Journal of Computational Mathematics, Vol.20, No.4, 2002, 437–448.

A REGULARIZED CONJUGATE GRADIENT METHOD FOR SYMMETRIC POSITIVE DEFINITE SYSTEM OF LINEAR EQUATIONS*[1)]

Zhong-zhi Bai
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Shao-liang Zhang
(*Department of Applied Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*)

Abstract

A class of regularized conjugate gradient methods is presented for solving the large sparse system of linear equations of which the coefficient matrix is an ill-conditioned symmetric positive definite matrix. The convergence properties of these methods are discussed in depth, and the best possible choices of the parameters involved in the new methods are investigated in detail. Numerical computations show that the new methods are more efficient and robust than both classical relaxation methods and classical conjugate direction methods.

Key words: Conjugate gradient method, Symmetric positive definite matrix, Regularization, Ill-conditioned linear system.

1. Introduction

Let $\mathbb{R}^n$ represent the real n-dimensional vector space, and $\mathbb{R}^{n\times n}$ the real $n\times n$ matrix space. In this paper, we will study iterative methods for solving the system of linear equations

$$Ax = b, \qquad A \in \mathbb{R}^{n\times n} \quad \text{and} \quad x, b \in \mathbb{R}^n, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is a large sparse and possibly very ill-conditioned *symmetric positive definite* (SPD) matrix, $x \in \mathbb{R}^n$ the unknown vector, and $b \in \mathbb{R}^n$ a given *right-hand side* (RHS) vector.

The *conjugate gradient* (CG) method [9] is an efficient solver for approximating the solution of the system of linear equations (1), provided the coefficient matrix $A \in \mathbb{R}^{n\times n}$ is well-conditioned, or a good preconditioner is cheaply obtainable when it is ill-conditioned. A preconditioner transforms the original linear system (1) by a suitable linear transformation such that the spectral property of the matrix $A \in \mathbb{R}^{n\times n}$ is largely improved, and therefore, the convergence speed of the CG method is considerably accelerated. Two typical ways of constructing a practical preconditioner for an SPD matrix are the *symmetric successive over-relaxation* (SSOR) iteration [16, 1, 2, 3] and the *incomplete Cholesky* (IC) factorization [2, 12]. However, both SSOR and IC preconditioners are only applicable and efficient for a special class of SPD system of linear equations, e.g., a diagonally dominant or an irreducibly weakly diagonally dominant one which may come from the discretization of a second-order self-adjoint elliptic boundary value problem by the finite difference method [16, 12, 2, 15]. Moreover, the IC factorization may break down even for an SPD matrix [11]. Therefore, the existence of an IC factor can not be guaranteed even if we neglect the influence of the rounding error, needless to say its stability and accuracy.

Considering that the CG method is quite efficient for solving an SPD system of linear equations whose coefficient matrix has tightly clustered spectrum [2, 6, 15], in this paper, we present a class of *regularized conjugate gradient* (RCG) method for solving the system of linear equations (1). In the RCG method, the linear system (1) is first regularized by reasonably

* Received March 12, 2000.
[1)] Subsidized by The Special Funds For Major State Basic Research Projects G1999032803.

shifting and contracting the spectrum of the coefficient matrix $A \in \mathbb{R}^{n\times n}$, and its solution is then approximated successively by a sequence of regularized linear systems. At each step of iteration, the regularized linear system itself is iteratively solved by the CG method. Therefore, the RCG method is actually an inner/outer iterative method [13, 14, 7, 5, 8] with a standard splitting iteration as its outer iteration, and the CG iteration as its inner iteration. Evidently, this new approach is quite different from the typical ones, such as the classical relaxation methods [16] and the classical conjugate direction methods [2, 15]. Moreover, the RCG method itself can be again preconditioned by employing an IC or an SSOR preconditioner to the regularized linear system. Then, the CG method is directly applied to this preconditioned regularized linear system at each outer iterate. This naturally leads to a so-called *preconditioned regularized conjugate gradient* (PRCG) method for solving the system of linear equations (1). In actual implementation of the PRCG method, we can suitably choose the shift and contract factors in such a way that the regularized linear system has reasonably good diagonally dominant property such that the IC or the SSOR preconditioner is existent, stable, and accurate, and hence, make it a highly efficient method for solving the system of linear equations (1).

We prove the convergence and estimate the relative residual and error of both RCG and PRCG methods. In particular, we discuss the best possible choices of the shift and contract factors, as well as the best possible number of the inner CG iteration steps. Both theoretical analyses and numerical experiments show that the new regularized conjugate gradient method and its preconditioned variant converge much faster and more robust to the exact solution of the system of linear equations (1) than both classical relaxation methods and classical conjugate direction methods.

2. The Regularized Conjugate Gradient Method

For an SPD matrix $A \in \mathbb{R}^{n\times n}$, we use $\sigma(A)$ to represent its spectrum set, and $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ its smallest and largest eigenvalues, respectively. Denote $\mathcal{I}(A) = [\lambda_{\min}(A), \lambda_{\max}(A)]$. Then any $\lambda \in \sigma(A)$ satisfies $\lambda \in \mathcal{I}(A)$. The condition number $\kappa_2(A)$ of the matrix $A \in \mathbb{R}^{n\times n}$ with respect to the Euclidean norm is given by $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 = \dfrac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$. For any $z \in \mathbb{R}^n$, its A-norm is defined by $\|z\|_A = \sqrt{z^T Az}$.

We use $\nu = \frac{\mu}{\eta}$ to represent the horizontal intercept of the linear transformation $f : \mathbb{R}^1 \to \mathbb{R}^1$,

$$f(t) = \mu + \eta t, \qquad \mu, \eta \in \mathbb{R}^1, \quad \eta \neq 0. \tag{2}$$

Evidently, $f(A) = \eta A(\nu)$, where $A(\nu) = \nu I + A$, is the transformed matrix. The linear transformation (2) maps the spectrum set $\sigma(A)$ of the matrix $A \in \mathbb{R}^{n\times n}$ onto a new set $f(\sigma(A)) = \mu + \eta\sigma(A)$ which is obviously contained in the interval $\mu + \eta\mathcal{I}(A)$. If we choose the reals μ and η such that $\eta \neq 0$ and $\nu > 0$, then it immediately holds that

$$\kappa_2(f(A)) \equiv \frac{\mu + \eta\lambda_{\max}(A)}{\mu + \eta\lambda_{\min}(A)} = \frac{\nu + \lambda_{\max}(A)}{\nu + \lambda_{\min}(A)} < \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \equiv \kappa_2(A).$$

Therefore, the linear transformation (2) may considerably improve the condition number of the matrix $A \in \mathbb{R}^{n\times n}$, provided a pair of suitable constants μ and η is easily obtainable. The constant μ is called a shift, and the constant η a contractor when $|\eta| < 1$ and an amplifier when $|\eta| \geq 1$.

By the linear transformation (2), we can rewrite the system of linear equations (1) as

$$(\nu I + A)x = \nu x + b, \tag{3}$$

where $I \in \mathbb{R}^{n\times n}$ is the identity matrix and $\nu \geq 0$ a constant. Hence, the system of linear equations (1) is equivalent to the system of linear equations (3).

The basic idea of our new *regularized conjugate gradient* (RCG) method is as follows: Given a starting vector $x^{(0)} \in \mathbb{R}^n$; Suppose that we have got approximations $x^{(0)}, x^{(1)}, \ldots, x^{(k)}$ to the solution x^* of the system of linear equations (1), then the next approximation $x^{(k+1)}$ to x^* is obtained through solving the system of linear equations

$$(\nu I + A)x = \nu x^{(k)} + b \tag{4}$$

iteratively, with the CG method, to certain arithmetic precision. More precisely, this RCG method can be described as follows: