AN ALTERNATING GRADIENT METHOD FOR SOLVING LINEAR SYSTEMS

by

EVERETT CRAMER RIGGLE

A THESIS submitted to
OREGON STATE COLLEGE

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

June 1959
APPROVED:

[Signature]
Associate Professor of Mathematics
In Charge of Major

[Signature]
Chairman of Department of Mathematics

[Signature]
Chairman of School of Science Graduate Committee

[Signature]
Dean of Graduate School

Date thesis is presented August 11, 1958

Typed by Everett Riggle
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. DEVELOPMENT OF FORMULAS</td>
<td>4</td>
</tr>
<tr>
<td>III. PROOF OF BASIC PROPERTIES</td>
<td>9</td>
</tr>
<tr>
<td>IV. ITERATIVE ROUTINE</td>
<td>16</td>
</tr>
<tr>
<td>V. EXAMPLES</td>
<td>18</td>
</tr>
<tr>
<td>VI. CORRECTION OF ROUND-OFF ERRORS</td>
<td>21</td>
</tr>
<tr>
<td>VII. TABLES</td>
<td>24</td>
</tr>
<tr>
<td>VIII. BIBLIOGRAPHY</td>
<td>38</td>
</tr>
</tbody>
</table>
AN ALTERNATING GRADIENT METHOD
FOR SOLVING LINEAR SYSTEMS

INTRODUCTION

Various gradient methods for solving systems of linear equations have been presented by Temple (5, p. 476-500), Stein (4, p. 407-413), Forsythe and Motzkin (1, p. 304-305), Hestenes and Stiefel (2, p. 409-436), and others. However, at present, there is no best method for the solution of systems of linear equations, as each method depends to some extent upon the particular system to be solved.

In the present paper, a modified gradient method is exploited to obtain the solution, if it exists, of m linear equations in n unknowns. If the solution does not exist, an indication is obtained by the process of solution.

Let the given system of equations be denoted in matrix notation by $Ax - b = 0$, where $A$ is any $m \times n$ matrix, $b$ and $0$ are column vectors with $m$ components, and $x$ is a column vector with $n$ components.

Define a function of $x$ as the sum of the absolute values of the components of $Ax - b$. That is,

$$ \phi(x) = |(Ax - b)_1| + |(Ax - b)_2| + \ldots + |(Ax - b)_m| . $$

This positive definite function has the value zero if and only if $x$ is the solution to the given system.
\( \phi = \text{constant} \), represents a family of surfaces in \( n \) dimensions, with \( x \) as its center. Given any arbitrary point, determined by the position vector \( x_0 \), that point lies on some surface of \( \phi = c_1 \). The problem of finding a solution to the system of equations now becomes one of finding the center of the surfaces.

This will be accomplished in the following manner: The negative of the gradient of the function at \( x_0 \) will be used as a first direction vector \( v_0 \). Proceeding in the direction of \( v_0 \), a point \( x_1 \) is reached such that the distance, \( |x-x_1| \), is a relative minimum with respect to this direction. The point \( x_1 \) lies in a hyperplane which is orthogonal to \( v_0 \) and passes through the center \( x \). It also lies on some surface of \( \phi = c_2 \), hence, the gradient of the function at this point can be determined. The second direction vector \( v_1 \) is made orthogonal to \( v_0 \) by taking a linear combination of \( \nabla \phi \) and \( v_0 \). As before, the direction of the vector \( v_1 \) is followed until a point \( x_2 \) is reached such that the distance, \( |x-x_2| \), is a relative minimum with respect to this direction. This procedure is continued, obtaining a set of mutually orthogonal direction vectors \( (v_0, v_1, \ldots) \) and successive new estimates \( x_1, x_2, \ldots \). It will be
proved later, that if a solution exists, the $M$th estimate, $x_M$, will be the solution $x$, for some value of $M \leq \min(m,n)$. 
To develop the formula for \( \text{grad} \ \phi \), the function 
\[
\phi(x) = |a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1| + \cdots + |a_{m1}x_1 + \cdots + a_{mn}x_n - b_m|
\]
The gradient of this function is the column vector
\[
\text{grad} \ \phi = \begin{bmatrix} \frac{\partial \phi}{\partial x_1} \\ \frac{\partial \phi}{\partial x_2} \\ \vdots \\ \frac{\partial \phi}{\partial x_n} \end{bmatrix}
\]
\[
= \begin{bmatrix} a_{11} \text{sgn}(a_{11}x_1 + \cdots + a_{1n}x_n - b_1) + \cdots + a_{m1} \text{sgn}(a_{m1}x_1 + \cdots + a_{mn}x_n - b_m) \\ a_{12} \text{sgn}(a_{11}x_1 + \cdots + a_{1n}x_n - b_1) + \cdots + a_{m2} \text{sgn}(a_{m1}x_1 + \cdots + a_{mn}x_n - b_m) \\ \vdots \\ a_{1n} \text{sgn}(a_{11}x_1 + \cdots + a_{1n}x_n - b_1) + \cdots + a_{mn} \text{sgn}(a_{m1}x_1 + \cdots + a_{mn}x_n - b_m) \end{bmatrix}
\]
For any vector \( x_i \), \( Ax_i - b = r_i \). That is,
\[
a_{11}x_{i1} + a_{12}x_{i2} + \cdots + a_{1n}x_{in} - b_1 = r_{i1} \\
a_{21}x_{i1} + a_{22}x_{i2} + \cdots + a_{2n}x_{in} - b_2 = r_{i2} \\
\vdots \\
a_{m1}x_{i1} + a_{m2}x_{i2} + \cdots + a_{mn}x_{in} - b_m = r_{im}
\]
Then, \( \text{grad } \phi_i = \begin{bmatrix} a_{11} \text{sgn } r_{il} + \cdots + a_{1m} \text{sgn } r_{im} \\ a_{21} \text{sgn } r_{il} + \cdots + a_{2m} \text{sgn } r_{im} \\ \vdots \\ a_{n1} \text{sgn } r_{il} + \cdots + a_{nm} \text{sgn } r_{im} \end{bmatrix} \).

1) \( \text{grad } \phi_i = A^T \text{sgn } r_i \).

The notation \( \text{sgn } r_i \) represents the vector whose components are \( \text{sgn } r_{ij} \) \((j=1,2,\ldots,m)\).

The direction vectors are developed by defining \( v_0 = \text{grad } \phi_0 \) and \( v_1 \) as a linear combination of \( \text{grad } \phi_1 \) and the previous direction vectors. That is,

2) \( v_i = \text{grad } \phi_i + e_i, i-lv_{i-1} + e_i, i-2v_{i-2} + \cdots + e_i, o v_0 \).

The direction vectors are to be mutually orthogonal, hence, \( v_i \cdot v_k = 0 \) \((i \neq k)\).

Then, \( v_i \cdot v_k = \text{grad } \phi_i \cdot v_k + e_i, i-lv_{i-1} \cdot v_k + \cdots + e_i, o v_0 \cdot v_k = 0 \).

Each term of this equation is zero except for the first term and the term containing \( v_k \cdot v_k \). Thus,

\[ \text{grad } \phi_i \cdot v_k + e_i, k v_k \cdot v_k = 0; \]

3) \( e_{i,k} = \frac{-\text{grad } \phi_i \cdot v_k}{v_k \cdot v_k} = \frac{-\text{grad } \phi_i^T v_k}{v_k^T v_k}, \quad (k=0,1,\ldots,i-1). \)

As previously stated, the initial estimate \( x_0 \) of the solution \( x \), is to be arbitrary. Then, the successive new estimates \( x_1, x_2, \ldots \) are obtained by the formula,

4) \( x_{i+1} = x_i - \lambda_i v_i \).
where \( \lambda_i \) is picked such that the distance, \(|x-x_{i+1}|\), will be a minimum with respect to the direction \( v_i \).

That is, \( v_i \) will be orthogonal to the vector \( x_{i+1}-x \).

Geometrically, this can be represented as follows:

\[
\begin{align*}
&\text{Let } d \text{ represent the distance } |x_{i+1}-x_i|, \text{ then,} \\
&d = \frac{(x_{i+1}-x) \cdot v_i}{|v_i|}.
\end{align*}
\]

The desired vector \( \lambda_i v_i \) which is represented in the diagram as \( x_i-x_{i+1} \), can be written as \( d \) multiplied by the normalized vector \( v_i \). That is,

\[
\frac{\lambda_i v_i}{|v_i|} = \frac{(x_{i+1}-x) \cdot v_i}{|v_i|} = \frac{(x_{i+1}-x) \cdot v_i}{v_i \cdot v_i} v_i;
\]

and

\[
\lambda_i = \frac{(x_{i+1}-x) \cdot v_i}{v_i \cdot v_i}.
\]

Introduce now a vector \( u_i \), such that,

\[
5) \quad v_i = A^T u_i.
\]

Then,

\[
\lambda_i = \frac{- (x-x_i)^T A^T u_i}{v_i \cdot v_i} = \frac{[-A(x-x_i)]^T u_i}{v_i \cdot v_i}.
\]

But,

\[
Ax-b=0,
\]

and for any vector \( x_i \), \( Ax_i-b=r_i \);
6) \[ A(x-x_i)=-r_i. \]

Making this substitution, we now have,
\[ \lambda_i = \frac{r_i^T u_i}{v_i v_i}. \]

However, this can be written in a more useful form as follows:
\[ x_i = x_{i-1} - \lambda_{i-1} v_{i-1}; \]
\[ x_{i-1} = x_{i-2} - \lambda_{i-2} v_{i-2}; \]
\[ \vdots \]
\[ x_i = x_0 - \lambda_0 v_0. \]

Hence,
\[ x_i = x_0 - \lambda_0 v_0 - \lambda_1 v_1 - \cdots - \lambda_{i-1} v_{i-1}; \]
\[ (x_i - x_0) \cdot v_i = 0; \]
\[ (x_i - x_j - x_0) \cdot v_i = 0; \]
\[ (x - x_i) \cdot v_i = (x - x_0) \cdot v_i; \]
\[ r_i \cdot u_i = r_0 \cdot u_1. \]

Thus,
7) \[ \lambda_i = \frac{r_i^T u_i}{v_i v_i}. \]

It is necessary to have \( u_i \) to determine \( \lambda_i \).

However, it is possible to determine \( u_i \) recursively as follows:

Since \[ v_0 = \text{grad} \ \phi_0 = A^T \text{sgn} \ r_0, \]
we have for \( u_0 \), the vector \( \text{sgn} \ r_0 \). If \( u_1 \) and hence
$v_j$ are determined up through the value $i-1$, we have,

$$v_i = \text{grad } \phi_i + e_i, i-1 v_{i-1} + e_i, i-2 v_{i-2} + \ldots + e_i, o v_o;$$

$$v_i = A^T\text{sgn } r_i + e_i, i-1 A^T u_{i-1} + e_i, i-2 A^T u_{i-2} + \ldots + e_i, o A^T u_o;$$

$$v_i = A^T(\text{sgn } r_i + e_i, i-1 u_{i-1} + e_i, i-2 u_{i-2} + \ldots + e_i, o u_o).$$

Thus the necessary recurrence relation for $u_i$ is defined by:

3) \hspace{1cm} u_i = \text{sgn } r_i + e_i, i-1 u_{i-1} + e_i, i-2 u_{i-2} + \ldots + e_i, o u_o.
Proof of the basic properties of this method are given in the following theorems.

Theorem I.

The set of vectors \( v_0, v_1, \ldots \) form a mutually orthogonal set.

The proof will be by induction.

By formula 2),

\[
v_1 = \text{grad } \phi_1 + e_1, v_0;
\]

\[
v_1 \cdot v_0 = \text{grad } \phi_1 \cdot v_0 + e_1, v_0 \cdot v_0;
\]

but, by formula 3),

\[
e_1, v_0 = \frac{-\text{grad } \phi_1, v_0}{v_0 \cdot v_0};
\]

hence \( v_1 \cdot v_0 = 0 \), which implies \( v_1 \) is orthogonal to \( v_0 \).

Assume that the direction vector \( v_{k-1} \) is orthogonal to all preceding direction vectors, that is,

\[
v_{k-1} \cdot v_j = 0 \quad (j=0,1,\ldots,k-2).
\]

By formula 2),

\[
v_k = \text{grad } \phi_k + e_k, v_{k-1} \cdot v_{k-1} + \cdots + e_k, v_0;
\]

\[
v_k \cdot v_{k-1} = \text{grad } \phi_k \cdot v_{k-1} + e_k, v_{k-1} \cdot v_{k-1} + \cdots + e_k, v_0 \cdot v_{k-1};
\]

but, by formula 3),

\[
e_k, v_{k-1} = \frac{-\text{grad } \phi_k, v_{k-1}}{v_{k-1} \cdot v_{k-1}};
\]

hence,

\[
v_k \cdot v_{k-1} = 0.
\]
Similarly, for case \(0 \leq j < k-1\),

\[
v_k \cdot v_j = \nabla \phi_k \cdot v_j + e_{k, k-l} v_{k-l} \cdot v_j + \cdots + e_{k, o} v_o \cdot v_j;
\]

\[
v_k \cdot v_j = \nabla \phi_k \cdot v_j + e_{k, j} v_j \cdot v_j;
\]

\[
e_{k, j} = \frac{-\nabla \phi_k \cdot v_j}{v_j \cdot v_j};
\]

therefore,

\[
v_k \cdot v_j = 0.
\]

Hence,

\[
v_k \cdot v_j = 0.
\]

Therefore, it is proved by induction that for all \(k\),

\(v_k\) is orthogonal to all preceding direction vectors,

and \(v_0, v_1, \ldots\) form a mutually orthogonal set.

**Corollary:**

Since they are orthogonal, if none of the vectors

\(v_0, v_1, \ldots, v_j\) is the zero vector, they are linearly independent.

**Theorem II.**

If a solution exists, \(v_i \neq 0\) when \(r_i \neq 0\), for \(i \leq R-1\).

The proof will be by contradiction. Assume a solution exists when some \(v_i = 0\) and \(r_i \neq 0\). Let \(s\) be the first positive integer such that \(v_s = 0\) when \(r_s \neq 0\).

By formula 5),

\[
v_s = A^T u_s = 0.
\]

This leads to two cases to be considered; \(A^T u_s = 0\) when \(u_s = 0\) and \(A^T u_s = 0\) when \(u_s \neq 0\).
Case I. \( A^T u_s = 0 \) when \( u_s = 0 \).

By formula 8),
\[
u_s = \text{sgn} \left( r_s + e_s, s-1 u_{s-1} + \cdots + e_s, o u_0 \right).
\]

Taking the inner product of both sides of the equation with the vector \( r_s \),
\[
u_s \cdot r_s = \text{sgn} \left( r_s + e_s, s-1 u_{s-1} \cdot r_s + \cdots + e_s, o u_0 \cdot r_s \right).
\]

However, \( Ax_s - b = r_s \),
and by formula 4),
\[
x_s = x_{s-1} - \lambda_{s-1} v_{s-1}.
\]

Substituting, we have,
\[
Ax_{s-1} - \lambda_{s-1} Av_{s-1} - b = r_s.
\]

But,
\[
Ax_{s-1} - b = r_{s-1};
\]

hence,
\[
r_s = r_{s-1} - \lambda_{s-1} Av_{s-1}.
\]

Similarly,
\[
r_{s-1} = r_{s-2} - \lambda_{s-2} Av_{s-2};
\]
\[
r_{s-2} = r_{s-3} - \lambda_{s-3} Av_{s-3};
\]
\[
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldot
For $j=0,1,\ldots,s-1$ every term on the right side of the equation is zero except for the first term and the term containing $v_j \cdot v_j$. Hence, $u_j \cdot r_s = u_j \cdot r_0 - \lambda_j v_j \cdot v_j$.

By formula 7),

$$\lambda_j = \frac{r_0 \cdot u_j}{v_j \cdot v_j}$$

and $u_j \cdot r_s = 0$ for $j=0,1,\ldots,s-1$. Making this substitution into equation 9), from the previous page,

$$u_s \cdot r_s = \text{sgn} \, r_s \cdot r_s.$$

By hypothesis $r_s \neq 0$, hence, $\text{sgn} \, r_s \cdot r_s > 0$.

But if $u_s = 0$, $u_s \cdot r_s = 0$. This is a contradiction, thus, $u_s \neq 0$.

Case II. $A^T u_s = 0$ when $u_s \neq 0$.

This implies that the vector $u_s$ is orthogonal to each of the column vectors of $A$. If $A$ is of rank $R$, there exists $R$ column vectors of $A$, $(\alpha_1, \alpha_2, \ldots, \alpha_R)$, which are linearly independent. Therefore, the vectors $(\alpha_1, \alpha_2, \ldots, \alpha_R, u_s)$ form a linearly independent set.

By formula 3),

$$u_s = \text{sgn} \, r_s \cdot u_s + e_s, s-1 u_s - 1 + \cdots + e_s, 0 u_0;$$

$$u_{s-1} = \text{sgn} \, r_{s-1} \cdot u_{s-1} + e_{s-1}, s-2 u_{s-2} + \cdots + e_{s-1}, 0 u_0;$$

$$u_0 = \text{sgn} \, r_0.$$
Upon substitution,
\[ u_s = \text{sgn } r_s + e_{s, s-1} \text{sgn } r_{s-1} + \cdots + (e_{s, 0} \cdots e_{1, 0}) \text{sgn } r_0. \]
Since \( u_s \neq 0 \), at least one of the constants is not equal to zero and the vectors \((u_s, \text{sgn } r_s, \text{sgn } r_{s-1}, \ldots, \text{sgn } r_0)\) form a linearly dependent set.

Now consider the equation \( r_j = Ax_j - b \). The components of \( r_j \) may be written,
\[ r_{ji} = \text{sgn } r_{ji} \left| r_{ji} \right| \quad (i=1, 2, \ldots, m). \]
If any \( r_{ji} = 0 \), then \( \text{sgn } r_{ji} = 0 \) and we may re-define \( r_{ji} \) as follows: \( r_{ji} = \text{sgn } r_{ji} \left| r_{ji}^* \right| \), where \( r_{ji}^* = \left| r_{ji} \right| \) if \( r_{ji} \neq 0 \), and \( r_{ji}^* = 1 \), if \( r_{ji} = 0 \). Dividing the \( i \)th row of this system by \( \left| r_{ji}^* \right| \), \( i=1, 2, \ldots, m \) respectively, the system becomes:
\[
\begin{bmatrix}
  a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\
  a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \\
\end{bmatrix}
\begin{bmatrix}
  x_{j1} \\
  x_{j1}^* \\
  \vdots \\
  x_{jm}^* \\
\end{bmatrix}
= \begin{bmatrix}
  b_1 \\
  b_2 \\
  \vdots \\
  b_m \\
\end{bmatrix}
\]

The vector \( \text{sgn } r_j \) is a linear combination of these column vectors and the matrix composed of them is of the same rank as the matrix \((Ab)\) since the rows have merely been divided by constants. By the assumption
that a solution exists, rank \((Ab) = \text{rank } A\). Hence, for \((j=0,1,\ldots,R-1)\), the set of vectors \((\text{sgn } r_0, \text{sgn } r_1, \\ \text{sgn } r_2, \ldots, \text{sgn } r_{R-1})\) lie in a subspace of the space spanned by the \(R\) independent column vectors of \(A\). Therefore, the vectors \((u_s, \alpha_1, \alpha_2, \ldots, \alpha_R)\) form a linearly dependent set. This is a contradiction, thus, if a solution exists, \(A^T u_s \neq 0\) for \(s \leq R-1\).

Theorem III.

The method presented gives an exact solution, if it exists, in \(M\) steps, where \(M \leq R\). (\(R\) is the dimension of the smallest space spanned by the column vectors of \(A\)).

Let \(M\) be the first natural number, if it exists, such that \(x-x_0\) is in the subspace spanned by \(v_0, v_1, \ldots, v_{M-1}\). Then,

\[
x-x_0 = k_0 v_0 + \ldots + k_{M-1} v_{M-1} \quad (k_i \text{ not all zero});
\]

\[
(x-x_o) \cdot v_i = k_i v_i \cdot v_i;
\]

\[
k_i = \frac{(x-x_o) \cdot v_i}{v_i \cdot v_i}.
\]

But, from formula 6),

\[
\lambda_i = \frac{-(x-x_o) \cdot v_i}{v_i \cdot v_i}
\]

hence,

\[
k_i = -\lambda_i.
\]
and \[ x = x_0 - \lambda_0 v_0 - \ldots - \lambda_{M-1} v_{M-1}. \]

From formula 4), \[ x_1 = x_0 - \lambda_0 v_0; \]
\[ x_2 = x_1 - \lambda_1 v_1 = x_0 - \lambda_0 v_0 - \lambda_1 v_1; \]
\[ \vdots \]
\[ x_M = x_0 - \lambda_0 v_0 - \ldots - \lambda_{M-1} v_{M-1}. \]

Therefore, \[ x = x_M. \]

**Theorem IV.**

If the direction vectors \( v_0, v_1, \ldots, v_{R-1} \) exist, then \( v_R \) is always equal to zero.

Since the direction vectors \((v_0, v_1, \ldots)\) are linearly independent, the vectors \((v_0, v_1, \ldots, v_{R-1})\) span the space and \( v_R \) must equal zero.

**Theorem V.**

If \( v_i = 0 \) when \( r_i \neq 0 \), then no solution exists.

For \( i \leq R-1 \) this theorem is just the contrapositive of Theorem II. For \( i=R \), Theorem IV states that \( v_R \) always equals zero if \( v_0, v_1, \ldots, v_{R-1} \) exist. If no solution exists, \( r_i \neq 0 \) and the theorem is proved.
ITERATIVE ROUTINE

The formulas developed can now be presented in the following routine:

0) Select an arbitrary vector \( x_0 \).

Compute:

1) \( r_i = Ax_i - b \quad (i=0,1,\ldots,M-1). \)

2) \( \text{grad } \phi_i = A^T \text{sgn } r_i. \)

3) \( e_{o,k} = 0, \quad e_{i,k} = \frac{-\text{grad } \phi_i^T v_k}{v_k^T v_k} \quad (k=0,1,\ldots,i-1). \)

4) \( u_i = \text{sgn } r_i + e_{i,1-i-1}u_{i-1} + \cdots + e_{i,0}u_0. \)

5) \( v_i = A^T u_i. \)

6) \( \lambda_i = \frac{r_0^T u_i}{v_i^T v_i}. \)

7) \( x_{i+1} = x_i - \lambda_i v_i. \)
A summary of the important attributes of this method are listed below:

1) No restrictions of any kind are placed upon the matrix $A$.

2) If a solution exists, it is obtained in $M \leq R$ iterations.

3) If no solution exists, it is indicated in the $i$th iteration ($i \leq R+1$).

4) The given matrix $A$ and the vector $b$ are unaltered during the process, so that a maximum of the original data is used. The advantage of having many zeros in the matrix is preserved.

5) At the end of each iteration an estimate of the solution is given, which is an improvement over the one given in the preceding iteration.

6) At the end of any iteration one can start anew, keeping the estimate last obtained as the initial estimate.
EXAMPLES

In this section various examples will be given to demonstrate the method and exploit its use on a high speed computer. Special attention will be given to round-off errors and a comparison will be made with the Conjugate Gradient Method of Hestenes and Stiefel.

With the exception of the first two, the results of all examples were obtained by use of the ALWAC III-C Computer, using a single length floating point routine.

I. Demonstration of Process.

The process of solution for a $3 \times 3$ system is shown in Table I. The exact solution is obtained in three iterations.

The second example, exhibits the process involved in demonstrating the inconsistency of this set of three equations in two unknowns.

II. Comparison With the Conjugate Gradient Method of Hestenes and Stiefel (2, p. 409-436).

The formulas presented by Hestenes and Stiefel in the non-symmetric case were coded using the same relative methods employed in coding the method presented in this paper. No attempt was made to optimize the coding of either method and no correction of round-off errors was made.

Table II, gives the number of main memory channels
used for storage and the running time of various systems of equations. The running time was calculated from the time of the last input until the last output was completed. This includes the time consumed in typing out each successive estimate of the solution.

Due to round-off errors the exact solution was not obtained in many cases of ill-conditioned matrices. The relative accuracy of the solution of a few of the many systems tried are given by Tables III, IV, V, and VI.

A summary of the results of the comparison are stated below:

1. The Alternating Gradient Method requires more storage space in a computer. That is, 2n+1 channels are needed in excess of the requirement for the Method of Conjugate Gradients.

2. The time involved in the use of the Alternating Gradient Method becomes increasingly greater over the time involved in the Conjugate Gradient Method, as the dimension of the system increases.

3. The Alternating Gradient Method is more accurate, especially in large systems of ill-conditioned equations. However, the Conjugate Gradient Method allows continuance past the nth step for greater accuracy. This cannot be done with the Method of Alternating Gradients. Both methods allow starting over at any
time using the last estimate calculated as the initial vector.

III. Inconsistent Systems.

The inconsistency of a system of equations is exhibited by the direction vector \( v_1 = 0 \) when \( r_1 \neq 0 \). Due to round-off errors this condition is not usually satisfied exactly. However, in all cases tried, the inner product of this vector with itself gave a zero divisor which caused the alarm to sound. Furthermore, in all cases tried where a solution did exist, none of the direction vectors was near enough to zero to sound the alarm. An example of the type of results obtained is given in Table VII.

Using the correction formulas that are developed in the next section, the direction vector did equal zero exactly in every inconsistent system tried.
CORRECTION OF ROUND-OFF ERRORS

The formula for the successive approximations of the solution was given as \( x_{i+1} = x_i - \lambda_i v_i \). This recursive formula presents two possible errors, namely:

1. The direction vector \( v_i \) does not give the correct direction due to round-off errors.
2. The distance, governed by \( \lambda_i \), from \( x_i \) to the new approximation \( x_{i+1} \) may be incorrect due to round-off errors.

To correct the first of these, note that \( v_0 = A^T \text{sgn} r_0 \) must be very accurate, as the only processes involved are the multiplying of each component of the columns of \( A \) by \( \pm 1 \) and adding. Hence, by insuring that every \( v_i \) that is generated after \( v_0 \), forms a mutually orthogonal set with the previously derived direction vectors, their directions are guaranteed. A method for this type of correction is given by Cornelius Lanczos (3, p. 271).

Let \( v_i = \) theoretically correct vector;
\( v_i^\prime = \) actual computed vector;
\( v_i^\ast = \) corrected vector;

then,

\[
v_i^\ast = v_i^\prime - \sum_{k=0}^{i-1} \frac{v_i^\prime \cdot v_k}{v_k^\ast \cdot v_k} v_k^\ast
\]

That this method actually corrects the generated error
is proved as follows.

Let \( e_{i,k} = e_{i,k} + \varepsilon_k \) \((k=0,1,\ldots,i-1)\),
represent the accumulation of the round-off error \( \varepsilon_k \).

From formula 2),

\[
\begin{align*}
\mathbf{v}_i'^* &= \text{grad } \phi_i + e_{i,i-1}^{(i)} \mathbf{v}_i - \cdots + e_{i,0}^{(i)} \mathbf{v}_0; \\
\mathbf{v}_i &= \text{grad } \phi_i + (e_{i,i-1}^{(i)} + \varepsilon_{i-1}) \mathbf{v}_{i-1} + \cdots + (e_{i,0}^{(i)} + \varepsilon_0) \mathbf{v}_0; \\
\mathbf{v}_i &= \text{grad } \phi_i + e_{i,i-1}^{(i)} \mathbf{v}_{i-1} + \cdots + e_{i,0}^{(i)} \mathbf{v}_0 + \sum_{k=0}^{i-1} \varepsilon_k \mathbf{v}_k; \\
\mathbf{v}_i &= \mathbf{v}_i + \sum_{k=0}^{i-1} \varepsilon_k \mathbf{v}_k.
\end{align*}
\]

Substituting this into the Lanczos formula,

\[
\mathbf{v}_i^* = \mathbf{v}_i + \sum_{k=0}^{i-1} \varepsilon_k \mathbf{v}_k - \sum_{k=0}^{i-1} \frac{(\mathbf{v}_i + \sum_{k=0}^{i-1} \varepsilon_k \mathbf{v}_k) \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k} \mathbf{v}_k.
\]

Since the \( \mathbf{v} \)'s form a mutually orthogonal set,

\[
\mathbf{v}_i^* = \mathbf{v}_i + \sum_{k=0}^{i-1} \varepsilon_k \mathbf{v}_k - \sum_{k=0}^{i-1} \frac{\varepsilon_k \mathbf{v}_k \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k} \mathbf{v}_k;
\]

\[
\mathbf{v}_i^* = \mathbf{v}_i.
\]

A formula for the correction of round-off errors of the second type can best be approached in the following geometrical way.
If $\alpha$ is the hyperplane which contains the solution $x$ and is orthogonal to the direction vector $v_i$, then $x_{i+1}^l$ is theoretically the next approximation of the solution $x$. Due to round-off errors assume that $x_{i+1}^l$ is the actual computed approximation, then the error, $\xi$, is the distance between $x_{i+1}^l$ and $x_{i+1}$. This distance can be represented by the inner product of the vector $(x_{i+1}^l-x)$ with the normalized vector $v_i$. That is,

$$\xi = \frac{(x_{i+1}^l-x) \cdot v_i}{|v_i|},$$

but by formula 6), this may be written,

$$\xi = \frac{r_{i+1}^l \cdot u_i}{|v_i|}.$$

The correction formula may now be written as:

$$x_{i+1}^* = x_{i+1}^l - \frac{(r_{i+1}^l \cdot u_i)}{|v_i|} \frac{v_i}{|v_i|}.$$

$$x_{i+1}^* = x_{i+1}^l - \frac{(r_{i+1}^l \cdot u_i)}{v_i \cdot v_i} v_i.$$

Examples of the corrections instituted by these formulas are given in Table VIII.
**TABLE I**

**DEMONSTRATION OF PROCESS**

Example 1.

\[
\begin{align*}
\mathbf{x}_1 + \mathbf{x}_2 - 2\mathbf{x}_3 &= -1 \\
-\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 &= -2 \\
\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 &= 2
\end{align*}
\]

\[
\mathbf{A} = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}
\]

00) \[ \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \]

11) \[ \mathbf{r}_0 = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \]

12) \[ \text{grad } \phi_0 = \begin{bmatrix} 1 & -1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

13) \[ \mathbf{e}_{0,0} = 0 \]

14) \[ \mathbf{u}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \]

15) \[ \mathbf{v}_0 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} \]

16) \[ \lambda_0 = \frac{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}} = \frac{1}{2} \]

17) \[ \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{2} \\ -1 \end{bmatrix} \]

21) \[ \mathbf{r}_1 = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix} \]
22) \[ \text{grad } \phi_1 = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix} \]

23) \[ e_{1,0} = \frac{-(-3 \quad -1 \quad 2)}{(1 \quad 3 \quad 0)} = \frac{3}{5} \]

24) \[ u_1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} + \frac{3}{5} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 8/5 \\ -2/5 \end{bmatrix} \]

25) \[ v_1 = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2/5 \\ 8/5 \\ -2/5 \end{bmatrix} = \begin{bmatrix} -12/5 \\ 4/5 \\ 10/5 \end{bmatrix} \]

26) \[ \lambda_1 = \frac{(1 \quad 3 \quad 1)}{(-12/5 \quad 4/5 \quad 10/5)} = \frac{5}{13} \]

27) \[ x_2 = \begin{bmatrix} 1/2 \\ -1/2 \\ 2/2 \end{bmatrix} - \frac{5}{13} \begin{bmatrix} -12/5 \\ 4/5 \\ 10/5 \end{bmatrix} = \begin{bmatrix} 37/26 \\ -21/26 \\ 6/26 \end{bmatrix} \]

31) \[ r_2 = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 37/26 \\ -21/26 \\ 6/26 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 15/13 \\ 0 \\ -15/13 \end{bmatrix} \]

32) \[ \text{grad } \phi_2 = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix} \]

33) \[ e_{2,1} = \frac{-(-3 \quad 0 \quad 0)}{(-12/5 \quad 4/5 \quad 10/5)} = \frac{15}{26} \]
\[
e_{2,0} = \frac{-(0 \quad 0 \quad 3)}{(1 \quad 3 \quad 0)} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 10 \end{bmatrix} = 0
\]

34) \[
u_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{15}{26} \begin{bmatrix} -2/5 \\ 8/5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 10/13 \\ 12/13 \end{bmatrix}
\]

35) \[
u_2 = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ -2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10/13 \\ 12/13 \\ -24/13 \end{bmatrix} = \begin{bmatrix} -18/13 \\ -6/13 \\ 5/12 \end{bmatrix}
\]

36) \[
\lambda_2 = \frac{(15/13 \quad 0 \quad -15/13)}{(-18/13 \quad 6/13 \quad -24/13)} = \begin{bmatrix} 10/13 \\ 12/13 \\ -16/13 \end{bmatrix} = \frac{5}{12}
\]

37) \[
x_3 = \begin{bmatrix} 37/26 \\ -21/26 \\ 6/26 \end{bmatrix} - \frac{5}{12} \begin{bmatrix} -18/13 \\ 6/13 \\ -24/13 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}
\]

This is the exact solution.

Example 2.
\[
x_1 - x_2 = -4
\]
\[
2x_1 + x_2 = 1
\]
\[
x_1 - x_2 = 2
\]

00) \[
x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\]

11) \[
r_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ -3 \end{bmatrix}
\]

12) \[
\text{grad } \phi = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}
\]

13) \[
e_{0,0} = 0
\]
14) \( u_o = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \)

15) \( v_o = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \)

16) \( \lambda_o = \frac{1}{2} \begin{pmatrix} 4 & 2 & -2 \\ 2 & 1 \end{pmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \frac{8}{5} \)

17) \( x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{8}{5} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -11/5 \\ -3/5 \end{bmatrix} \)

21) \( r_1 = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -11/5 \\ -3/5 \end{bmatrix} - \begin{bmatrix} 4 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 12/5 \\ -6 \end{bmatrix} - \begin{bmatrix} 18/5 \\ -12/5 \end{bmatrix} \)

22) \( \text{grad } \phi_1 = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix} \)

23) \( e_{1,0} = \frac{-(-2 & -1)}{2 & 1} \frac{2}{1} = 1 \)

24) \( u_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \)

25) \( v_1 = \begin{bmatrix} -2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \)

This indicates that no solution exists.
**TABLE II**  
**COMPARISON OF METHODS**

Storage Space in Main Memory of Computer:

<table>
<thead>
<tr>
<th></th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>control routine</td>
<td>4 channels</td>
<td>4 channels</td>
</tr>
<tr>
<td>sub-routines</td>
<td>5 channels</td>
<td>5 channels</td>
</tr>
<tr>
<td>(floating pt.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>general storage</td>
<td>3n+6 channels</td>
<td>n+5 channels</td>
</tr>
<tr>
<td>totals</td>
<td>3n+15 channels</td>
<td>n+14 channels</td>
</tr>
</tbody>
</table>

Running Time:

<table>
<thead>
<tr>
<th>Dimensions of System</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 x 2</td>
<td>31 sec.</td>
<td>30 sec.</td>
</tr>
<tr>
<td>3 x 3</td>
<td>1 min. 14 sec.</td>
<td>1 min. 10 sec.</td>
</tr>
<tr>
<td>4 x 4</td>
<td>2 min. 22 sec.</td>
<td>2 min. 12 sec.</td>
</tr>
<tr>
<td>6 x 6</td>
<td>6 min. 26 sec.</td>
<td>5 min. 26 sec.</td>
</tr>
<tr>
<td>8 x 8</td>
<td>13 min. 40 sec.</td>
<td>11 min. 6 sec.</td>
</tr>
<tr>
<td>16 x 16</td>
<td>1 hr. 27 min. 52 sec.</td>
<td>1 hr. 6 min. 53 sec.</td>
</tr>
</tbody>
</table>
TABLE III
COMPARISON OF METHODS

Two Equations in Two Unknowns:

Example 1. (well-conditioned)

Ratio of largest to smallest eigenvalues is 2.

\[
\begin{bmatrix}
1 & 2 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
=
\begin{bmatrix}
3 \\
1
\end{bmatrix}
\]

Initial vector = (5 8).

<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1 = 1.0000000)</td>
<td>1.0000000</td>
<td>0.9999994</td>
</tr>
<tr>
<td>(x_2 = 1.0000000)</td>
<td>1.0000000</td>
<td>0.9999991</td>
</tr>
</tbody>
</table>

Example 2. (ill-conditioned)

Ratio of largest to smallest eigenvalues is 20,000.

\[
\begin{bmatrix}
200 & 256 \\
155.5625 & 199.08
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
=
\begin{bmatrix}
456 \\
354.6425
\end{bmatrix}
\]

Initial vector = (5 8).

<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1 = 1.0000000)</td>
<td>1.0018341</td>
<td>0.8772738</td>
</tr>
<tr>
<td>(x_2 = 1.0000000)</td>
<td>0.9985004</td>
<td>1.7127669</td>
</tr>
</tbody>
</table>
### TABLE IV

**COMPARISON OF METHODS**

Four Equations in Four Unknowns:

**Example 1.** (well-conditioned)

Ratio of largest to smallest eigenvalues is 9.47.

\[
\begin{bmatrix}
-2 & 1 & 0 & 0 \\
1 & -2 & 1 & 0 \\
0 & 1 & -2 & 1 \\
0 & 0 & 1 & -2
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
= 
\begin{bmatrix}
-4 \\
5 \\
1 \\
-6
\end{bmatrix}
\]

Initial vector = \((1 \ 1 \ 1 \ 1)\).

<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1) = 1.0000000</td>
<td>0.99999999</td>
<td>0.99999978</td>
</tr>
<tr>
<td>(x_2) = -2.0000000</td>
<td>-1.99999999</td>
<td>-1.99999988</td>
</tr>
<tr>
<td>(x_3) = 0.0000000</td>
<td>0.00000000</td>
<td>0.00000012</td>
</tr>
<tr>
<td>(x_4) = 3.0000000</td>
<td>2.99999998</td>
<td>2.99999977</td>
</tr>
</tbody>
</table>

**Example 2.** (ill-conditioned)

Ratio of largest to smallest eigenvalues is 2984.

\[
\begin{bmatrix}
10 & 7 & 8 & 7 \\
7 & 5 & 6 & 5 \\
8 & 6 & 10 & 9 \\
7 & 5 & 9 & 10
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix}
= 
\begin{bmatrix}
32 \\
23 \\
33 \\
31
\end{bmatrix}
\]

Initial vector = \((0 \ 0 \ 0 \ 0)\).

<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1) = 1.0000000</td>
<td>0.9999721</td>
<td>1.0953395</td>
</tr>
<tr>
<td>(x_2) = 1.0000000</td>
<td>1.0000532</td>
<td>0.8401629</td>
</tr>
<tr>
<td>(x_3) = 1.0000000</td>
<td>0.999796</td>
<td>1.0385410</td>
</tr>
<tr>
<td>(x_4) = 1.0000000</td>
<td>1.000264</td>
<td>0.9757126</td>
</tr>
</tbody>
</table>
TABLE V
COMPARISON OF METHODS

Eight Equations in Eight Unknowns:

Example 1. (well-conditioned)

Ratio of largest to smallest eigenvalues is $30.4859$.

$$
\begin{pmatrix}
10 & 7 & 8 & 7 & 6 & 8 & 5 & 4 \\
7 & 5 & 6 & 5 & 7 & 7 & 4 & 9 \\
8 & 6 & 10 & 9 & 8 & 5 & 8 & 2 \\
7 & 5 & 9 & 10 & 5 & 9 & 6 & 8 \\
6 & 7 & 8 & 5 & 5 & 4 & 7 & 6 \\
8 & 7 & 5 & 9 & 4 & 10 & 9 & 5 \\
5 & 4 & 8 & 6 & 7 & 9 & 10 & 1 \\
4 & 9 & 2 & 8 & 6 & 5 & 1 & 5
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6 \\
x_7 \\
x_8
\end{pmatrix}
= \begin{pmatrix}
55 \\
50 \\
56 \\
59 \\
48 \\
57 \\
50 \\
40
\end{pmatrix}

\[ x_0 = (5 \ 2 \ -1 \ 4 \ -3 \ 1.5 \ 2 \ -4). \]

<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$ = 1.0000000</td>
<td>0.9999989</td>
<td>1.0033344</td>
</tr>
<tr>
<td>$x_2$ = 1.0000000</td>
<td>1.0000005</td>
<td>1.0004616</td>
</tr>
<tr>
<td>$x_3$ = 1.0000000</td>
<td>0.9999998</td>
<td>0.9980522</td>
</tr>
<tr>
<td>$x_4$ = 1.0000000</td>
<td>0.9999988</td>
<td>0.9957286</td>
</tr>
<tr>
<td>$x_5$ = 1.0000000</td>
<td>1.0000008</td>
<td>1.0026373</td>
</tr>
<tr>
<td>$x_6$ = 1.0000000</td>
<td>1.0000002</td>
<td>1.0021600</td>
</tr>
<tr>
<td>$x_7$ = 1.0000000</td>
<td>0.9999993</td>
<td>0.9993067</td>
</tr>
<tr>
<td>$x_8$ = 1.0000000</td>
<td>0.9999996</td>
<td>0.9988578</td>
</tr>
</tbody>
</table>
Example 2. (ill-conditioned)

Ratio of largest to smallest eigenvalues is 667.69.

\[
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6 \\
x_7 \\
x_8
\end{pmatrix}
=
\begin{pmatrix}
8.0 \\
7.9 \\
7.8 \\
7.7 \\
7.6 \\
7.5 \\
7.4 \\
7.3
\end{pmatrix}
\]

\[x_0 = (5 \ 2 \ -1 \ 4 \ -3 \ 1.5 \ 2 \ -4)\].

<table>
<thead>
<tr>
<th>Exact Solution (x_i)</th>
<th>Alternating Gradient (x_i)</th>
<th>Conjugate Gradient (x_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(x_1 = 1.0000000)</td>
<td>1.0000303</td>
<td>0.9862018</td>
</tr>
<tr>
<td>(x_2 = 1.0000000)</td>
<td>0.9999988</td>
<td>1.0203175</td>
</tr>
<tr>
<td>(x_3 = 1.0000000)</td>
<td>1.0000105</td>
<td>0.9771331</td>
</tr>
<tr>
<td>(x_4 = 1.0000000)</td>
<td>0.9999942</td>
<td>1.0148934</td>
</tr>
<tr>
<td>(x_5 = 1.0000000)</td>
<td>1.0000108</td>
<td>0.9934853</td>
</tr>
<tr>
<td>(x_6 = 1.0000000)</td>
<td>0.9999983</td>
<td>1.0067546</td>
</tr>
<tr>
<td>(x_7 = 1.0000000)</td>
<td>0.9999886</td>
<td>0.9880737</td>
</tr>
<tr>
<td>(x_8 = 1.0000000)</td>
<td>1.0000003</td>
<td>1.0132430</td>
</tr>
</tbody>
</table>
TABLE VI
COMPARISON OF METHODS

Sixteen Equations in Sixteen Unknowns:

Example 1.

\[
\begin{bmatrix}
1 & 2 & -7 & 4 & -1 & 0 & -2 & 1 & 4 & 3 & -2 & 0 & -1 & 2 & 1 & 0 \\
2 & 5 & 4 & 3 & -1 & 4 & -2 & 2 & 1 & 2 & -1 & 0 & -4 & 1 & -1 & 0 \\
-1 & -1 & 3 & 1 & 1 & 3 & -3 & 1 & 1 & 1 & 0 & 2 & 1 & 2 & 0 & -1 \\
1 & 7 & 2 & 2 & 1 & 1 & -1 & 0 & -2 & 4 & 0 & 1 & -2 & 4 & -1 & -2 \\
3 & 3 & 1 & 3 & 2 & -1 & 1 & 0 & 3 & 2 & 0 & -3 & -1 & -2 & -1 & -1 \\
1 & 1 & 7 & -1 & 2 & 0 & 1 & 3 & -4 & 1 & 1 & 0 & 1 & 1 & -1 & -2 \\
-2 & 2 & 1 & 4 & -1 & 7 & 1 & 2 & 2 & 1 & -1 & 0 & 1 & 0 & -1 & 1 \\
4 & 8 & 1 & 1 & -1 & -2 & -2 & 1 & -1 & 1 & 1 & 0 & -1 & 3 & -2 & 0 \\
1 & -1 & -1 & 1 & 0 & 1 & 1 & -1 & 4 & 1 & 1 & 3 & 1 & -1 & -3 & 1 & 0 \\
2 & 4 & -1 & -1 & 4 & 4 & 2 & 4 & 5 & 0 & 4 & 0 & -1 & 3 & -1 & -1 \\
-2 & 1 & -1 & -2 & 1 & 2 & 3 & 1 & -6 & 0 & 1 & 0 & 2 & 2 & -1 & 1 \\
3 & 1 & 1 & 2 & 0 & 3 & 1 & 0 & -5 & 0 & -2 & -1 & 1 & -1 & 1 & 0 \\
1 & 1 & 1 & 2 & 1 & -1 & 1 & 0 & -1 & -1 & 3 & 4 & 0 & -2 & 0 & 1 \\
4 & 3 & -2 & 3 & 1 & 4 & -1 & -3 & 1 & 0 & -2 & -1 & 0 & 1 & 0 & -2 \\
2 & -2 & 1 & 1 & -1 & 0 & 0 & 2 & 0 & -1 & 1 & 1 & 3 & 2 & -1 & -1 \\
1 & 1 & -1 & 4 & 0 & -1 & 0 & 1 & 1 & 1 & 1 & 3 & -1 & 0 & -1 & -2
\end{bmatrix}
\]

\[b = (21 \ 10 \ 6 \ 14 \ 0 \ -2 \ 14 \ 8 \ 5 \ 16 \ 10 \ 6 \ 11 \ 6 \ -7 \ 13).\]

\[x_0 = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1).\]
<table>
<thead>
<tr>
<th>Exact Solution</th>
<th>Alternating Gradient</th>
<th>Conjugate Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$ = -1.0000000</td>
<td>-0.9999985</td>
<td>0.8892401</td>
</tr>
<tr>
<td>$x_2$ = 2.0000000</td>
<td>2.0000001</td>
<td>1.8488161</td>
</tr>
<tr>
<td>$x_3$ = -2.0000000</td>
<td>-2.0000007</td>
<td>1.8247232</td>
</tr>
<tr>
<td>$x_4$ = 1.0000000</td>
<td>0.9999993</td>
<td>1.011682</td>
</tr>
<tr>
<td>$x_5$ = 2.0000000</td>
<td>1.9999999</td>
<td>2.0104537</td>
</tr>
<tr>
<td>$x_6$ = 1.0000000</td>
<td>1.0000009</td>
<td>0.7997739</td>
</tr>
<tr>
<td>$x_7$ = -1.0000000</td>
<td>-1.0000017</td>
<td>-0.8470299</td>
</tr>
<tr>
<td>$x_8$ = 2.0000000</td>
<td>2.0000014</td>
<td>1.6983474</td>
</tr>
<tr>
<td>$x_9$ = -1.0000000</td>
<td>-0.9999993</td>
<td>-1.0599308</td>
</tr>
<tr>
<td>$x_{10}$ = 1.0000000</td>
<td>0.9999978</td>
<td>1.3231497</td>
</tr>
<tr>
<td>$x_{11}$ = 0.0000000</td>
<td>0.0000050</td>
<td>-0.6304432</td>
</tr>
<tr>
<td>$x_{12}$ = 2.0000000</td>
<td>1.9999982</td>
<td>2.2931508</td>
</tr>
<tr>
<td>$x_{13}$ = 0.0000000</td>
<td>-0.0000036</td>
<td>0.1149432</td>
</tr>
<tr>
<td>$x_{14}$ =-1.0000000</td>
<td>-1.0000004</td>
<td>-0.9168063</td>
</tr>
<tr>
<td>$x_{15}$ = 1.0000000</td>
<td>1.0000024</td>
<td>0.5898514</td>
</tr>
<tr>
<td>$x_{16}$ = 0.0000000</td>
<td>-0.0000009</td>
<td>0.4474895</td>
</tr>
</tbody>
</table>
### TABLE VII

#### INCONSISTENT SYSTEMS

Example 1.

\[
\begin{pmatrix}
1 & 4 & -3 & 9 \\
-2 & 7 & 1 & 6 \\
-1 & 3 & -2 & 1 \\
2 & 8 & -6 & 18
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{pmatrix}
= 
\begin{pmatrix}
11 \\
12 \\
1 \\
0
\end{pmatrix}
\]

The ratio of the largest to smallest eigenvalue excluding the zero eigenvalue is 7.9269.

\[x_0 = (0, 0, 0, 0)\].

\[\lambda_0 = 0.0508474\]

\[\lambda_1 = 0.6399669\]

\[\lambda_2 = 0.3295198\]

\[r_3 = (-5.5000054, 10.9999942, -5.5000019, 10.9999890)\].

\[v_3 = (0.0000006, -0.0000005, 0.0000012, 0.0000013)\].

Example 2.

\[
\begin{pmatrix}
1 & .5 & 333333333 & .25 \\
.5 & 333333333 & .25 & .2 \\
333333333 & .25 & .2 & 166666667 \\
571428571 & 285714286 & .19047619 & .142857143
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
0 \\
0 \\
0
\end{pmatrix}
\]

The ratio of the largest to smallest eigenvalues excluding the zero eigenvalue is 105.2273959.

\[x_0 = (1, 1, 1, 1)\].

\[\lambda_0 = 0.49084837\]

\[\lambda_1 = 37.0555620\]

\[r_2 = (-0.0000000, -0.1815374, -0.3898911, 0.5714285)\].

\[v_2 = (-0.0000000, -0.0000000, -0.0000000, -0.0000000)\].
Table VIII
CORRECTION OF ROUND-OFF ERRORS

Example 1.
Two Equations in Two Unknowns:

The system given in Table III, Example 2, is used as the first example. This system was ill-conditioned.

\[ x_0 = (5 \ 8). \]

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact Solution</td>
<td>Method of Alternating Gradients Without Correction</td>
</tr>
<tr>
<td>1.00000000</td>
<td>1.0018341</td>
</tr>
<tr>
<td>0.9985004</td>
<td>0.9985353</td>
</tr>
</tbody>
</table>

Example 2.
Four Equations in Four Unknowns:

This system is given in Table IV, Example 2.

The system is ill-conditioned.

\[ x_0 = (0 \ 0 \ 0 \ 0). \]

<table>
<thead>
<tr>
<th>( x_1 )</th>
<th>( x_2 )</th>
<th>( x_3 )</th>
<th>( x_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact Solution</td>
<td>Method of Alternating Gradients Without Correction</td>
<td>Method of Alternating Gradients With Correction</td>
<td></td>
</tr>
<tr>
<td>1.00000000</td>
<td>1.0000072</td>
<td>1.00000072</td>
<td></td>
</tr>
<tr>
<td>0.9999721</td>
<td>0.9999938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0000532</td>
<td>1.0000041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9999976</td>
<td>0.9999975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0000264</td>
<td>0.9999975</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example 3.

Eight Equations in Eight Unknowns:

This system is given in Table V, Example 2.

It was not too badly conditioned, the ratio of its largest to smallest eigenvalues being 667.69.

\[ x_0 = (5 \ 2 \ -1 \ 4 \ -3 \ 1.5 \ 2 \ -4). \]

<table>
<thead>
<tr>
<th>( x_i )</th>
<th>Exact Solution</th>
<th>Method of Alternating Gradients Without Correction</th>
<th>Method of Alternating Gradients With Correction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_1 = 1.0000000 )</td>
<td>1.0000000000</td>
<td>1.00000303</td>
<td>1.00000036</td>
</tr>
<tr>
<td>( x_2 = 1.0000000 )</td>
<td>0.99999988</td>
<td>0.99999986</td>
<td>0.99999996</td>
</tr>
<tr>
<td>( x_3 = 1.0000000 )</td>
<td>1.00000105</td>
<td>0.9999992</td>
<td>0.99999992</td>
</tr>
<tr>
<td>( x_4 = 1.0000000 )</td>
<td>0.99999942</td>
<td>0.9999986</td>
<td>0.99999986</td>
</tr>
<tr>
<td>( x_5 = 1.0000000 )</td>
<td>1.00000108</td>
<td>0.9999993</td>
<td>0.99999993</td>
</tr>
<tr>
<td>( x_6 = 1.0000000 )</td>
<td>0.99999983</td>
<td>0.9999995</td>
<td>0.99999995</td>
</tr>
<tr>
<td>( x_7 = 1.0000000 )</td>
<td>0.9999986</td>
<td>0.9999997</td>
<td>0.99999997</td>
</tr>
<tr>
<td>( x_8 = 1.0000000 )</td>
<td>1.00000003</td>
<td>0.9999998</td>
<td>0.99999998</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


