Proseminar
Hierarchische Matrizen

Arithmetik der $\mathcal{H}$-Matrizen
Die LU- und Choleskyzerlegung

Crameri Remo, 11. Juni 2004

Gliederung des Vortrages

0. Einführung und Motivation

1. Wiederholung der LU- und Choleskyzerlegung

2. Die LU-Zerlegung der $\mathcal{H}$-Matrizen

3. Zusammenfassung

Ziel des Vortrages:

Eine gegebene $\mathcal{H}$-Matrix A möglichst effizient in ihre LU-Zerlegung darstellen.

0. Einführung und Motivation

Ein wichtiger Bestandteil unseres Seminars ist das Lösen linearer Gleichungsysteme. Aus der numerischen Mathematik wissen wir, dass wir ein gegebenes Gleichungssystem

$$Ax = b$$

mit Hilfe von Zerlegungen der Matrix A effizient lösen können.

In der Numerik haben wir zwei solche Zerlegungen kennengelernt:

die LU-Zerlegung
die Choleskyzerlegung.

In diesem Vortrag beschäftigen wir uns damit, wie wir die LU-Zerlegung einer gegebenen $\mathcal{H}$-Matrix A effizient berechnen können.

Ein Verfahren, um die Choleskyzerlegung zu berechnen, kann analog entwickelt werden.

1. Wiederholung der LU- und Choleskyzerlegung

Die LU-Zerlegung

Um das Verfahren des Rückwärtseinsetzens anwenden zu können, ist eine Dreiecksmatrix U notwendig. Im allgemeinen hat jedoch die Matrix A des Gleichungssystems nicht diese Form. Mit Hilfe des Gauss-schen Eliminationsverfahren können wir aber dies erreichen.

Sei das Gleichungssystem $Ax = b$ gegeben, wobei die Matrix $A = (a_{ij})$ ist. Wir führen die Bezeichnung

$$l_{ik} = \frac{a_{ik}}{a_{kk}}$$

ein.

Mit Hilfe von l_{ik} können wir dann das Gauss-schen Eliminationsverfahren durchführen und erhalten daraus eine obere Dreiecksmatrix, die wir mit U bezeichnen.

Wir definieren weiter die untere Dreiecksmatrix L wie folgt:

$$L = (l_{ij}) \text{ mit } l_{ij} = \begin{cases} \frac{a_{ik}}{a_{kk}} \text{ falls } i > j \\ 1 \text{ falls } , i = j \\ \quad O \text{ sonst} \end{cases}$$

Es sei an dieser Stelle noch bemerkt, dass die neue, schrittweise erhaltene Matrizen im Speicherplatz der ursprünglichen Matrix A gespeichert werden. Mit (a_{ij}) bezeichnet man daher die Elemente der neuen Matrix!

Aus der numerischen Mathematik wissen wir aber (unter Voraussetzung, dass die Gauss-schen Elimination sich durchführen lässt), dass zwischen A, L und U folgende Beziehung gilt:

$$A = L \cdot U$$

Wir haben also die Matrix A in das Produkt $L \cdot U$ zerlegt. Eine solche Zerlegung (die eindeutig ist!) nennen wir

die LU-Zerlegung von A.

Beispiel:

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix}$$

Führt man die obige Schritte durch, dann erhalten wir die folgende LU-Zerlegung für A

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 & 0 \\ 0 & -\frac{2}{3} & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{4} & 1 & 0 \\ 0 & 0 & 0 & -\frac{4}{5} & 1 \end{pmatrix}, \mathbf{U} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & -1 & 0 & 0 \\ 0 & 0 & \frac{4}{3} & -1 & 0 \\ 0 & 0 & 0 & \frac{5}{4} & -1 \\ 0 & 0 & 0 & 0 & \frac{6}{5} \end{pmatrix}$$

In der Praxis wird es hier am Ort gearbeitet, das heisst, die neuen Matrizen L und U werden im Speicher von der ursprünglichen Matrix A gespeichert.

Die Hauptarbeit beim Lösen eines Gleichungssystems mit Hilfe der LU-Zerlegung, wird für die Berechnung der Zerlegung selbst verwendet.

Falls aber die Matirx A schon als zerlegt vorausgesetzt wird, dann reduziert sich der Arbeitsaufwand sehr:

Sei $A = L \cdot U$ gegeben.

Um die Lösung des Gleichungssystems zu berechnen, müssen wir nur folgende Schritte durchführen:

$$Ax = L \underbrace{Ux}_{:=y} = b$$

Löse $Ly = b$ auf (Vorwärtseinsetzen)

Löse $Ux = y$ auf (Rückwärtseinsetzen)

Eine genauere Betrachtung der Punktoperationen für das Lösen eines gegebenen Gleichgunssystems zeigt ganz eindeutig, dass das Verfahren mit der LU-Zerlegung viel effizienter als die Berechnung der exakten Inverse A^{-1} wird, wenn grosse dünnbesetzte Matrizen vorliegen.

Die Cholesky-Zerlegung

Für eine bestimmte Klasse von Matrizen lässt sich der Arbeits- und Speicheraufwand der LU-Zerlegung noch weiter reduzieren, insbesondere falls symmetrische positiv-definite Matrizen vorliegen.

Lemma 1: Zu jeder reellen symmetrischen, positiv definiten Matrix A gibt es genau eine obere Dreiecksmatrix $R = (r_{ij})$ mit folgenden Eigenschaften:

$$r_{ij} = 0 \text{ für } i > j$$
$$r_{ii} > 0 \text{ für } i = 1, 2, ..., n$$
$$A = R^T R.$$

Eine solche Zerlegung von A nennt man

die Cholesky-Zerlegung

der Matrix A. Seine Existenz wird unter explizite Konstruktion gezeigt.

Zusammenhang LU- und Cholesky-Zerlegung

Sei A symmetrisch positiv definit. Da sich das Gauss-Algorithmus durchführen lässt, können wir dann die eindeutige LU von A bestimmen:

$$A = L \cdot U$$

Mit einer kleinen Rechnung (und unter Verwendung, dass die LU-Zerlegung eindeutig ist) kann man zeigen, dass durch

$$D = \text{diag}(u_{11}, u_{22}, ..., u_{nn}) \text{ und } V = L^T.$$

eine Zerlegung von U definiert wird, d.h es gilt:

$$U = D \cdot V$$

Somit haben wir für die symmetrische Matrix A eine symmetrische Zerlegung der Form

$$A = L \cdot D \cdot L^T$$

gefunden.

Um den Zusammenhang mit der Cholesky-Zerlegung ganz zeigen zu können, müssen wir aus D (oder besser gesagt, aus ihren Diagonalelementen) die Wurzel ziehen, d.h wir müssen D in das folgende Produkt zerlegen:

$$D = D^{1/2} \cdot D^{1/2}.$$

Nun ist es aber klar (mit der Eindeutigkeit der Cholesky-Zerlegung) dass wir mit

$$R = D^{1/2} \cdot L^T$$

die Cholesy-Zerlegung von A gefunden haben, denn

$$R^T \cdot R = (D^{1/2}L^T)^T(D^{1/2}L^T) = L(D^{1/2})^T D^{1/2}L^T = LU = A.$$

2. Die LU-Zerlegung von $\mathcal{H}$-Matrizen

Eine approximative $\mathcal{H}-LU$ Zerlegung einer $\mathcal{H}$-Matrix A ist definiert als die Zerlegung der folgenden Form:

$$A \approx L_{\mathcal{H}} U_{\mathcal{H}}$$

wobei die zwei Matrizen (obere- und untere Dreiecksmatrizen) $L_{\mathcal{H}}$ und $U_{\mathcal{H}}$ das $\mathcal{H}$-Matrix Format haben.

Um diese Zerlegung durchzuführen, müssen wir uns zuerst überlegen, wie wir ein trianguläres $\mathcal{H}$-Matrixsystems $LX = B$ (wobei die Matrix L das $\mathcal{H}$-Format hat) lösen können:

Das Lösen eines triangulären Systems der Form $LX = B$ (X, B sind normale Matrizen und L ist eine untere Dreiecksmatrix mit $\mathcal{H}$-Format) wird rekursiv durchgeführt:

1.

$$\mathbf{L} = \begin{pmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{pmatrix}, \mathbf{X} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

2. Löse zuerst das System

$$L_{11}X_{11} = B_{11}$$

und

$$L_{11}X_{12} = B_{12}$$

nach X_{11} und X_{12} auf (oder setze voraus, dass die Lösung schon bekannt ist, Induktionsannahme, ev. mit Befehl `dtrsv`).

3. Danach löse

$$L_{22}X_{21} = B_{21} - L_{21}X_{11}$$

und

$$L_{22}X_{22} = B_{22} - L_{21}X_{12}$$

nach X_{21} und X_{22} auf.

Schliesslich sind wir im Stande, die gewünschte Zerlegung zu berechnen: Wir erhlaten die $\mathcal{H}$-LU Zerlegung $A = LU$ rekursiv wie folgt:

1.

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

2. Wir setzen voraus (Induktionsannahme, ev. mit Befehl LAPACK `dgetrf`), dass wir die LU-zerlegung für A_{11} schon kennen, das heisst

$$L_{11}U_{11} = A_{11}$$

ist bekannt, was uns die Elemente L_{11}, U_{11} von der Faktorisierung LU liefert.

3. Löse nun das trianguläre System

$$L_{11}U_{12} = A_{12}$$

und

$$L_{21}U_{11} = A_{21}$$

um U_{12} und L_{21} zu berechnen.

Bemerkungen:

(a) Bei jedem Schritt, den wir durchführen (Berechnung einer Inverse, Multiplikation, Addition...) benützen wir nicht die gewöhnliche Arithmetik für Matrizen! Damit das gewünschte Format nicht verloren geht, verwenden wie **immer** die **formattierte Arithmetik** für die $\mathcal{H}$-Matrizen.

(b) Damit dies aber möglich ist, müssen wir zuerst sicher sein, dass die Bedingungen für die formatierte Arithmetik (gleiche block cluster Strukturen!) erfüllt sind.

(c) Sollten die Strukturen der für L und U zugrundliegenden Bäumen nicht gleich sein, dann können wir die formatierte Arithmetik nicht verwenden. Das zwingt uns dazu, die Zerlegung von A zu berechnen, ohne die Eigenschaft ausnützen zu können, dass A eine $\mathcal{H}$-Matrix ist (Befehl: LAPACK `dgetrf`).

Schliesslich erhalten wir aus

$$L_{22}U_{22} = A_{22} - L_{21}U_{12}$$

die Blöcke L_{22} und U_{22} indem wir die LU-Zerlegung von

$$A_{22} - L_{21}U_{12}$$

bilden.

Bemerkung:
Wiederum setzten wir voraus, dass die Strukturen der Bäume gleich sind, damit wir die formatierte Arithmetik anwenden können.

Mit diesem rekursiv definierten Verfahren haben wir alle Blöcke der Zerlegung $L \cdot U$ berechnet und somit die gewünschte Zerlegung gefunden.

Bemerkung:
Ausgehend von der LU-Zerlegung kann man analog ein Verfahren entwickeln, um die Cholesky-Zerlegung einer Matrix A zu finden.

Übung :-)

Zusammenfassung

THE END