Iterative Methods for Nonlinear Operator Equations

A. T. Chronopoulos

Department of Computer Science
University of Minnesota at Minneapolis
Minneapolis, Minnesota 55455

and

Z. Zlatev

Airpollution Laboratory
Danish Agency of Environmental Protection
Risoe National Laboratory
DK-4000 Roskilde, Denmark

ABSTRACT

A nonlinear conjugate gradient method has been introduced and analyzed by J. W. Daniel. This method applies to nonlinear operators with symmetric Jacobians. The conjugate gradient method applied to the normal equations can be used to approximate the soultion of general nonsymmetric linear systems of equations if the condition of the coefficient matrix is small. In this article, we obtain nonlinear generalizations of this method which apply directly to nonlinear operator equations. Under conditions on the Hessian and the Jacobian of the operators, we prove that these methods converge to a unique solution. Error bounds and local convergence results are also obtained.

1. INTRODUCTION

Nonlinear systems of equations often arise when solving initial or boundary value problems in ordinary or partial differential equations. We consider the nonlinear system of equations

$$F(x) = 0 \tag{1.1}$$

where $F(x)$ is a nonlinear operator from a real Euclidean space of dimension N or Hilbert space into itself. The Newton method coupled with Gaussian elimination is an efficient way to solve such nonlinear systems when the

APPLIED MATHEMATICS AND COMPUTATION 51:167–180 (1992)

655 Avenue of the Americas, New York, NY 10010

0096-3003/92/$5.00

dimension of the Jacobian is of small. When the Jacobian is large and sparse, some kind of iterative method may be used. This can be a nonlinear iteration (for example, functional iteration for contractive operators) or an inexact Newton method. In an inexact Newton, the solution of the resulting linear systems is approximated by a linear iterative method. (c.f. [15], [6])

Nonlinear steepest descent methods for the minimal residual and normal equations have been studied by many authors (c.f. [12] and [14]). J. Fletcher and C. M. Reeves [8], and J. W. Daniel [4] have obtained a nonlinear conjugate gradient method that converges if the Jacobian is symmetric and uniformly positive definite. These nonlinear methods reduce to the standard conjugate gradient methods for linear systems. These methods are based on exact line search at each iteration and thus must solve a scalar nonlinear minimization problem in order to determine the steplengths. Several authors have suggested inexact line search and have given conditions under which these methods would still converge [8]. This is done to avoid solving exactly the scalar minimization problem whose derivative evaluation involves evaluation of the nonlinear operator.

The conjugate gradient method applied to normal equations can be used to solve iteratively nonsymmetric linear systems when the condition number of the Jacobian is small. Some preconditioning applied to the original linear system can be used to achieve this goal. Two algorithms exist for the conjugate gradient method applied to the normal equations: the CGNR [1, 11] and CGNE [1, 3] (or Craigs method).

In this article we obtain a nonlinear extension of the Conjugate Gradient methods applied to the normal equations. We assume that the Jacobian and the Hessian of the nonlinear operator are uniformly bounded. We prove global convergence and local convergence results for the nonlinear algorithms. We also give asymptotic steplength estimates and error bounds. These steplengths can be used in implementing these methods. In section 2, we review the CGNR and CGNE methods. In section 3, we derive a nonlinear CGNR method and prove global convergence. In section 4, we derive a nonlinear CGNE method and prove local convergence. In section 5, we obtain asymptotic steplength and error estimates.

2. THE CONJUGATE GRADIENT APPLIED TO THE NORMAL EQUATIONS

Let us consider the system of linear equations $Ax = f$, where A is nonsingular, nonsymmetric matrix of order N. This system can be solved by either of the two normal equations systems:

$$A^T A x = A^t f \tag{2.1}$$

$$AA^T y = f, \quad x = A^T y \tag{2.2}$$

Since both A^TA and AA^T have the same spectrum, we can apply CG to either system to obtain an approximate solution of $Ax = f$.

The CGNR method [11] applies CG to (2.1). Then x_n minimizes the norm of the residual error $E(x_n) = \|f - Ax_n\|_2$ over the affine Krylov subspace

$$x_0 + \left\{A^T r_0, \ldots, (A^TA)^{n-1} A^T r_0\right\}$$

and the resulting algorithm is the following.

ALGORITHM 2.1. The CGNR algorithm.
Initial vector x_0
$r_0 = f - Ax_0$, $p_0 = A^T r_0$
For $n = 0$ **Until** Convergence **Do**
1. $a_n = \dfrac{(A^T r_n, A^T r_n)}{(Ap_n, Ap_n)}$
2. $x_{n+1} = x_n + a_n p_n$ and $r_{n+1} = r_n - a_n Ap_n$.
3. $p_{n+1} = A^T r_{n+1} + b_n p_n$ where $b_n = -\dfrac{(AA^T r_{n+1}, Ap_n)}{\|Ap_n\|^2}$.

EndFor.

The CGNE method [3] applies CG to (2.2). Then x_n minimizes the norm of the error $E(x_n) = \|x^* - x_n\|^2$ over the same affine Krylov subspace as CGNR, and the resulting algorithm is the following.

ALGORITHM 2.2. The CGNE algorithm.
Initial vector x_0
$r_0 = f - Ax_0$, $p_0 = A^T r_0$
For $n = 0$ **Until** Convergence **Do**
1. $a_n = \dfrac{(r_n, r_n)}{(p_n, p_n)}$
2. $x_{n+1} = x_n + a_n p_n$ and $r_{n+1} = r_n - a_n Ap_n$.
3. $p_{n+1} = A^T r_{n+1} + b_n p_n$ where $b_n = -\dfrac{(A^T r_{n+1}, p_n)}{\|p_n\|^2}$.

EndFor.

Since the spectrum of the matrices AA^T and A^TA are the same, we should expect that the performance of CGNR and CGNE is the same. However, CGNE minimizes the norm of the error and may yield better performance.

$$AA^Ty = f, \quad x = A^Ty \tag{2.2}$$

Since both A^TA and AA^T have the same spectrum, we can apply CG to either system to obtain an approximate solution of $Ax = f$.

The CGNR method [11] applies CG to (2.1). Then x_n minimizes the norm of the residual error $E(x_n) = \|f - Ax_n\|_2$ over the affine Krylov subspace

$$x_0 + \left\{A^Tr_0, \ldots, (A^TA)^{n-1}A^Tr_0\right\}$$

and the resulting algorithm is the following.

ALGORITHM 2.1. The CGNR algorithm.
Initial vector x_0
$r_0 = f - Ax_0,\ p_0 = A^Tr_0$
For $n = 0$ **Until** Convergence **Do**
1. $a_n = \dfrac{(A^Tr_n, A^Tr_n)}{(Ap_n, Ap_n)}$
2. $x_{n+1} = x_n + a_n p_n$ and $r_{n+1} = r_n - a_n Ap_n$.
3. $p_{n+1} = A^Tr_{n+1} + b_n p_n$ where $b_n = -\dfrac{(AA^Tr_{n+1}, Ap_n)}{\|Ap_n\|^2}$.

EndFor.

The CGNE method [3] applies CG to (2.2). Then x_n minimizes the norm of the error $E(x_n) = \|x^* - x_n\|^2$ over the same affine Krylov subspace as CGNR, and the resulting algorithm is the following.

ALGORITHM 2.2. The CGNE algorithm.
Initial vector x_0
$r_0 = f - Ax_0,\ p_0 = A^Tr_0$
For $n = 0$ **Until** Convergence **Do**
1. $a_n = \dfrac{(r_n, r_n)}{(p_n, p_n)}$
2. $x_{n+1} = x_n + a_n p_n$ and $r_{n+1} = r_n - a_n Ap_n$.
3. $p_{n+1} = A^Tr_{n+1} + b_n p_n$ where $b_n = -\dfrac{(A^Tr_{n+1}, p_n)}{\|p_n\|^2}$.

EndFor.

Since the spectrum of the matrices AA^T and A^TA are the same, we should expect that the performance of CGNR and CGNE is the same. However, CGNE minimizes the norm of the error and may yield better performance.

The CGNE method is sometimes called Craigs method because it was first proposed by E. J. Craig.

The following bound error can be obtained [5] for the error functional $E(x)$:

$$E(x_n) \leqslant 2\left(\frac{1 - 1/\rho}{1 + 1/\rho}\right)^n E(x_0) \tag{2.3}$$

where $\rho = \|A\|_2 \|A^{-1}\|_2$ is the condition number of the matrix.

3. THE NONLINEAR CGNR METHOD

In this section, we generalize the CGNR iteration to a nonlinear iteration which requires the solution of a scalar equation to determine the steplength. We then prove a global convergence result under assumptions that the Hessian and the Jacobian are uniformly bounded.

Let $F(x)$ be an operator mapping of the Euclidean space R^n (or, even more generally a real Hilbert space) into itself. The notation $F'(x)$ and $F''(x)$ will be used to denote the Frechet and Gateaux derivatives respectively. Also, for simplicity F_n' and F_n'' will denote $F'(x_n)$ and $F''(x_n)$ respectively. We seek to solve iteratively the nonlinear system of equations: $F(x) = 0$. In the linear case $F(x) = Ax - b$ and $F'(x) = A$.

Assume that $F'(x)$ and $F''(x)$ exist at all x and that there exist scalars $0 < m \leqslant M$, $0 < B$ independent of x so that the following conditions are satisfied for any vectors x and v:

$$m^2\|v\|^2 \leqslant \left(\left(F'(x)^T F'(x)\right)v, v\right) \leqslant M^2\|v\|^2 \tag{3.1a}$$

$$\|F''(x)\| \leqslant B \tag{3.1b}$$

REMARK 3.0. (i) The symmetric definite operators $F'(x)^T F'(x)$ and $F'(x)F'(x)^T$ have the same eigenvalues. Thus, the following inequality holds:

$$m^2\|v\|^2 \leqslant \left(F'(x)F'(x)^T v, v\right) \leqslant M^2\|v\|^2.$$

(ii) The left inequality in (3.1a) and the inverse function theorem for differential operators imply that the inverse $F^{-1}(x)$ exists and it is differentiable.

From the left inequality in (3.1a) and the inverse function theorem, we conclude that $F^{-1}(x)$ exists and it is differentiable at all x. We use the mean value theorem for the operator $F^{-1}(x)$ to obtain the following equation

$$(y - x, (y - x)) = (F'^{-1}(z)(f(y) - f(x)), (y - x)).$$

Combining this with the right inequality in (3.1a) we obtain:

$$\|y - x\|^2 \leqslant \frac{1}{m}\|F(y) - F(x)\| \, \|y - x\|.$$

This inequality implies that

$$m\|y - x\| \leqslant \|F(y) - F(x)\| \tag{3.2}$$

By use of the mean value theorem for the operator $F(x)$ and assumption (3.1a), we obtain the following inequality

$$\|F(y) - F(x)\| \leqslant M\|y - x\| \tag{3.3}$$

Under assumptions (3.1), we consider the following nonlinear generalization of CGNR.

ALGORITHM 3.1. The Nonlinear CGNR Algorithm.
Initial vector x_0
$r_0 = -F(x_0)$, $p_0 = F_0^T r_0$
For $n = 0$ **Until** Convergence **Do**
1. Select the smallest positive c_n to minimize $\|F(x_n + cp_n)\|_2$, $c > 0$
2. $x_{n+1} = x_n + c_n p_n$ and $r_{n+1} = -F(x_{n+1})$
3. $b_n = -\dfrac{(F_{n+1}F_{n+1}^T r_{n+1}, F_{n+1}p_n)}{\|F_{n+1}p_n\|^2}$ where $p_{n+1} = F_{n+1}^T r_{n+1} + b_n p_n$

EndFor

The scalars c_n and b_n are defined to guarantee the following two orthogonality relations:

$$(r_n, F_n' p_{n-1}) = 0 \tag{3.4}$$

and

$$(F_n' p_n, F_n' p_{n-1}) = 0. \tag{3.5}$$

Under the assumptions (3.1), the following lemma holds.

LEMMA 3.1. *Let* $\{r_n\}$ *be the nonlinear residuals and* $\{p_n\}$ *be the direction vectors in Algorithm* 3.1 *then the following identities hold*:

(*i*) $(r_n, F_n' p_n) = \|F_n' r_n\|^2$
(*ii*) $\|p_n\|^2 = \|F_n'^T r_n\|^2 + b_{n-1}^2 \|p_{n-1}\|^2$
(*iii*) $\|F_n' F_n'^T r_n\|^2 = \|F_n' p_n\|^2 + b_{n-1}^2 \|F_n' p_{n-1}\|^2$
(*iv*) $m\|r_n\| \leqslant \|F_n'^T r_n\| \leqslant \|p_n\|$
(*v*) $\|p_n\| \leqslant \dfrac{M^2}{m}\|r_n\|$
(*vi*) $\|r_{n+1}\| \leqslant \|r_n\|$.

PROOF. The orthogonality relations (3.4) and (3.5) combined with Step 3 of Algorithm 3.1 imply (i)–(iii). Equality (ii) and (3.1a) are used in proving inequality in (iv) as follows:

$$m\|r_n\|^2 \leqslant \|F_n'^T r_n\|^2 \leqslant \|p_n\|^2$$

Equality (iii) and (3.1a) are used in proving inequality (v) as follows:

$$m\|p_n\| \leqslant \|F_n' p_n\| \leqslant \|F_n' F_n'^T r_n\| \leqslant M^2 \|r_n\|$$

Inequality (vi) follows from the definition of c_n. □

REMARK 3.1. Let $f_n(c)$ denote the scalar function: $\frac{1}{2}\|F(x_n + cp_n)\|^2$. Its first and second derivatives are given by:

$$f_n'(c) = (F(x_n + cp_n), F'(x_n + cp_n) p_n) \tag{3.6}$$

$$f_n''(c) = ((F''(x_n + cp_n) p_n, p_n), F(x_n + cp_n)) + \|F'(x + cp_n) p_n\|^2 \tag{3.7}$$

The following upper and lower bounds on $f_n''(c)$ can be computed from (3.6), the assumptions (3.1), and Lemma 3.1 (vi).

$$\|p_n\|^2(m^2 - B\|r_0\|) \leqslant m^2\|p_n\|^2 - B\|p_n\|^2\|r_n\| \leqslant f_n''(c) \qquad (3.8)$$

$$f_n''(c) \leqslant \|p_n\|^2(M^2 + B\|r_n\|) \leqslant \|p_n\|^2(M^2 + B\|r_0\|) \qquad (3.9)$$

We next prove that under assumptions (3.1) the nonlinear CGNR iteration converges globally to a unique solution.

THEOREM 3.1. *Under the assumptions* (3.1) *on the nonlinear operator* $F(x)$ *the sequence* x_n *generated by Algorithm* 3.1 *is well-defined for any* x_0, *it converges to a unique solution* x^* *of the nonlinear system* $F(x) = 0$ *and*

$$\|x_n - x^*\| < \frac{1}{m}\|F(x_n)\|.$$

PROOF. The proof is divided in four parts.

Firstly, we prove the existence of c_n in 1 of Algorithm 3.1. The derivative of the real function f_n at zero [because of Lemma 3.1 (i)] is:

$$f_n'(0) = -\|F_n'^T r_n\|^2 < 0$$

So there exists $c > 0$ such that $\|F(x_n + cp_n)\| < \|r_n\|$. We must prove that there is a $c > 0$ such that $f_n(0) \leqslant f_n(c)$. This would imply that there exists $c_n > 0$ where $f_n(c)$ assumes a local minimum. From inequality (3.2) by inserting $x = x_n$ and $y = x_n + cp_n$ we conclude that $F(y)$ grows unbounded for $c \to \infty$. This proves that there is a $0 < c$ such that $f_n(0) \leqslant f_n(c)$.

Secondly, we obtain a lower bound on the steplength c_n. Taylor's expansion gives $f_n'(c_n) = 0 = f_n'(0) + c_n f_n''(\bar{c}_n)$, where $\bar{c}_n = t_n c_n$ for some t in $(0, 1)$. We solve for c_n. We then use the upper bound in inequality (3.9), (3.1a), and Lemma 3.1 (v) to obtain

$$\frac{m^4}{M^4(M^2 + B\|r_0\|)} \leqslant \frac{m^2\|r_n\|^2}{\|p_n\|^2(M^2 + B\|r_0\|)} \leqslant \frac{\|F_n'^T r_n\|^2}{\|p_n\|^2(M^2 + B\|r_n\|)} \leqslant c_n. \qquad (3.10)$$

Thirdly, we prove that the sequence of residual norms decreases to zero. For $\bar{c} = tc$ for some t in [0, 1] we have

$$f_n(c) = \frac{1}{2}\|r_n\|^2 - c\|F_n'^T r_n\|^2 + \frac{c^2}{2} f_n''(\bar{c})$$

Now by inserting $c = \dfrac{\|F_n'^T r_n\|^2}{\|p_n\|^2 (M^2 + B\|r_n\|)}$ we obtain

$$\frac{1}{2}\|r_{n+1}\|^2 = f_n(c_n) \leqslant f_n(c) \leqslant \frac{1}{2}\left[\|r_n\|^2 - \frac{\|F_n'^T r_n\|^4}{\|p_n\|^2 (M^2 + B\|r_n\|)}\right].$$

Now using $m^2\|r_n\|^2 \leqslant \|F_n'^T r_n\|^2$ [from Lemma 3.1 (iv)] we obtain

$$\frac{1}{2}\|r_{n+1}\| \leqslant \frac{1}{2}\left[1 - m^2 \frac{\|F_n'^T r_n\|^2}{\|p_n\|^2 (M^2 + B\|r_n\|)}\right]\|r_n\|^2$$

If we substitute the fraction term in the square brackets by the left most term in (3.10), we prove that the norm of the residual is reduced (at each iteration) by a constant factor that is less than one. This implies that $\|r_n\|$ converges to zero.

Finally, we prove that the sequence of iterates converges to a unique solution of the nonlinear operator equation. By use of (3.2) with $x = x_n$ and $y = x_{n+k}$, we obtain that the sequence x_n is a Cauchy sequence. Thus, it converges to x^* and $F(x^*) = 0$. The uniqueness and the error bound inequality in the theorem statement follow from (3.2) with $x = x_n$ and $y = x^*$. □

4. THE NONLINEAR CGNE METHOD

Let us assume that (3.1a) and (3.1b) hold in this section. From Theorem 3.1, it follows that a unique solution of $F(x) = 0$ exists. Next, we introduce a nonlinear version of CGNE, and we prove a local convergence theorem.

ALGORITHM 4.1. The Nonlinear CGNE Algorithm.
Initial vector x_0
$r_0 = -F(x_0)$, $p_0 = F_0^T r_0$

For $n = 0$ **Until** Convergence **Do**

1. Select the smallest positive c_n to minimize $\|x^* - (x_n + cp_n)\|_2$, $c > 0$
2. $x_{n+1} = x_n + c_n p_n$ and $r_{n+1} = -F(x_{n+1})$
3. $b_n = -\dfrac{(F_{n+1}^T r_{n+1}, p_n)}{\|p_n\|^2}$ where $p_{n+1} = F_{n+1}^T r_{n+1} + b_n p_n$

EndFor

REMARK 4.0. The error function in step 1 of the algorithm is not computable because it uses the exact solution. However, it is possible to determine an approximation to c_n that guarantees local convergence of the algorithm to the solution.

Let us denote the true error $x^* - x_n$ by e_n. The scalars c_n and b_n by definition imply the following two orthogonality relations:

$$(e_{n+1}, p_n) = 0 \tag{4.1}$$

and

$$(p_{n+1}, p_n) = 0. \tag{4.2}$$

Under the assumptions (3.1) the following lemma holds for Algorithm 4.1.

LEMMA 4.1. *Let* $\{r_n\}$ *be the nonlinear residuals and* $\{p_n\}$ *be the direction vectors in Algorithm* 4.1 *then the following identities hold true*:

(*i*) $(p_n, e_n) = (F_n^T r_n, e_n)$
(*ii*) $\|F_n^T r_n\|^2 = \|p_n\|^2 + b_{n-1}^2 \|p_{n-1}\|^2$
(*iii*) $M\|r_n\| \geqslant \|p_n\|$
(*iv*) $M\|e_n\| \geqslant \|r_n\| \geqslant m\|e_n\|$
(*v*) $\|e_{n+1}\| \leqslant \|e_n\|$

PROOF. (i) follows from relation (4.1) and equality 5 of Algorithm (4.1). We prove (ii) from equality 5 of Algorithm (4.1) and relation (4.2). Part (iii) follows from (ii) and (3.1). By using equalities (3.2) and (3.3) (at the beginning of section 3) with $y = x_n$ and $x = x^*$ we prove (iv). Part (v) follows from the selection of c_n. □

REMARK 4.1. Let us denote by $f_n(c)$ the scalar function

$$\frac{1}{2}\|x^* - (x_n + cp_n)\|^2 = \frac{1}{2}\|e_n - cp_n\|^2.$$

The first and second derivatives are:

$$f_n'(c) = c\|p_n\|^2 - (p_n, e_n), f_n''(c) = \|p_n\|^2 \tag{4.3}$$

We expand $F(x^*) = 0$ in Taylor series around x_n to obtain:

$$r_n = F_n e_n + (F_{\tilde{n}}'' e_n, e_n), \tag{4.4}$$

where $\tilde{n} = x_n + te_n$. Using Lemma 4.1 (i) and (4.3), we obtain:

$$f_n'(c) = c\|p_n\|^2 + \|r_n\|^2 - (r_n, (F_{\tilde{n}}'' e_n, e_n)). \tag{4.5}$$

We next prove that under the assumptions (3.1) that Algorithm (4.1) converges locally to the unique solution of the nonlinear operator equation.

THEOREM 4.1. *Assume that conditions* (3.1) *hold. Also, assume that* x_0 *is selected such that* $\|F(x_0)\| < \frac{m^2}{2B}$. *Then the sequence* x_n *generated by Algorithm* (4.1) *is well-defined converges to the unique solution* x^* *of the nonlinear operator equation* $F(x) = 0$.

PROOF. Firstly, we prove the existence of the nonlinear steplengths c_n. It suffices to prove that the first derivative of $f_n(c)$ is negative at $c = 0$ and its second derivative is positive in an interval $[0, c_1)$. By using Lemma 4.1 (iv) and (v) and the assumption of the theorem we prove the following inequality:

$$|(r_n, (F_{\tilde{n}}'' e_n, e_n))| \leqslant \frac{B}{m}\|r_n\|^2\|e_0\| \leqslant \frac{B\|r_0\|}{m^2}\|r_n\|^2 < \frac{1}{2}\|r_n\|^2. \tag{4.6}$$

Combining (4.5) and (4.6), we conclude that $f_n'(0)$ is negative. Also, (4.3) and (4.6) imply that

$$|f_n'(0)| = |(e_n, p_n)| > \frac{1}{2}\|r_n\|^2. \tag{4.7}$$

Now, using Lemma 4.1 (iv), we obtain:

$$\|p_n\| \geqslant \frac{m}{2}\|r_n\|. \tag{4.8}$$

This indequality shows that the second derivative of $f_n(c)$ is positive if convergence has not been reached (i.e., $r_n \neq 0$).

Secondly, we obtain a lower bound on the nonlinear steplength c_n. Since $f_n(c_n) = 0$, we used (4.5) and (4.7) to obtain

$$c_n = \frac{-(p_n, e_n)}{\|p_n\|^2} \geqslant \frac{\|r_n\|^2}{2\|p_n\|^2} \geqslant \frac{1}{2M^2}.$$

We insert $c = \frac{1}{2M^2}$ in $f_n(c) = \frac{1}{2}\|e_n - cp_n\|^2$ and we use Lemma 4.1 (iii) and (iv) to obtain the following error bound:

$$\|e_{n+1}\|^2 = 2f(c_n) = \|e_n\|^2 - \frac{\|r_n\|^2}{2M^2} + \frac{\|p_n\|^2}{4M^4} \leqslant \|e_n\|^2 \frac{(1-m^2)}{4M^2}$$

This proves the convergence of the iteration and gives also a bound on the factor of the linear convergence. □

5. ASYMPTOTIC STEPLENGTH ESTIMATES AND ERROR BOUNDS

In this section, we obtain asymptotic estimates of the steplengths c_n near the solution. We also obtain an asymptotic error factor estimate. We only obtain these results for Algorithm 4.1. Similar results can be obtained for Algorithm 3.1.

We next obtain asymptotic estimates of the steplengths c_n under the assumptions of Theorem 4.1.

PROPOSITION 5.1.

$$\frac{\|r_n\|^2}{\|p_n\|^2}\frac{1}{(1+\varepsilon_n)} \leqslant c_n \leqslant \frac{\|r_n\|^2}{\|p_n\|^2}\frac{1}{(1-\varepsilon_n)},$$

where $\varepsilon_n = O(\|r_n\|)$.

PROOF. We will prove only the rightmost inequality. The leftmost inequality is proved similarly. From equality (4.5) and $f_n'(c_n) = 0$, we obtain the following inequality:

$$c_n \leqslant \frac{\|r_n\|^2}{\|p_n\|^2 - B\|r_n\|\,\|e_n\|^2} = \frac{\|r_n\|^2}{\|p_n\|^2(1 - \varepsilon_n)},$$

where ε_n is the fraction $\dfrac{B\|r_n\|\,\|e_n\|^2}{\|p_n\|^2}$. Now, we use Lemma 4.1 (iv) and inequality (4.8) to obtain:

$$\varepsilon_n \leqslant \frac{4B}{m^4}\|r_n\|.$$

□

We next obtain an asymptotic error bound for iterates in Algorithm 4.1.

PROPOSITION 5.2. *Under the assumptions of Theorem* 4.1, *we obtain the following inequality on the residual errors*:

$$\|e_{n+1}\|^2 \leqslant \|e_n\|^2 d_n,$$

where

$$d_n = \left[1 - \frac{m^2}{M^2} + \sigma_n\right]$$

and

$$\sigma_n = 2M^2\|r_n\|.$$

PROOF. We note that by using relation (4.1) and Lemma 4.1 (i) we obtain:

$$\|e_{n+1}\|^2 = (e_{n+1}, e_n - c_n p_n) = -c_n(e_n, p_n) = -c_n(r_n, F_n' e_n)$$

Now using equality (4.4) and Lemma 4.1 (iv) we obtain:

$$\|e_{n+1}\|^2 - \|e_n\|^2 \leqslant -c_n\|r_n\|^2 + B\|r_n\|\,\|e_n\|^2. \tag{5.1}$$

Using Proposition 5.1 and Lemma 4.1 (iii), we prove the following inequality:

$$c_n\|r_n\|^2 \geqslant \frac{\|r_n\|^4}{\|p_n\|^2}\frac{1}{(1+\varepsilon_n)} \geqslant \frac{m^2}{M^2}\|e_n\|^2\frac{1}{(1+\varepsilon_n)} \tag{5.2}$$

Now using (5.2) in (5.1) we obtain:

$$\|e_{n+1}\|^2 \leqslant \|e_n\|^2\left[1 - \frac{m^2}{M^2}\frac{1}{(1+\varepsilon_n)}\right] + B\|r_n\|\,\|e_n\|^2$$

The last term in this inequality is less than

$$\|e_n\|^2\left[1 - \frac{m^2}{M^2} + \sigma_n\right],$$

where

$$\sigma_n = B\|r_n\|.$$

□

6. CONCLUSIONS

We have presented and analyzed nonlinear generalizations of the CGNR and CGNE method. These nonlinear methods apply to nonlinear operator equations with nonsymmetric Jacobian. We show that under certain uniform assumptions on the Jacobians and Hessians the nonlinear CGNR is guaranteed to converge globally to a unique solution. For the nonlinear CGNE under the same assumptions as CGNR, we prove local convergence results and give asymptotic error bound estimates. These results extend the work of other authors [4, 8] to deriving nonlinear methods for nonsymmetric Jacobians.

The research was partially supported by NSF under grant CCR-8722260.

A conference proceedings version of this paper appeared in the proceedings of the S.E. Conference on Theory of Approximation, March 1991, Marcel-Dekker, *Pure and Applied Mathematics*, Vol. 138, edited by George A. Anastassion.

REFERENCES

1 S. F. Ashby, T. A. Manteuffel, and P. E. Saylor, A taxonomy for conjugate gradient methods, *SIAM J. Numer. Anal.* 27:1542–1568 (1990).

2 P. N. Brown, A local convergence theory for combined inexact-Newton/finite-difference projection methods, *SIAM J. Numer. Anal.* 24:610–638 (1987).

3 E. J. Craig, The N-step iteration procedures, *J. Math. Phys.* 34:64–73 (1955).

4 J. W. Daniel, The conjugate gradient method for linear and nonlinear operator equations, *SIAM J. Numer. Anal.* 4:10–26 (1967).

5 J. W. Daniel, *The Approximate Minimization of Functionals*, Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

6 R. S. Dembo, S. C. Eisenstat, and T. Steighaug, Inexact Newton methods, *SIAM J. Numer. Anal.* 19:400–408 (1982).

7 S. C. Eisenstat, H. C. Elman, and M. H. Schultz, Variational iterative methods for nonsymmetric systems of linear equations, *SIAM J. Numer. Anal.* 20:345–357 (1983).

8 R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, *Compt. J.* 7:149–154 (1964).

9 R. Fletcher, *Practical Methods in Optimization, Vol. 2, Unconstrained Optimization*, Wiley, Chichester, 1980.

10 G. H. Golub and R. Kannan, Convergence of a two-stage Richardson process for nonlinear equations, *BIT* 209–219 (1986).

11 M. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards* 49:409–436 (1952).

12 M. Z. Nashed, On general iterative methods for the solutions of a class of nonlinear operator equations, *Math. Comp.* 19:14–24 (1965).

13 M. Z. Nashed, The convergence of the method of steepest descent for nonlinear equations with variational or quasi-variational operators, *J. Math. Mech.* 13:765–794 (1964).

14 T. L. Saaty, *Modern Nonlinear Equations*, Dover Publications Inc., New York, 1981.

15 D. P. O'leary, A discrete Newton Algorithm for minimizing a function of many variables, *Math. Programming* 23:20–33 (1982).