16

F·R·O·N·T·I·E·R·S
IN APPLIED MATHEMATICS

Iterative Methods for Linear and Nonlinear Equations

C.T. Kelley

siam

Iterative Methods for Linear and Nonlinear Equations

C.T. Kelley
North Carolina State University

Society for Industrial and Applied Mathematics
siam. Philadelphia 1995

10 9 8 7 6 5 4 3 2

Library of Congress Cataloging-in-Publication Data

Kelley, C. T.
Iterative methods for linear and nonlinear equations / C. T. Kelley
p. cm. — (Frontiers in applied mathematics ; vol. 16)
Includes bibliographical references and index.
ISBN 0-89871-352-8 (pbk.)
1. Iterative methods (Mathematics) I. Title. II. Series: Frontiers in applied mathematics ; 16.
QA297.8.K45 1995
519.4—dc20 95-32249

Preface

This book on iterative methods for linear and nonlinear equations can be used as a tutorial and a reference by anyone who needs to solve nonlinear systems of equations or large linear systems. It may also be used as a textbook for introductory courses in nonlinear equations or iterative methods or as source material for an introductory course in numerical analysis at the graduate level. We assume that the reader is familiar with elementary numerical analysis, linear algebra, and the central ideas of direct methods for the numerical solution of dense linear systems as described in standard texts such as [7], [105], or [184].

Our approach is to focus on a small number of methods and treat them in depth. Though this book is written in a finite-dimensional setting, we have selected for coverage mostly algorithms and methods of analysis which extend directly to the infinite-dimensional case and whose convergence can be thoroughly analyzed. For example, the matrix-free formulation and analysis for GMRES and conjugate gradient is almost unchanged in an infinite-dimensional setting. The analysis of Broyden's method presented in Chapter 7 and the implementations presented in Chapters 7 and 8 are different from the classical ones and also extend directly to an infinite-dimensional setting. The computational examples and exercises focus on discretizations of infinite-dimensional problems such as integral and differential equations.

We present a limited number of computational examples. These examples are intended to provide results that can be used to validate the reader's own implementations and to give a sense of how the algorithms perform. The examples are not designed to give a complete picture of performance or to be a suite of test problems.

The computational examples in this book were done with MATLAB® (version 4.0a on various SUN SPARCstations and version 4.1 on an Apple Macintosh Powerbook 180) and the MATLAB environment is an excellent one for getting experience with the algorithms, for doing the exercises, and for small-to-medium scale production work.[1] MATLAB codes for many of the algorithms are available by anonymous ftp. A good introduction to the latest

[1]MATLAB is a registered trademark of The MathWorks, Inc.

version (version 4.2) of MATLAB is the MATLAB Primer [178]; [43] is also a useful resource. If the reader has no access to MATLAB or will be solving very large problems, the general algorithmic descriptions or even the MATLAB codes can easily be translated to another language.

Parts of this book are based upon work supported by the National Science Foundation and the Air Force Office of Scientific Research over several years, most recently under National Science Foundation Grant Nos. DMS-9024622 and DMS-9321938. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation or of the Air Force Office of Scientific Research.

Many of my students and colleagues discussed various aspects of this project with me and provided important corrections, ideas, suggestions, and pointers to the literature. I am especially indebted to Jim Banoczi, Jeff Butera, Steve Campbell, Tony Choi, Moody Chu, Andreas Griewank, Laura Helfrich, Ilse Ipsen, Vickie Kearn, Debbie Lockhart, Carl Meyer, Casey Miller, Ekkehard Sachs, Jeff Scroggs, Mike Tocci, Homer Walker, Steve Wright, Zhaqing Xue, Yue Zhang, and an anonymous reviewer for their contributions and encouragement.

Most importantly, I thank Chung-Wei Ng and my parents for over one hundred years of patience and support.

C. T. Kelley
Raleigh, North Carolina
January, 1995

How to get the software

A collection of MATLAB codes has been written to accompany this book. The MATLAB codes can be obtained by anonymous ftp from the MathWorks server `ftp.mathworks.com` in the directory `pub/books/kelley`, from the MathWorks World Wide Web site,

`http://www.mathworks.com`

or from SIAM's World Wide Web site

`http://www.siam.org/books/kelley/kelley.html`

One can obtain MATLAB from
The MathWorks, Inc.
24 Prime Park Way
Natick, MA 01760,
Phone: (508) 653-1415
Fax: (508) 653-2997
E-mail: info@mathworks.com
WWW: http://www.mathworks.com

Contents

Linear Equations

Chapter 1

Basic Concepts and Stationary Iterative Methods

1.1. Review and notation

We begin by setting notation and reviewing some ideas from numerical linear algebra that we expect the reader to be familiar with. An excellent reference for the basic ideas of numerical linear algebra and direct methods for linear equations is [184].

We will write linear equations as

$$Ax = b, \tag{1.1}$$

where A is a nonsingular $N \times N$ matrix, $b \in R^N$ is given, and

$$x^* = A^{-1}b \in R^N$$

is to be found.

Throughout this chapter x will denote a potential solution and $\{x_k\}_{k\geq 0}$ the sequence of iterates. We will denote the ith component of a vector x by $(x)_i$ (note the parentheses) and the ith component of x_k by $(x_k)_i$. We will rarely need to refer to individual components of vectors.

In this chapter $\|\cdot\|$ will denote a norm on R^N as well as the *induced matrix norm*.

DEFINITION 1.1.1. *Let $\|\cdot\|$ be a norm on R^N The* induced matrix norm *of an $N \times N$ matrix A is defined by*

$$\|A\| = \max_{\|x\|=1} \|Ax\|.$$

Induced norms have the important property that

$$\|Ax\| \leq \|A\|\|x\|.$$

Recall that the *condition number* of A relative to the norm $\|\cdot\|$ is

$$\kappa(A) = \|A\|\|A^{-1}\|,$$

where $\kappa(A)$ is understood to be infinite if A is singular. If $\|\cdot\|$ is the l^p norm

$$\|x\|_p = \left(\sum_{j=1}^{N} |(x)_i|^p \right)^{1/p}$$

we will write the condition number as κ_p.

Most iterative methods terminate when the residual

$$r = b - Ax$$

is sufficiently small. One termination criterion is

$$\frac{\|r_k\|}{\|r_0\|} < \tau, \tag{1.2}$$

which can be related to the error

$$e = x - x^*$$

in terms of the condition number.

LEMMA 1.1.1. *Let $b, x, x_0 \in R^N$. Let A be nonsingular and let $x^* = A^{-1}b$.*

$$\frac{\|e\|}{\|e_0\|} \le \kappa(A)\frac{\|r\|}{\|r_0\|}. \tag{1.3}$$

Proof. Since

$$r = b - Ax = -Ae$$

we have

$$\|e\| = \|A^{-1}Ae\| \le \|A^{-1}\|\|Ae\| = \|A^{-1}\|\|r\|$$

and

$$\|r_0\| = \|Ae_0\| \le \|A\|\|e_0\|.$$

Hence

$$\frac{\|e\|}{\|e_0\|} \le \frac{\|A^{-1}\|\|r\|}{\|A\|^{-1}\|r_0\|} = \kappa(A)\frac{\|r\|}{\|r_0\|},$$

as asserted. □

The termination criterion (1.2) depends on the initial iterate and may result in unnecessary work when the initial iterate is good and a poor result when the initial iterate is far from the solution. For this reason we prefer to terminate the iteration when

$$\frac{\|r_k\|}{\|b\|} < \tau. \tag{1.4}$$

The two conditions (1.2) and (1.4) are the same when $x_0 = 0$, which is a common choice, particularly when the linear iteration is being used as part of a nonlinear solver.

1.2. The Banach Lemma and approximate inverses

The most straightforward approach to an iterative solution of a linear system is to rewrite (1.1) as a linear fixed-point iteration. One way to do this is to write $Ax = b$ as

$$x = (I - A)x + b, \tag{1.5}$$

and to define the *Richardson iteration*

$$x_{k+1} = (I - A)x_k + b. \tag{1.6}$$

We will discuss more general methods in which $\{x_k\}$ is given by

$$x_{k+1} = Mx_k + c. \tag{1.7}$$

In (1.7) M is an $N \times N$ matrix called the *iteration matrix*. Iterative methods of this form are called *stationary iterative methods* because the transition from x_k to x_{k+1} does not depend on the history of the iteration. The Krylov methods discussed in Chapters 2 and 3 are not stationary iterative methods.

All our results are based on the following lemma.

LEMMA 1.2.1. *If M is an $N \times N$ matrix with $\|M\| < 1$ then $I - M$ is nonsingular and*

$$\|(I - M)^{-1}\| \le \frac{1}{1 - \|M\|}. \tag{1.8}$$

Proof. We will show that $I - M$ is nonsingular and that (1.8) holds by showing that the series

$$\sum_{l=0}^{\infty} M^l = (I - M)^{-1}.$$

The partial sums

$$S_k = \sum_{l=0}^{k} M^l$$

form a Cauchy sequence in $R^{N \times N}$. To see this note that for all $m > k$

$$\|S_k - S_m\| \le \sum_{l=k+1}^{m} \|M^l\|.$$

Now, $\|M^l\| \le \|M\|^l$ because $\|\cdot\|$ is an matrix norm that is induced by a vector norm. Hence

$$\|S_k - S_m\| \le \sum_{l=k+1}^{m} \|M\|^l = \|M\|^{k+1} \left(\frac{1 - \|M\|^{m-k}}{1 - \|M\|} \right) \to 0$$

as $m, k \to \infty$. Hence the sequence S_k converges, say to S. Since $MS_k + I = S_{k+1}$, we must have $MS + I = S$ and hence $(I - M)S = I$. This proves that $I - M$ is nonsingular and that $S = (I - M)^{-1}$.

Noting that

$$\|(I - M)^{-1}\| \leq \sum_{l=0}^{\infty} \|M\|^l = (1 - \|M\|)^{-1}.$$

proves (1.8) and completes the proof. □

The following corollary is a direct consequence of Lemma 1.2.1.

COROLLARY 1.2.1. *If* $\|M\| < 1$ *then the iteration* (1.7) *converges to* $x = (I - M)^{-1}c$ *for all initial iterates* x_0.

A consequence of Corollary 1.2.1 is that Richardson iteration (1.6) will converge if $\|I - A\| < 1$. It is sometimes possible to *precondition* a linear equation by multiplying both sides of (1.1) by a matrix B

$$BAx = Bb$$

so that convergence of iterative methods is improved. In the context of Richardson iteration, the matrices B that allow us to apply the Banach lemma and its corollary are called *approximate inverses.*

DEFINITION 1.2.1. *B is an* approximate inverse *of A if* $\|I - BA\| < 1$.

The following theorem is often referred to as the *Banach Lemma.*

THEOREM 1.2.1. *If A and B are $N \times N$ matrices and B is an approximate inverse of A. Then A and B are both nonsingular and*

$$\|A^{-1}\| \leq \frac{\|B\|}{1 - \|I - BA\|}, \quad \|B^{-1}\| \leq \frac{\|A\|}{1 - \|I - BA\|}, \tag{1.9}$$

and

$$\|A^{-1} - B\| \leq \frac{\|B\|\|I - BA\|}{1 - \|I - BA\|}, \quad \|A - B^{-1}\| \leq \frac{\|A\|\|I - BA\|}{1 - \|I - BA\|}. \tag{1.10}$$

Proof. Let $M = I - BA$. By Lemma 1.2.1 $I - M = I - (I - BA) = BA$ is nonsingular. Hence both A and B are nonsingular. By (1.8)

$$\|A^{-1}B^{-1}\| = \|(I - M)^{-1}\| \leq \frac{1}{1 - \|M\|} = \frac{1}{1 - \|I - BA\|}. \tag{1.11}$$

Since $A^{-1} = (I - M)^{-1}B$, inequality (1.11) implies the first part of (1.9). The second part follows in a similar way from $B^{-1} = A(I - M)^{-1}$.

To complete the proof note that

$$A^{-1} - B = (I - BA)A^{-1}, A - B^{-1} = B^{-1}(I - BA),$$

and use (1.9). □

Richardson iteration, preconditioned with approximate inversion, has the form

$$x_{k+1} = (I - BA)x_k + Bb. \tag{1.12}$$

If the norm of $I - BA$ is small, then not only will the iteration converge rapidly, but, as Lemma 1.1.1 indicates, termination decisions based on the preconditioned residual $Bb - BAx$ will better reflect the actual error. This method is a very effective technique for solving differential equations, integral equations, and related problems [15], [6], [100], [117], [111]. Multigrid methods [19], [99], [126], can also be interpreted in this light. We mention one other approach, *polynomial preconditioning*, which tries to approximate A^{-1} by a polynomial in A [123], [179], [169].

1.3. The spectral radius

The analysis in § 1.2 related convergence of the iteration (1.7) to the norm of the matrix M. However the norm of M could be small in some norms and quite large in others. Hence the performance of the iteration is not completely described by $\|M\|$. The concept of spectral radius allows us to make a complete description.

We let $\sigma(A)$ denote the set of eigenvalues of A.

DEFINITION 1.3.1. *The spectral radius of an $N \times N$ matrix A is*

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda| = \lim_{n \to \infty} \|A^n\|^{1/n}. \tag{1.13}$$

The term on the right-hand side of the second equality in (1.13) is the limit used by the radical test for convergence of the series $\sum A^n$.

The spectral radius of M is independent of any particular matrix norm of M. It is clear, in fact, that

$$\rho(A) \le \|A\| \tag{1.14}$$

for any induced matrix norm. The inequality (1.14) has a partial converse that allows us to completely describe the performance of iteration (1.7) in terms of spectral radius. We state that converse as a theorem and refer to [105] for a proof.

THEOREM 1.3.1. *Let A be an $N \times N$ matrix. Then for any $\epsilon > 0$ there is a norm $\|\cdot\|$ on R^N such that*

$$\rho(A) > \|A\| - \epsilon.$$

A consequence of Theorem 1.3.1, Lemma 1.2.1, and Exercise 1.5.1 is a characterization of convergent stationary iterative methods. The proof is left as an exercise.

THEOREM 1.3.2. *Let M be an $N \times N$ matrix. The iteration* (1.7) *converges for all $c \in R^N$ if and only if $\rho(M) < 1$.*

1.4. Matrix splittings and classical stationary iterative methods

There are ways to convert $Ax = b$ to a linear fixed-point iteration that are different from (1.5). Methods such as Jacobi, Gauss-Seidel, and sucessive overrelaxation (SOR) iteration are based on *splittings* of A of the form

$$A = A_1 + A_2,$$

where A_1 is a nonsingular matrix constructed so that equations with A_1 as coefficient matrix are easy to solve. Then $Ax = b$ is converted to the fixed-point problem

$$x = A_1^{-1}(b - A_2x).$$

The analysis of the method is based on an estimation of the spectral radius of the iteration matrix $M = -A_1^{-1}A_2$.

For a detailed description of the classical stationary iterative methods the reader may consult [89], [105], [144], [193], or [200]. These methods are usually less efficient than the Krylov methods discussed in Chapters 2 and 3 or the more modern stationary methods based on multigrid ideas. However the classical methods have a role as preconditioners. The limited description in this section is intended as a review that will set some notation to be used later.

As a first example we consider the Jacobi iteration that uses the splitting

$$A_1 = D, A_2 = L + U,$$

where D is the diagonal of A and L and U are the (strict) lower and upper triangular parts. This leads to the iteration matrix

$$M_{JAC} = -D^{-1}(L + U).$$

Letting $(x_k)_i$ denote the ith component of the kth iterate we can express Jacobi iteration concretely as

$$(x_{k+1})_i = a_{ii}^{-1}\left(b_i - \sum_{j \neq i} a_{ij}(x_k)_j\right). \tag{1.15}$$

Note that A_1 is diagonal and hence trivial to invert.

We present only one convergence result for the classical stationary iterative methods.

THEOREM 1.4.1. *Let A be an $N \times N$ matrix and assume that for all $1 \leq i \leq N$*

$$0 < \sum_{j \neq i} |a_{ij}| < |a_{ii}|. \tag{1.16}$$

Then A is nonsingular and the Jacobi iteration (1.15) *converges to $x^* = A^{-1}b$ for all b.*

Proof. Note that the ith row sum of $M = M_{JAC}$ satisfies

$$\sum_{j=1}^{N} |m_{ij}| = \frac{\sum_{j \neq i} |a_{ij}|}{|a_{ii}|} < 1.$$

Hence $\|M_{JAC}\|_\infty < 1$ and the iteration converges to the unique solution of $x = Mx + D^{-1}b$. Also $I - M = D^{-1}A$ is nonsingular and therefore A is nonsingular. ☐