Preconditioned Arnoldi methods for systems of nonlinear equations

Author(s):
Jaschke, Leonhard

Publication Date:
2004

Permanent Link:
https://doi.org/10.3929/ethz-a-004668316

DISS. ETH NO. 15223

Preconditioned Arnoldi Methods for Systems of Nonlinear Equations

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Technical Sciences

presented by
LEONHARD JASCHKE

Dipl. Informatik-Ing. ETH
born on March 24, 1971
citizen of Austria

accepted on the recommendation of
Prof. Dr. Walter Gander, examiner
Prof. Dr. Martin Gutknecht, co-examiner

2003

Zusammenfassung

Krylov-Raum-Methoden zum Lösen linearer Gleichungssysteme haben die Eigenschaft, dass sie implizit das minimale Polynom der Systemmatrix in bezug auf den ersten Residuumsvektor bestimmen.

Vektorextrapolationsmethoden werden verwendet, um die Konvergenz von Vektorsequenzen zu beschleunigen. Sie benötigen keine Information, wie eine solche Sequenz erzeugt wurde. Wenn aber die Sequenz durch eine lineare Iteration erzeugt wurde, kann gezeigt werden, dass der extrapolierte Grenzwert einer Vektorextrapolationsmethode die Lösung eines linearen Gleichungssystems ist. Vektorextrapolationsmethoden approximieren explizit die Koeffizienten des minimalen Polynoms dieses Gleichungssystems ohne Zuhilfenahme der Systemmatrix und des Vektors auf der rechten Seite. Es ist bereits bekannt, dass diese Methoden und Krylov-Raum-Methoden in exakter Arithmetik dieselben Iterierten erzeugen, falls sie zum Lösen von linearen Gleichungssystemen verwendet werden.

Die explizite Bestimmung der Koeffizienten des minimalen Polynoms ist numerisch nicht stabil. Krylov-Raum-Methoden verwenden einen genaueren, impliziten Ansatz. Auf der anderen Seite aber sind Vektorextrapolationsmethoden im Stande, *nichtlineare Gleichungssysteme zu lösen*, die einen dominanten linearen Anteil in der Nähe des zu extrapolierenden Grenzwertes aufweisen.

Der erzeugende Prozess der Vektorsequenz einer Vektorextrapolationsmethode kann im Krylov-Raum-Kontext als Vorkonditionierer angesehen werden. Unter Verwendung der zentralen Idee von Vektorextrapolationsmethoden (Differenzvektoren der Sequenzvektoren) kann der Vorkonditionierer um eine rekursive Vorschrift erweitert werden, um die Matrix-Vektor-Multiplikation von wiedergestarteten Krylov-Raum-Methoden zu ersetzen. Auf diese Weise können Krylov-Raum-Methoden zum Lösen linearer Gleichungssysteme auf das Lösen nichtlinearer Gleichungssysteme adaptiert werden. Sie besitzen dann ähnliche Konvergenzeigenschaften wie Vektorextrapolationsmethoden.

er wissenschaftliche Hauptbeitrag dieser Doktorarbeit ist die Entwicklung und Umsetzung dieses Ansatzes für die allgemeinen Arnoldi-Methoden *FOM* und *GMRES*. Dadurch erhalten wir zwei neue Methoden zum Lösen nichtlinearer Gleichungssysteme — *vorkonditionierte Arnoldi-Methoden zum Lösen nichtlinearer Gleichungssysteme*, die für gewisse Problemstellungen eine bessere Leistung erbringen können als etablierte Methoden wie z.B. inexakte Newton-Methoden. Die neuen Algorithmen wurden anhand der Chandrasekhar H-Gleichung und einem Wärmestrahlungsproblem des Forschungszentrums der Firma ABB in Dättwil, Schweiz getestet.

Ein weiterer Beitrag dieser Doktorarbeit ist die Entwicklung einer allgemeinen Theorie für Vektorextrapolationsmethoden. Diese kann auch auf die neu entwickelten Algorithmen angewendet werden. Die Hauptaussage dieser Theorie ist: Vektorextrapolationsmethoden sind Implementationen der Methode von Henrici (eine Verallgemeinerung der Methode von Steffensen auf Vektorsequenzen). Neu in dieser Theorie ist die Formulierung eines Kontorovich-Theorems für die Methode von Henrici, das die Bedingungen für ihre Konvergenz festlegt. Aufgrund unserer geometrischen Untersuchengen waren wir ausserdem in der Lage, eine Klasse von skalaren Iterationsvorschriften zu beschreiben, für welche die Schmidt-Shanks-Transformation *in nur einem Schritt* konvergiert.

In dieser Doktorarbeit werden die wichtigsten herkömmlichen Verfahren zum Lösen nichtlinearer Gleichungssysteme beschrieben. Überdies wird ein Überblick über die Theorie der linearen Krylov-Raum-Methoden gegeben. Beispiele und Graphiken illustrieren unsere Ausführungen.

Summary

Krylov space methods for solving a nonsingular linear system of equations are known to have the property to determine implicitly the *minimal polynomial* of the system matrix with respect to the initial residual vector, i.e. a polynomial of minimum degree, which is annihilated by the system matrix of the linear system when multiplied with the first residual vector.

Vector extrapolation methods have the purpose to accelerate the convergence of a vector sequence. No knowledge is needed on how the vector sequence is generated. However, if the sequence is generated implicitly by a linear iteration, then the extrapolated limit of the sequence is the fixed-point of a linear system. Without making use of the system matrix and the right-hand side of this system these methods explicitly estimate the coefficients of the minimal polynomial of the system matrix with respect to the first residual vector. It is already a known fact that vector extrapolation methods and corresponding Krylov space methods generate the same iterates in exact arithmetic for linear systems of equations.

The explicit determination of the coefficient of the minimal polynomial is not a stable process. The implicit Krylov approach is better. On the other hand vector extrapolation methods are capable of solving systems of

nonlinear equations with a dominant linear part in the area of interest.

The generating process of the vector sequence of vector extrapolation methods can be regarded as a preconditioner in terms of Krylov space methods. By using the central ideas of vector extrapolation methods (taking differences of sequence elements) the preconditioner can be expanded into a recursive scheme to replace the matrix vector multiplication in a *restarted Krylov method*. This way common linear Arnoldi methods can be extended to nonlinear ones, which have the same convergence properties as vector extrapolation methods.

The main scientific contribution of this thesis is the development and implementation of this approach for the general Arnoldi type Krylov space methods *FOM* and *GMRES*. Thus we get two new methods for solving systems of nonlinear equations — *preconditioned Arnoldi methods for solving systems of nonlinear equations*. For particular problems they have a better performance than common methods, e.g. *inexact Newton methods*. We have tested the new algorithms on the Chandrasekhar H-equation and a heat transfer problem problem of the research center of the ABB company in Dättwil, Switzerland.

Another contribution of this thesis is the development of a general theory of vector extrapolation methods, which can be also applied to the new methods. The main statement of the theory is that vector extrapolation methods are implementations of *Henrici's method* (a generalization of Steffensen's method to vector sequences). New in this thesis is also a Kantorovich like theorem for Henrici's method stating the conditions for its convergence. Furthermore we examined the geometric properties of the Schmidt-Shanks transformation (another generalization of Steffensen's method) and were thus able to specify a class of scalar iterations for which this transformation converges within *only one* step.

Several common methods for solving systems of nonlinear equations are reviewed in this thesis as well as the theory of Krylov space methods. Our considerations are illustrated by many examples and figures.