste und letzte Spalte, so ergibt sich für $B = A([5\ 2\ 3\ 4\ 1], [5\ 2\ 3\ 4\ 1])$ eine viel schwächer besetzte LR-Zerlegung,

$$B = \begin{bmatrix} 1.0 & & & & \\ & 1.0 & & & \\ & & 1.0 & & \\ & & & 1.0 & \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} 1.0 & & & & 1.0 \\ & 1.0 & & & 1.0 \\ & & 1.0 & & 1.0 \\ & & & 1.0 & 1.0 \\ & & & & -3.0 \end{bmatrix},$$

also keinerlei Fill-in.

Wir beschreiben mit der **umgekehrten Cuthill-McKee-Nummerierung** eine Heuristik, die sich bei symmetrischen Problemen bewährt hat. Ziel ist es, die Zeilen und Spalten von A so umzunummerieren, dass die Bandbreite möglichst klein wird.

Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen

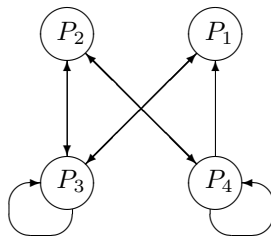
- Wie benötigen den Begriff des **gerichteten Graphen einer Matrix** $A \in \mathbb{C}^{n \times n}$ (vorläufig beliebig).
- Der Graph $G(A) = (N, V)$ von A besteht aus $n = \dim(A)$ **Knoten** P_1, P_2, \dots, P_n , die zusammen die **Knotenmenge** N bilden.
- Für jedes Element $a_{i,j} \neq 0$ wird P_i mit P_j durch eine **gerichtete Kante** verbunden (von P_i nach P_j).
- Falls $a_{i,i} \neq 0$ ist, so wird P_i mit einer geschlossenen Schleife versehen. Die Gesamtheit aller Kanten bildet die **Kantenmenge** V . (Bei symmetrischen Matrizen oder allgemeiner bei Matrizen mit symmetrischer Besetzungsstruktur kann man auf die Orientierung der Kanten natürlich verzichten.)
- Der **Grad** d_i eines Knotens P_i ist durch $d_i := |\{j : a_{i,j} \neq 0\}|$ definiert.

Beispiel.

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

\Rightarrow

$G(A) :$



$$d_1 = 1, d_2 = 2, d_3 = d_4 = 3.$$

Cuthill-McKee-Nummerierung (CMK) für symmetrisches $A \in \mathbb{R}^{n \times n}$:

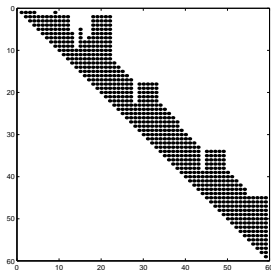
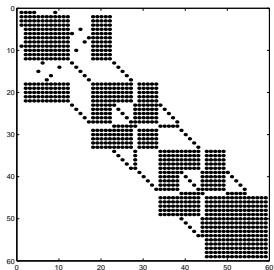
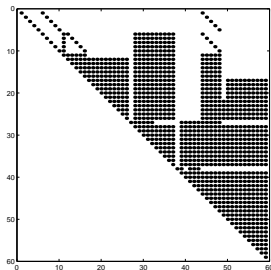
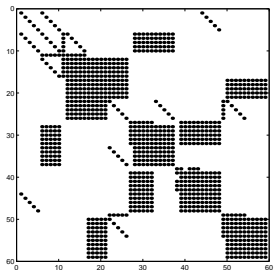
- 1: Wähle (beliebigen) Knoten aus $G(A)$ als Knoten P_1
- 2: **for** $i = 1, 2, \dots, n - 1$ **do**
- 3: Nummeriere die (noch unnummerierten) Nachbarn von P_i nach aufsteigendem Grad
- 4: Entferne alle Kanten zwischen bereits nummerierten Knoten und bestimme die Grade der noch unnummerierten Knoten bez. dieses neuen Graphen.
- 5: **end for**

Bemerkungen.

- (1) Diese Nummerierung ist nur möglich, wenn $G(A)$ **zusammenhängend** ist, d.h. wenn jeder Knoten von jedem anderen Knoten aus in $G(A)$ erreichbar ist.
- (2) Ist $[\pi(1), \pi(2), \dots, \pi(n)]$ die CMK-Nummerierung, so ist die **umgekehrte Cuthill-McKee-Nummerierung** durch $[\pi(n), \pi(n - 1), \dots, \pi(1)]$ definiert.

Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen



Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen

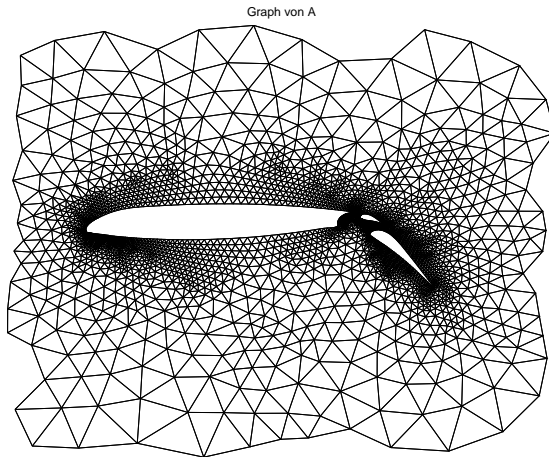
Legende.

- Links oben wird die Struktur der symmetrischen positiv definiten Matrix $A = CC^T$ (auf die Bedeutung dieser Matrix wird später eingegangen) gezeigt, wobei die nichtsymmetrische Matrix $C \in \mathbb{R}^{59 \times 59}$ aus der Modellierung eines chemischen Betriebs stammt (zu Details siehe <http://math.nist.gov:80/MatrixMarket/data/Harwell-Boeing/chemimp/impcolb.html>).
- Rechts oben ist das Besetzungsmuster des Cholesky-Faktors L_A^T von A dargestellt.

Sei P die Permutationsmatrix, die zur umgekehrten Cuthill-McKee-Nummerierung gehört, dann ist $B = PAP^T$ (links unten) und L_B^T der zugehörige Cholesky-Faktor.

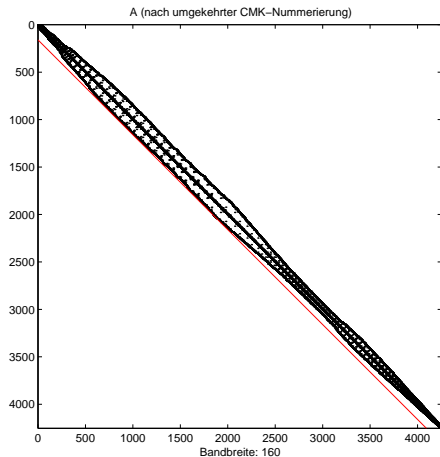
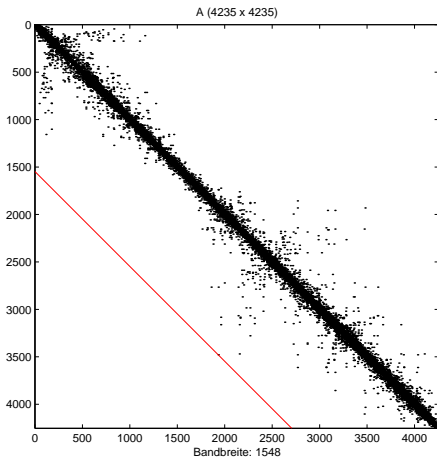
| | A | L_A | B | L_B |
|---|-----------|------------|-----------|-----------|
| # entries (i, j) $\neq 0$ ($i \leq j$) | 503 (28%) | 1073 (61%) | 503 (28%) | 657 (37%) |

Weiteres Beispiel: Finite-Element Modell eines Tragflächenprofils.



Direkte Verfahren für spezielle Systeme

Schwach besetzte Matrizen



- ① Einführung und Begriffe
- ② Gleitpunktarithmetik und Fehleranalyse
- ③ Direkte Verfahren zur Lösung linearer Gleichungssysteme
- ④ Direkte Verfahren für spezielle Systeme
 - 4.1 Die Cholesky-Zerlegung
 - 4.2 Bandmatrizen, Tridiagonalmatrizen
 - 4.3 Schwach besetzte Matrizen
 - 4.4 **Vandermonde-Matrizen**
 - 4.5 Toeplitz-Matrizen

Eine Matrix der Bauart

$$V = V(x_0, x_1, x_2, \dots, x_n) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_0 & x_1 & x_2 & & x_n \\ x_0^2 & x_1^2 & x_2^2 & & x_n^2 \\ \vdots & & & \ddots & \vdots \\ x_0^n & x_1^n & x_2^n & \cdots & x_n^n \end{bmatrix} \in \mathbb{C}^{(n+1) \times (n+1)}$$

heißt **Vandermonde-Matrix**. Es gilt

$$\det(V(x_0, x_1, \dots, x_n)) = \prod_{i>j} (x_i - x_j).$$

Insbesondere ist V genau dann regulär, wenn die x_j paarweise verschieden sind.

Direkte Verfahren für spezielle Systeme

Vandermonde-Matrizen

V ist zwar voll besetzt, besitzt aber eine sehr spezielle Struktur. Man erwartet, dass diese ausgenutzt werden kann zur Konstruktion von Algorithmen, die LGSe mit Koeffizientenmatrix V bzw. LGSe der Form

$$V^T \mathbf{a} = \mathbf{f} \quad \left(V \in \mathbb{R}^{(n+1) \times (n+1)}, \mathbf{f} \in \mathbb{R}^{n+1} \right) \quad (\text{LGS})$$

mit deutlich weniger als $O(n^3)$ Flops lösen.

Wir betrachten hierzu folgende **polynomiale Interpolationsaufgabe**: Zu gegebenen (paarweise verschiedenen) Knoten $x_0, x_1, \dots, x_n \in \mathbb{R}$ und gegebenen Funktionswerten $f_0, f_1, \dots, f_n \in \mathbb{R}$ soll ein **Interpolationspolynom** vom Grad n

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in \mathcal{P}_n$$

mit reellen Koeffizienten a_0, a_1, \dots, a_n (also $n + 1$ Freiheitsgrade) konstruiert werden, das die $n + 1$ Interpolationsbedingungen

$$p(x_i) = f_i, \quad i = 0, 1, \dots, n. \quad (\text{IP})$$

erfüllt.

Satz 4.5

Die Aufgaben (LGS) und (IP) sind äquivalent: Der Vektor

$$\mathbf{a} = [a_0, a_1, \dots, a_n]^\top \in \mathbb{R}^{n+1} \quad \text{löst das LGS} \quad V^\top \mathbf{a} = \mathbf{f}$$

genau dann, wenn das Polynom

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \in \mathcal{P}_n$$

die Interpolationsbedingungen (IP) erfüllt.

Bemerkung 4.6

Damit ist unsere Interpolationsaufgabe genau dann eindeutig lösbar, wenn die Knoten $\{x_i\}_{i=0}^n$ paarweise verschieden sind, d.h. es gibt in diesem Fall genau ein Polynom $p \in \mathcal{P}_n$, das den Bedingungen (IP) genügt.

Direkte Verfahren für spezielle Systeme

Vandermonde-Matrizen

Wir lösen $V^T \mathbf{a} = \mathbf{f}$ dadurch, dass wir die Koeffizienten des Interpolationspolynoms p in **Newton-Form** berechnen. Letztere verwendet die **dividierten Differenzen** der Wertepaare $\{(x_i, f_i)\}_{i=0}^n$, welche wie folgt definiert sind:

Sind $i_0, i_1, \dots, i_k \in \{0, 1, \dots, n\}$ paarweise verschieden, so setzen wir

$$f_{i_0, i_1, \dots, i_k} := \frac{f_{i_1, i_2, \dots, i_k} - f_{i_0, i_1, \dots, i_{k-1}}}{x_{i_k} - x_{i_0}} \quad (k \geq 1).$$

Satz 4.7 (Newton, 1669)

Bezüglich der dividierten Differenzen lässt sich das Interpolationspolynom p in **Newton-Form**

$$\begin{aligned} p(x) &= f_0 + f_{0,1}(x - x_0) + f_{0,1,2}(x - x_0)(x - x_1) \\ &\quad + \dots \\ &\quad + f_{0,1,\dots,n}(x - x_0)(x - x_1) \cdots (x - x_{n-1}) \end{aligned}$$

darstellen.

Direkte Verfahren für spezielle Systeme

Vandermonde-Matrizen

Die dividierten Differenzen der Wertepaare $\{(x_i, f_i)\}_{i=0}^n$ kann man sukzessive durch spaltenweises Auffüllen des **Newton-Tableaus** generieren. Für $n = 4$ erhält man beispielsweise

| x_i | $k = 0$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ |
|-------|---------|-----------|-------------|---------------|-----------------|
| x_0 | f_0 | | | | |
| x_1 | f_1 | $f_{0,1}$ | $f_{0,1,2}$ | | |
| x_2 | f_2 | $f_{1,2}$ | $f_{1,2,3}$ | $f_{0,1,2,3}$ | |
| x_3 | f_3 | $f_{2,3}$ | $f_{2,3,4}$ | $f_{1,2,3,4}$ | $f_{0,1,2,3,4}$ |
| x_4 | f_4 | $f_{3,4}$ | | | |

Folgender Algorithmus berechnet die dividierten Differenzen:

Gegeben: Wertepaare $\{x_i, f_i\}_{i=0}^n$, x_i paarweise verschieden

```
for  $k = 0 : n - 1$  do  
  for  $j = n : -1 : k + 1$  do  
     $f_j := (f_j - f_{j-1}) / (x_j - x_{j-k-1})$   
  end for  
end for
```

- Die im weiteren Verlauf nicht mehr benötigten Differenzen werden im obigen Algorithmus überschrieben. Dadurch wird kein zusätzlicher Speicherplatz benötigt.
- Am Ende stehen die dividierten Differenzen $f_0, f_{0,1}, f_{0,1,\dots,n}$ in den Variablen f_0, f_1, \dots, f_n .
- **Aufwand:** $\frac{3}{2}(n^2 + n)$ Flops.

Um aus der Newton-Darstellung

$$p(x) = \sum_{k=0}^n f_{0,1,\dots,k} \prod_{j=0}^{k-1} (x - x_j)$$

die Koeffizienten in $p(x) = a_0 + a_1x + \dots + a_nx^n$ zu ermitteln, beachte man, dass die Rekursion

$$p_n(x) := f_{0,1,\dots,n},$$

$$p_k(x) := f_{0,1,\dots,k} + (x - x_k)p_{k+1}(x) \quad (k = n - 1, n - 2, \dots, 0)$$

mit $p_0(x) = p(x)$ endet.

Durch Koeffizientenvergleich in dieser Rekursion erhält man folgenden Algorithmus.

Direkte Verfahren für spezielle Systeme

Vandermonde-Matrizen

Gegeben: Dividierte Differenzen der Newton-Darstellung in den Variablen f_0, \dots, f_n , zugehörige x -Werte x_0, \dots, x_n paarweise verschieden.

```
for  $k = n - 1 : 1 : 0$  do
  for  $j = k : n - 1$  do
     $f_j := f_j - f_{j+1}x_k$ 
  end for
end for
```

- Am Ende stehen die Koeffizienten a_0, \dots, a_n in den Variablen f_0, \dots, f_n .
- **Aufwand:** $n^2 + n$ Flops.
- **Fazit:** Da das Hintereinanderausführen der beiden vorangehenden Algorithmen die Koeffizienten $\{a_i\}_{i=0}^n$ des Interpolationspolynoms der Wertepaare $\{(x_i, f_i)\}_{i=0}^n$ in $\frac{5}{2}n^2 + O(n)$ Flops liefert, ist damit auch die äquivalente Aufgabe, das LGS $V^\top \mathbf{a} = \mathbf{f}$ zu lösen, mit derselben Komplexität gelöst.

Mit den $(n + 1) \times (n + 1)$ -Matrizen

$$D_k := \text{diag}(\underbrace{1, \dots, 1}_{k+1}, x_{k+1} - x_0, \dots, x_n - x_{n-k+1}),$$

$$L_k(\alpha) := \left[\begin{array}{c|c} I_k & O \\ \hline O & T_{n+1-k} \end{array} \right],$$

$k = 0, 1, \dots, n - 1$, wobei

$$T_{n+1-k} = \text{tridiag}(-\alpha, 1, 0) \in \mathbb{R}^{(n+1-k) \times (n+1-k)}, \quad \alpha \in \mathbb{R},$$

definieren wir eine normierte untere Δ -Matrix L sowie eine obere Δ -Matrix R gemäß

$$L := L_{n-1}(x_{n-1})L_{n-2}(x_{n-2}) \dots L_0(x_0),$$

$$R := L_0(1)^\top D_0^{-1} L_1(1)^\top D_1^{-1} \dots L_{n-1}(1)^\top D_{n-1}^{-1}.$$

Satz 4.8

Für $V = V(x_0, x_1, \dots, x_n)$ ist

$$V^{-1} = RL.$$

Mit anderen Worten: Der Algorithmus zur Lösung von $V^T \mathbf{a} = \mathbf{f}$ berechnet implizit eine **RL-Zerlegung** von V^{-1} .

- Dies kann zur Lösung von LGSen der Bauart $V\mathbf{z} = \mathbf{b}$ verwendet werden.
- Der folgende Algorithmus berechnet $\mathbf{z} = V^{-1}\mathbf{b}$ in $5n^2/2$ Flops.
- Eingabedaten sind x_0, x_1, \dots, x_n (paarweise verschieden), so dass $V = V(x_0, x_1, \dots, x_n)$, und $\mathbf{b} = [b_0, b_1, \dots, b_n]^T$.
- Die rechte Seite \mathbf{b} wird mit der Lösung \mathbf{z} überschrieben.

Direkte Verfahren für spezielle Systeme

Vandermonde-Matrizen

Gegeben: x_0, x_1, \dots, x_n sodass $V = V(x_0, x_1, \dots, x_n)$, $\mathbf{b} = [b_0, b_1, \dots, b_n]^T$.

```
for  $k = 0 : n - 1$  do  
  for  $j = n : -1 : k + 1$  do  
     $b_j = b_j - x_k * b_{j-1}$   
  end for  
end for  
for  $k = n - 1 : -1 : 0$  do  
  for  $j = k + 1 : n$  do  
     $b_j = b_j / (x_j - x_{j-k-1})$   
  end for  
  for  $j = k : n - 1$  do  
     $b_j = b_j - b_{j+1}$   
  end for  
end for
```

- 1 Einführung und Begriffe
- 2 Gleitpunktarithmetik und Fehleranalyse
- 3 Direkte Verfahren zur Lösung linearer Gleichungssysteme
- 4 Direkte Verfahren für spezielle Systeme
 - 4.1 Die Cholesky-Zerlegung
 - 4.2 Bandmatrizen, Tridiagonalmatrizen
 - 4.3 Schwach besetzte Matrizen
 - 4.4 Vandermonde-Matrizen
 - 4.5 Toeplitz-Matrizen

Eine Matrix T heißt **Toeplitz-Matrix**, wenn sie die Form

$$T = \begin{bmatrix} t_0 & t_1 & t_2 & \cdots & \cdots & t_{n-1} \\ t_{-1} & t_0 & t_1 & & & \vdots \\ t_{-2} & t_{-1} & t_0 & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & t_0 & t_1 \\ t_{1-n} & \cdots & \cdots & \cdots & t_{-1} & t_0 \end{bmatrix},$$

also konstante Einträge auf jeder Diagonalen besitzt:

$$T = [t_{i-j}]_{1 \leq i, j \leq n} \in \mathbb{C}^{n \times n}.$$

Toeplitz-Matrizen treten in vielen Anwendungen, etwa der Diskretisierung von Integraloperatoren, aber auch in vielen Algorithmen der digitalen Signalverarbeitung auf.

- Symmetrisch-positiv definite Toeplitz-Systeme mit spezieller rechter Seite (sog. **Yule-Walker-Gleichungen**):

$$\begin{bmatrix} 1 & t_1 & \cdots & t_{n-1} \\ t_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{n-1} & \cdots & t_1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} = - \begin{bmatrix} t_1 \\ \vdots \\ t_{n-1} \\ t_n \end{bmatrix},$$

oder kurz, $T_n \mathbf{y} = -\mathbf{t}_n$.

lassen sich durch den **Algorithmus von Durbin (1960)** in $O(n^2)$ Operationen lösen.

- Symmetrisch positiv definite Toeplitz-Systeme mit beliebiger rechter Seite

$$\begin{bmatrix} 1 & t_1 & \cdots & t_{n-1} \\ t_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{n-1} & \cdots & t_1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix}, \quad \text{, oder kurz } T_n \mathbf{x} = \mathbf{b},$$

lassen sich mit dem **Algorithmus von Levinson (1947)** ebenfalls in $O(n^2)$ Operationen lösen.

- Mit dem **Algorithmus von Trench (1964)** läßt sich sogar die Inverse T_n^{-1} einer symmetrisch positiv-definiten Toeplitz-Matrix T_n in ebenfalls $O(n^2)$ Operationen berechnen.
- Details nachzulesen in [\[Golub & van Loan, 2012, Section 4.7\]](#)