CONVERGENCE OF THE CONJUGATE
GRADIENT METHOD: A CORRECTION
AND A DIFFERENT APPROACH

by

James W. Daniel
Technical Report #52

December 1968

1. **INTRODUCTION**

It has been observed by J. Ortega and W. Rheinboldt [3] that there is an error in this author's previously published paper [1]. In particular, the fifth line from the bottom of page 22 is false since the inductive method being used at that point breaks down at \( i = 0 \); thus the Theorem 2.1.2 has not been proved. Since the application of this theorem to finite dimensional problems yields the only known results concerning superlinear convergence of the conjugate gradient method for nonlinear equations [2], the theorem is worth further investigation. In this note we prove the theorem correctly using a method which, when slightly generalized, appears to give some hope of proving superlinear convergence for linear or nonlinear equations in infinite dimensional spaces.

2. **A VIEW OF THE METHOD**

We use the notation in [1,2] without further explanation; for convenience we restrict ourselves to the simplest form of the CG method, although the results apply in total more generally.

We seek to solve the equation

\[
Mx = k
\]

in a real, separable Hilbert space \( \mathcal{H} \), where \( M \) is a bounded, self-adjoint, positive definite linear operator from \( \mathcal{H} \) onto \( \mathcal{H} \), having a
All the known theory of the CG method for linear equations applies here and we can in particular deduce that

\[ E_1(x_n) \leq \omega_n^2 E_1(x_0) \]

where

\[ \omega_n = \frac{2(1 - \alpha_n)^n}{(1 + \sqrt{\alpha_n})^{2n} + (1 - \sqrt{\alpha_n})^{2n}} \]

\[ \alpha_n = \frac{\text{a}_n}{\text{A}_n} \]

\[ E_1(x) = [h - x, h - x] = \langle h - x, M(h - x) \rangle. \]

A straightforward calculation shows that the iterates \( x_n \) generated by this general algorithm on \( \mathcal{M}_1 \) in \( \mathcal{H}_1 \) are precisely the same as the iterates generated by using the standard simple algorithm on \( \mathcal{M} \) in \( \mathcal{H} \) if the initial direction \( p_0 \) in the simple algorithm is not chosen as \( r_0 = k - Mx_0 = k \) as usual but by the formula

\[ p_0 = P_I r_0 = r_0 + b_{-1} d, \quad b_{-1} = -\frac{\langle r_0, M d \rangle}{\langle d, M d \rangle}, \]

that is, by the usual way of generating CG directions if we identify \( d \) with \( p_{-1} \).

All that this says is that the standard CG method, modified to require the first direction \( p_0 \) to be conjugate to \( d \), is equivalent to a general CG method in a space \( M \)-conjugate to \( d \); therefore the modified standard method converges and in fact, since \( E_1(x) = [h - x, h - x] = \)
\(<h - x, M(h - x)\rangle \equiv E(x)\) in standard notation,

\[E(x_n) \leq \omega_n^2 E(x_0).\]

More generally, if we have proceeded through standard CG directions \(p_0, p_1, \ldots, p_{L-1}\) to arrive at \(x_L = 0\), then the solution \(h\) is \(M\)-conjugate to \(p_i\), \(0 \leq i \leq L-1\), and we can define \(P_F\) as the orthogonal projection (in the \([\cdot, \cdot]\) sense) onto the span of \(\{p_0, \ldots, p_{L-1}\} \), \(P_I = I - P_F\), \(M_1 = P_I M\). Then the remainder of the standard CG iterates are precisely the same as those generated by the more general CG method applied to \(M_1\) in \(H_1\) and therefore our convergence estimates can make use of the spectral bounds of \(M_1\) on \(H_1\) rather than of \(M\) on \(H\). Since the projections \(P_I\) are "contracting" as we do this analysis after each new standard CG step, the spectral bounds on the operators \(M_1\) might be contracting, allowing a proof of superlinear convergence. While we have not been successful in accomplishing this, it seems a worthwhile approach.

3. THE CORRECTED PROOF

Let \(J\) be a continuous nonlinear operator form \(H\) into \(H'\), satisfying \(0 < aI \leq J_x \leq aI, \|J''_x\| \leq B, \ J'_x\) self adjoint, for all \(x \in S(x_0, R_0), R_0 = (\sqrt{A/a}(1 - q_0)) \varepsilon_0, x_0\) such that \(q_0^2 \equiv q^2 + \sigma_0 < 1, q = \frac{A-a}{A+a}\). Then for any \(m\) there exists an integer \(n_m\) such that for all \(n \geq n_m\), we have

\[E_n(x_{n+m}) \leq \left(\omega_m^2 + \delta_n\right) E_n(x_n)\]
where \( \delta_n \) tends to zero, \( \omega_m = \frac{2(1 - \alpha)^m}{(1 + \sqrt{\alpha})^{2m} + (1 - \sqrt{\alpha})^{2m}} \), \( \alpha = \frac{a}{A} \),

\[
E_n(x) = \langle r_n, J_n^{-1} r_n \rangle, \quad r_n = r(x) = -J(x), \text{ and the } x_n \text{ are the iterates generated by the standard CG method to solve } J(x) = 0.
\]

**Proof:** Consider the iterate \( x_n \) and the linear equation \( J_n z_n = J_n x_n + r_n \) for \( z_n \), having solution \( h_n \equiv x_n + J_n^{-1} r_n \). We note that \( h_n - x_n \) is \( J_n \)-conjugate to \( p_{n-1} \). If we consider the standard CG method to compute \( z = h_n \) starting with \( z_0 = x_n \) but requiring that the first direction \( \tilde{p}_0 \) be \( J_n \)-conjugate to the given direction \( d = p_{n-1} \), we have precisely the situation discussed in the preceding section. Therefore the sequence of such iterates \( z_i \) converges to \( h_n \) and, since \( \alpha_1 \geq \alpha \),

\[
\langle h_n - z_m, J_n (h_n - z_m) \rangle \leq \omega_m \langle r_n, J_n^{-1} r_n \rangle.
\]

The first direction \( \tilde{p}_0 \) in the modified method is the projection of \( J_n x_n + r_n - J_n z_0 = r_n \) onto the \( J_n \)-conjugate complement of \( p_{n-1} \), that is,

\[
\tilde{p}_0 = p_n.
\]

Recall that \( \epsilon_n^2 \equiv E_n(x_n) \).

If we show that

\[
|\langle h_n - z_m, J_n (h_n - z_m) \rangle - E_{n+m}(x_{n+m})|,
\]

which equals

\[
|\langle h_n - z_m, J_n (h_n - z_m) \rangle - \langle h_{n+m} - x_{n+m}, J_{n+m} (h_{n+m} - x_{n+m}) \rangle|,
\]

is of order \( \epsilon_n^3 \), then we will have
\[ E_{n+m}(x_{n+m}) = \langle h_n - z_m, J_n'(h_n - z_m) \rangle \]

\[ + \left[ E_{n+m}(x_{n+m}) - \langle h_n - z_m, J_n'(h_n - z_m) \rangle \right] \]

\[ \leq (\varepsilon_n^2 + O(\varepsilon_n)) E_n(x_n). \]

We indicate the proof of the order of magnitude. The sum to be estimated splits into

\[ | \langle h_n - z_m, (J_n' - J_{n+m}')(h_n - z_m) \rangle | \]

and

\[ | \langle h_n - h_{n+m} + x_{n+m} - z_m, J_{n+m}'(h_n - z_m + h_{n+m} - x_{n+m}) \rangle | , \]

the first of which is less than

\[ B \| h_n - z_m \|^2 \| x_{n+m} - x_n \| = O(\varepsilon_n^3) , \]

by (2.1.1) and the proof of Theorem 2.1.1 (for big n) in [1]. Clearly the second part of the sum is less than

\[ \| h_n - h_{n+m} + x_{n+m} - z_m \| O(\varepsilon_n) ; \]

we estimate the normed term. First

\[ \| h_n - h_{n+m} \| = \sum_{i=0}^{m-1} \| h_{n+1} - h_{n+i+1} \| , \]

while

\[ \| h_{j+1} - h_j \| = \| x_{j+1} - x_j + J_{j+1}' r_{j+1} - J_j' r_j \| \]

\[ = \| c_j p_j + J_{j+1}'(r_{j+1} - r_j) + (J_{j+1}' - J_j') r_j \| = O(\varepsilon_n^2) , \]

since

\[ r_{j+1} - r_j = -J_{j+1}'(c_j p_j) + O(\varepsilon_n^2) . \]

We still must estimate

\[ \| x_{n+m} - z_m \| = \| x_{n+m} - c_{n+m-1} p_{n+m-1} - z_m - \tilde{c}_{m-1} \tilde{p}_{m-1} \| , \]
where the $\sim$ indicates the $z_i$ iteration. Since $\tilde{p}_0 = p_n$, a simple inductive argument yields
\[
\|c_{n+i} p_{n+i} - \tilde{c}_{i} \tilde{p}_{i}\| = O(\varepsilon_n^2)
\]
for all $i$, which leads to
\[
\|x_{n+m} - z_m\| = O(\varepsilon_n^2).
\]
Q.E.D.
REFERENCES


3. Ortega, J., Rheinboldt, W., private communication.
**REPORT TITLE**

CONVERGENCE OF THE CONJUGATE GRADIENT METHOD: A CORRECTION AND A DIFFERENT APPROACH

**AUTHOR(S)**

James W. Daniel

**REPORT DATE**

December 1968

**CONTRACT OR GRANT NO.**

N00014-67-A-0128-0004

**PROJECT NO.**

c.

d.

**DISTRIBUTION STATEMENT**

Releasable without limitations on dissemination

**ABSTRACT**

An error in a convergence proof for the conjugate gradient method is corrected.
<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROLE</td>
<td>WT</td>
<td>ROLE</td>
</tr>
<tr>
<td>Conjugate gradient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonlinear operator equations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear operator equations</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>