The Conjugate Gradient Method

A Tutorial Note

by

Erik Thompson

OVERVIEW

In the following, we will point out and explain that:

1. The solution of a set of n simultaneous, symmetric, linear equations is equivalent to the minimization of a quadratic function in the same n-space.

2. That this minimization can be performed using n directions and if these directions are A-conjugate, that these searches can be made independent of each other.

3. Finally, we will show how these A-conjugate directions can be easily obtained during the solution process.

NOTATION

$$
\begin{aligned}
\{\mathrm{A}\} &= \text{A column matrix.} \\
\lfloor \mathrm{A} \rfloor &= \text{A row matrix.} \\
[\mathrm{A}] &= \text{A square matrix.} \\
[\,\mathrm{I}\,] &= \text{The Identity matrix} \\
[A]^T &= \text{Transpose of } [A]. \\
\{A\}^T &= \lfloor \mathrm{A} \rfloor \\
\lfloor \mathrm{A} \rfloor^T &= \{\mathrm{A}\}
\end{aligned}
$$

Equations that are used directly in the computer code will be placed in boxes such as this.

DEVELOPEMENT

The Minimizational Function

We seek the solution to the set of n linear equations:

$$[\mathrm{A}]\{\mathrm{x}\} = \{\mathrm{f}\}$$

where [A] is symmetric and positive-definite. This is equivalent to finding the value of {x} that minimizes

$$\Phi = \frac{1}{2}\lfloor \mathrm{x} \rfloor [\mathrm{A}]\{\mathrm{x}\} - \lfloor \mathrm{x} \rfloor \{\mathrm{f}\}$$

The gradient of Φ is

$$\{\nabla\Phi\} = \frac{1}{2}[\,\mathrm{I}\,][\mathrm{A}]\{\mathrm{x}\} + \frac{1}{2}\lfloor \mathrm{x} \rfloor [\mathrm{A}][\,\mathrm{I}\,] - [\,\mathrm{I}\,]\{\mathrm{f}\}$$

where

$$\{\nabla\Phi\} = \left\{ \begin{array}{c} \dfrac{\partial\Phi}{\partial x_1} \\ \dfrac{\partial\Phi}{\partial x_2} \\ \vdots \\ \dfrac{\partial\Phi}{\partial x_n} \end{array} \right\}$$

Because [A] is symmetric,

$$\{\nabla\Phi\} = [\,\mathrm{I}\,][\mathrm{A}]\{\mathrm{x}\} - [\,\mathrm{I}\,]\{\mathrm{f}\}$$

or, simply

$$\{\nabla\Phi\} = [\mathrm{A}]\{\mathrm{x}\} - \{\mathrm{f}\}$$

Hence the gradient of Φ is zero at the point {x} that satisfies our set of simultaneous equations. At some other point, however, it is not zero and gives us the direction of the steepest ascent of the surface Φ. When not equal to zero, it is also referred to as the (negative) residual of our set of equations. Hence,

$$\{\mathrm{g}\} = -\{\mathrm{r}\} = \left\{\frac{\partial\Phi}{\partial x}\right\} = \{\nabla\Phi\}$$

where

$$\boxed{\{\mathrm{r}\} = \{\mathrm{f}\} - [\mathrm{A}]\{\mathrm{x}\}}$$

Linear Independent Vectors

In what follows we will make use of the concept that any vector in our n-space can be considered a linear function of n linear independent vectors in that space. Let such a set of n-linearly independent vectors be

$$\{p\}_1, \quad \{p\}_2, \quad \{p\}_3, \quad \{p\}_4, \quad \cdots \quad \{p\}_n$$

Then any vector $\{x\}$ can be represented as

$$\{x\} = \{\mathrm{x}\}_0 + s_1\{p\}_1 + s_2\{p\}_2 + s_3\{p\}_3 + s_4\{p\}_4 \cdots + s_n\{p\}_n$$

Or, more compactly,

$$\{x\} = \{\mathrm{x}\}_0 + [\mathrm{P}]\{\mathrm{S}\}$$

Here, $\{\mathrm{x}\}_0$ is an initial reference vector (a constant term in the expansion), [P] is the matrix of the independent vectors $\{\mathrm{p}\}_i$, and {S} is a column vector containing the values of s_i. When this representation of {x} is substituted into the expression for Φ, Φ becomes a function of {s}. The solution to the set of simultaneous equations is then found by minimizing Φ with respect to the s values.

Conjugate Vectors and the Diagonalization of [A]

We now introduce the basic idea behind the conjugate gradient method. A set of vectors p_i are said to be A-conjugate if

$$\lfloor \mathrm{p} \rfloor_i [\mathrm{A}] \{\mathrm{p}\}_j = 0 \qquad \text{for} \quad i \neq j$$

Hence, if the set of linear independent vectors in [P] are A-conjugate then

$$[\mathrm{P}]^T [\mathrm{A}][\mathrm{P}] = [\mathrm{D}]$$

where [D] is a diagonal matrix. Its i^{th} diagonal is

$$d_i = \lfloor \mathrm{p} \rfloor_i [\mathrm{A}] \{\mathrm{p}\}_i$$

The gradient of Φ can now be written as

$$\{\nabla\Phi\} = [\mathrm{A}]\{\mathrm{x}\} - \{\mathrm{f}\} = [\mathrm{A}]\{\mathrm{x}\}_0 + [\mathrm{A}][\mathrm{P}]\{\mathrm{S}\} - \{\mathrm{f}\}$$

or

$$\{\nabla\Phi\} = [\mathrm{A}][\mathrm{P}]\{\mathrm{S}\} + \{\mathrm{r}\}_0$$

where

$$\{\mathrm{r}\}_0 = \{\mathrm{f}\} - [\mathrm{A}]\{\mathrm{x}\}_0$$

Multiplication by the transpose of [P] and setting the gradient equal to zero gives us

$$[\mathrm{P}]^T [\mathrm{A}][\mathrm{P}]\{\mathrm{S}\} + [\mathrm{P}]^T \{\mathrm{r}\}_0 = 0$$

If the vectors in [P] are A-conjugate then

$$[\mathrm{D}]\{\mathrm{S}\} + [\mathrm{P}]^T\{\mathrm{r}\}_0 = 0$$

We have decoupled our set of equations for the values of s in $\{\mathrm{S}\}$. These values minimize Φ and thus give us the solution to our original equation. They are given by

$$s_i = \frac{\lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0}{\lfloor \mathrm{p} \rfloor_i [\mathrm{A}] \{\mathrm{p}\}_i}$$

Each s can now be solved for independently and used to write the solution to our set of simultaneous equations as

$$\{\mathrm{x}\} = \{\mathrm{x}\}_0 + [\mathrm{P}]\{\mathrm{S}\}$$

The Concept of Line Searches

An essential feature of the conjugate gradient method is the intepretation that each s_i in $\{\mathrm{S}\}$ found by globally minimizing Φ, locally minimizes Φ along $\{p\}_i$. Thus the solution

$$\{x\} = \{x\}_0 + s_1\{p\}_1 + s_2\{p\}_2 + s_3\{p\}_3 + s_4\{p\}_4 \cdots + s_n\{p\}_n$$

can be thought of as a sequence of steps such that

$$\{x\}_1 = \{x\}_0 + s_1\{p\}_1$$

is the location of the minimum value of Φ along a line from $\{\mathrm{x}\}_0$ in the direction $\{p\}_1$. Next

$$\{x\}_2 = \{x\}_0 + s_1\{p\}_1 + s_2\{p\}_2$$

minimizes Φ by a search from $\{\mathrm{x}\}_0$ in two directions, $\{p\}_1$ and then $\{p\}_2$. However, the above can be written as

$$\{x\}_2 = \{x\}_1 + s_2\{p\}_2$$

This step then minimizes Φ along a line from $\{x\}_1$ in the direction $\{p\}_2$. The next step can thus be written

$$\{x\}_3 = \{x\}_2 + s_3\{p\}_3$$

and so forth.

The structure of conjugate gradient algorithms is thus: Starting at some point $\{\mathrm{x}\}_0$, a search for the minimum value of Φ in the direction of $\{p\}_1$ is begun. This will be at the point s_1 along $\{p\}_1$, given by

$$s_1 = \frac{\lfloor \mathrm{p} \rfloor_1 \{\mathrm{r}\}_0}{\lfloor \mathrm{p} \rfloor_1 [\mathrm{A}] \{\mathrm{p}\}_1}$$

The next step begins a new search for the minimum value of Φ staring at $\{x\}_1$ in the direction of $\{p\}_2$. The value of s_2 that locates this point will now be given by

$$s_2 = \frac{\lfloor \mathrm{p} \rfloor_2 \{\mathrm{r}\}_1}{\lfloor \mathrm{p} \rfloor_1 [\mathrm{A}] \{\mathrm{p}\}_1}$$

where $\{\mathrm{r}\}_1$ is the residual at the new reference point $\{\mathrm{x}\}_1$. The procedure continues in this manner until all conjugate vectors have been used as search directions. At that time the true minimum of Φ will have been obtained. For this reason, the conjugate vectors are referred to as search directions. In general then, from some point $\{\mathrm{x}\}_i$, the next point is found as:

$$\{\mathrm{x}\}_{i+1} = \{\mathrm{x}\}_i + s_i\{\mathrm{p}\}_i$$

where

$$s_i = \frac{\lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_i}{\lfloor \mathrm{p} \rfloor_i [\mathrm{A}] \{\mathrm{p}\}_i}$$

We now show how these conjugate search directions are obtained.

Selection of Search Directions

In general, there are an infinite number of sets of conjugate vectors for a symmetric, positive definite matrix. For example

$$\begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 6 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 6 \end{bmatrix}$$

Note also that the Eigenvectors of [A] are conjugate so we know that there must be at least one set.

Conjugate gradient algorithms determine each direction vector as needed and then that vector is discarded as soon as the next direction is determined. Hence, we direct our attention as to how each new direction is chosen based on the current point and the previous direction. We do this in a manner that will leave one question temporarily unanswered, but we will return to it later.

Consider that we have reached the point $\{x\}_i$ which has just minimized Φ along $\{p\}_i$ and we are now ready to determine the next search direction $\{p\}_{i+1}$. Because our primary objective is to locate the global minimum of Φ, a logical choice for the next direction would be the direction of the steepest decent. However, this direction is probably not conjugate to our previous search directions; hence, such a move would undo what we have already accomplished. We therefore wish to move in the general direction of the steepest decent, but change it sufficiently so as to make the new direction conjugate to our previous direction. This will decouple the current search from our previous calculations.

We now show how to make the next search direction have a component in the direction of the negative gradient and be conjugate to the previous search direction. Later we will prove that this procedure is sufficient to insure that this new search direction is conjugate to all previous search directions.

We choose the new direction as made up of two components. One in the direction of steepest decent (the residual) and the other in the direction of the previous search. We write this as

$$\{p\}_i = \{r\}_i - \lambda\{p\}_{i-1}$$

That is, we plan to make a correction in the direction of steepest decent so that the new direction will be conjugate to the direction just traversed. Hence we seek a value of λ that will accomplish this and write

$$\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{\mathrm{p}\}_i = 0$$

Substitution of the general equation for our new search direction gives us

$$\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\left\{\{r\}_i - \lambda\{p\}_{i-1}\right\} = 0$$

This give us

$$\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{r\}_i = \lambda\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{p\}_{i-1}$$

which we solve for λ to obtain

$$\lambda = \frac{\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{r\}_i}{\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{p\}_{i-1}}$$

The new search direction is therefore

$$\{P\}_i = \{r\} - \lambda\{p\}_{i-1}$$

where

$$\lambda = \frac{\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{r\}_i}{\lfloor \mathrm{p} \rfloor_{i-1}[\mathrm{A}]\{p\}_{i-1}}$$

We must now address the question as to whether the above procedure produces a set of mutually conjugate directions. In order to prove this we must first prove that when such a set of directions is used to arrive at a minimum, then they are orthogonal to the gradients at each point along the search. And second, we must prove that the the gradients (residuals) are mutually orthogonal

<u>The Orthogonality of Gradients and Conjugate Directions</u>

We now set out to prove that at any point, $\{\mathrm{x}\}_j$, the gradient, $\{\mathrm{g}\}_j$, is orthogonal to all the previous search directions. Therefore, we seek to prove

$$\lfloor \mathrm{p} \rfloor_i \{\mathrm{g}\}_j = 0 \qquad\qquad i < j$$

We begin by first noting that

$$\{\mathrm{g}\}_j = \{\mathrm{f}\}_j - [\mathrm{A}]\{\mathrm{x}\}_j$$

and that

$$\{\mathrm{x}\}_j = \{\mathrm{x}\}_0 + [\mathrm{P(j)}]\{\mathrm{s(j)}\}$$

where (j) indicates that all elements in the array up to the j^{th} element are considered. Hence, the above simply states that

$$\{\mathrm{x}\}_j = \{\mathrm{x}\}_0 + \{\mathrm{p}\}_1 s_1 + \{\mathrm{p}\}_2 s_2 \cdots + \{\mathrm{p}\}_j s_j$$

We now have

$$\lfloor \mathrm{p} \rfloor_i \{\mathrm{g}\}_j = \lfloor \mathrm{p} \rfloor_i \{\mathrm{f}\} - \lfloor \mathrm{p} \rfloor_i [\mathrm{A}][\mathrm{x}]_0 - \lfloor \mathrm{p} \rfloor_i [\mathrm{A}][\mathrm{P(j)}]\{\mathrm{S(j)}\}$$

But, because the $\{\mathrm{p}\}$'s are conjugate, this simplifies to

$$\lfloor \mathrm{p} \rfloor_i \{\mathrm{g}\}_j = \lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0 - \lfloor \mathrm{p} \rfloor_i [\mathrm{A}]\{\mathrm{p}\}_i s_i$$

Now we use our expression for s_i to obtain

$$\lfloor \mathrm{p} \rfloor_i \{\mathrm{g}\}_j = \lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0 - \lfloor \mathrm{p} \rfloor_i [\mathrm{A}]\{\mathrm{p}\}_i \frac{\lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0}{\lfloor \mathrm{p} \rfloor_i [\mathrm{A}]\{\mathrm{p}\}_i}$$

$$\lfloor \mathrm{p} \rfloor_i \{\mathrm{g}\}_j = \lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0 - \lfloor \mathrm{p} \rfloor_i \{\mathrm{r}\}_0 \equiv 0$$

This states that the gradient $\{\mathrm{g}\}_j$ at point $\{\mathrm{x}\}_j$ is orthogonal to every search direction $\{\mathrm{p}\}_i$ for $i < j$. Therefore, it is orthogonal to every search direction used to arrive at point $\{\mathrm{x}\}_j$ from $\{\mathrm{x}\}_0$.. It is based on the assumptions that the search directions $\{\mathrm{p}\}_i$ are A-conjugate and that for each direction, Φ has been minimized.

The Orthogonality of Gradients

In this section we set out to prove that the gradients at points located during the conjugate gradient method are mutually orthogonal. That is, we seek to prove

$$\lfloor \mathrm{g} \rfloor_j \{\mathrm{g}\}_i = 0 \qquad i \neq j$$

Because the order of multiplication is arbitrary, we assume $j > i$. Now, we recall that each new search direction is determined by

$$\{\mathrm{p}\}_i = \{\mathrm{g}\}_i - \lambda_i \{\mathrm{p}\}_{i-1}$$

therefore the gradient at any point j is

$$\{g\}_i = \lambda_i \{p\}_{i-1} + \{p\}_i$$

That is, the gradient at any point is a linear combination of the current search direction and the previous search direction, hence

$$\lfloor g \rfloor_j \{p\}_i = \lfloor g \rfloor_j \lambda_{i-1} \{p\}_{i-1} + \lfloor g \rfloor_j \{p\}_i$$

However, in the previous section we proved that a gradient at any point is orthogonal to all previous search directions. Because we have taken $j > i$, the left hand side of the above equation is indeed zero and we have proved what we set out to prove, i.e. that the gradients are mutually orthogonal.

<u>Are the Search Directions Mutually Conjugate?</u>

We now ask the question, If we choose each new direction (in a sequence of directions) by the method proposed, thus making each new direction conjugate to the direction just previously traversed, does that insure that the new direction is conjugate to all previous directions? The answer is yes and is shown by induction as follows.

It is possible to choose the first search direction arbitrarily, hence we always select the direction of steepest decent at the initial reference point. Our proposed method for determining the next direction assures us that it is conjugate to the previous one, hence to all previous ones (there has only been one other). We now show that if we have traversed $(j-1)$ search directions, and that all previous search directions are mutually A-conjugate, then the proposed procedure for the next direction makes it A-conjugate to all previous direcition. Hence the method results in a set of mutually A-conjugate directions.

At point $\{x\}_j$, the new search direction is

$$\{p\}_j = \{g\}_j - \lambda_j \{p\}_{j-1}$$

We now test to see if it is conjugate to all previous search direction $\{p\}_i$, where $i < (j-1)$. We therefore test

$$\{p\}_i [A] \{p\}_j = \{p\}_i [A] \{g\}_j - \{p\}_i [A] \lambda_j \{p\}_{j-1}$$

The last term on the RHS is zero because of $\{\mathrm{p}\}_i$ and $\{\mathrm{p}\}_{j-1}$ are A-conjugate. Therefore

$$\{\mathrm{p}\}_i[\mathrm{A}]\{\mathrm{p}\}_j = \{\mathrm{p}\}_i[\mathrm{A}]\{\mathrm{g}\}_j$$

To complete our proof we only need to show that the term on the right is zero. Note that $\{\mathrm{p}\}_i$ represents the direction between points $\{\mathrm{x}\}_i$ and $\{\mathrm{x}\}_{i+1}$. Specifically

$$\{\mathrm{p}\}_i = \frac{1}{\lambda_i}\left(\{\mathrm{x}\}_{i+1} - \{\mathrm{x}\}_i\right)$$

Therefore

$$\{\mathrm{p}\}_i[\mathrm{A}] = \frac{1}{\lambda_i}\left([\mathrm{A}]\{\mathrm{x}\}_{i+1} - [\mathrm{A}]\{\mathrm{x}\}_i\right)$$

Now,

$$[\mathrm{A}]\{\mathrm{x}\}_{i+1} = \{\mathrm{f}\} - \{\mathrm{g}\}_{i+1}$$

and

$$[\mathrm{A}]\{\mathrm{x}\}_i = \{\mathrm{f}\} - \{\mathrm{g}\}_i$$

hence

$$\{\mathrm{p}\}_i[\mathrm{A}] = \frac{1}{\lambda_{(j)}}\left(-\{\mathrm{g}\}_{i+1} + \{\mathrm{g}\}_i\right)$$

and

$$\{\mathrm{p}\}_i[\mathrm{A}]\{\mathrm{p}\}_j = \frac{1}{\lambda_{(j)}}\left(-\{\mathrm{g}\}_{i+1}\{\mathrm{g}\}_j + \{\mathrm{g}\}_i\{\mathrm{g}\}_j\right)$$

But, because the gradients are orthogonal the right-hand side is identically zero for any $i < (j-1)$ and we have proved what we set out to prove. Namely, the procedure for creating a new conjugate direction based on just the previous direction, creates a vector conjugate to all previous directions.

Algorithm

We have now shown the theoretical development of the conjugate gradient method. Although it has been rather lengthy, the actual implementation of the procedure is quite simple.

read $\{x\}$, ϵ	Enter starting point and tolerance
$\{r\} = \{f\} - [A]\{x\}$	Calculate initial residual
WHILE $\vert\{r\}\vert < \epsilon$	
IF starting point	
$\{p\} = \{r\}$	
ELSE	
$\mathrm{rAp} = \lfloor \mathrm{r} \rfloor \{\mathrm{Ap}\}$	
$\beta = \mathrm{rAp}/\mathrm{pAp}$	
$\{p\} = \{r\} - \beta\{p\}$	new search direction
ENDIF	
$\mathrm{rp} = \lfloor \mathrm{r} \rfloor \{\mathrm{p}\}$	
$\{\mathrm{Ap}\} = [\mathrm{A}]\{\mathrm{p}\}$	first matrix multiplication
$\mathrm{pAp} = \lfloor \mathrm{p} \rfloor \{\mathrm{Ap}\}$	
$\mathrm{s} = (\mathrm{rp})/(\mathrm{pAp})$	distance along search direction
$\{\mathrm{x}\} = \{\mathrm{x}\} + \mathrm{s}\{\mathrm{p}\}$	
$\{\mathrm{r}\} = \{\mathrm{f}\} - [\mathrm{A}]\{\mathrm{x}\}$	second matrix multiplication
$\vert\{r\}\vert < \epsilon$?	
CONTINUE	
Print Results	
STOP	

How to Eliminate the Second Matrix Multiplication

In the previous algorithm two matrix multiplications per iteration were shown. These multiplication represent the most computational intense part of the algorithm; hence, we now wish to show how to eliminate the second multiplication. Rather than calculate a new residual vector, it is possible to update the old residual. Here is the method:

$$\{r\}_{old} = \{f\} - [A]\{x\}_{old}$$

$$\{r\}_{new} = \{f\} - [A]\{x\}_{new}$$

but $$\{x\}_{new} = \{x\}_{old} + s\{p\}$$

hence $$\{r\}_{new} = \{f\} - [A]\{x\}_{old} - s[A]\{p\}$$

or, simply $$\{r\}_{new} = \{r\}_{old} - s\{Ap\}$$

Therefore, the line:

$\{r\} = \{f\} - [A]\{x\}$ second matrix multiplication

can be replaced by:

$\{r\} = \{r\} - s\{Ap\}$ a simple vector multiplcation

APPENDIX A

Relationship Between Conjugate Directions and Changes in Gradients.
We show in this section that when transversing in a direction $\{b\}$ that is conjugate to a direction $\{a\}$, the change in the gradient along $\{b\}$ never has a component in the $\{a\}$ direction. Thus, once we have found a minimum of Φ along $\{a\}$ such that there is no longer a component of the gradient in the direction of $\{a\}$, then as we move along $\{b\}$, the gradient will never develop a component in the direction of $\{a\}$; therefore, the gradient at everypoint along $\{b\}$ will be orthogonal to $\{a\}$.

Let the minimum point along $\{a\}$ be $\{b_o\}$ and consider the vector $\{b\}$ to be of variable length in a direction conjugate to $\{a\}$. Its gradient at any point will thus be

$$\{g\} = \{f\} - [A]\{\{b_o\} + \{b\}\}$$

The change in this gradient as we move from

$$\{b_o + b\}$$

to

$$\{b_o + b + db\}$$

is

$$\{dg\} = [A]\{db\}$$

But if $\{b\}$ is conjugate to $\{a\}$ then each increment along $\{b\}$ is conjugate to $\{a\}$. Therefore:

$$\lfloor a \rfloor [A]\{db\} = 0$$

hence

$$\lfloor a \rfloor \{dg\} = \lfloor a \rfloor [A]\{db\} = 0$$

QED

This relationship expresses the underlying concept behind the conjugate gradient method. It states that once we have found a minimum point along a direction $\{a\}$, that as we traverse a new direction $\{b\}$ conjugate to $\{a\}$, we will not undo our minimization along $\{a\}$ by moving in a direction that has a gadient to $\{a\}$. This implies what we have already found out - namely, that conjugate directions decouple the process of minimization.

APPENDIX B

FORTRAN Program for the Conjugate Gradient Method

```
                           Program Conjugate
c-----------------------------------------------------------------
c     A program for the solution of simultaneous linear
c     equations by the conjugate gradient method (cgm).
c
c                     [A]{x} - {f}  =  0
c
c-----------------------------------------------------------------

c     ----------------------------------------------------------
      parameter (INNI=50)
      COMMON/CsdmR/A(INNI,INNI),x(INNI),f(INNI),r(INNI),p(INNI),
     *             Ap(INNI)
      COMMON/CsdmI/numeq,mni,llpt,lcrd
c     ----------------------------------------------------------

      call openup
      call data

c     --------------------------
c     Calculate initial residual
c         {r} = {f} - [A]{x}
c     --------------------------
      do 120 j=1,numeq
         r(j)=  f(j)
120   continue

      do 130 j=1,numeq
          do 125 k=1,numeq
             r(j)=r(j)-a(j,k)*x(k)
125      continue
130   continue


c     ---------------------
c     Begin iterations
c     ---------------------
      write(llpt,4)
      do 500 i=1,mni

c        -----------------------------------------------------
c        Determine new search direction {h}
c        -----------------------------------------------------
         if(i.eq.1) then
c           ---------------------------------
c           Use steepest decent, {r}
c           ---------------------------------
            do 220 j=1,numeq
               p(j)=r(j)
220         continue
         else
c           ----------------------------------------
```

```
c           Make new direction conjugate to last {h}
c           --------------------------------------
            rAp=0.0
            do 230 j=1,numeq
               rAp=rAp + r(j)*Ap(j)
230         continue
            Beta =  rAp/pAp

            do 240 j=1,numeq
               p(j)=r(j)-Beta*p(j)
240         continue
         endif


c        -----------------------------------------------------
c        Determine distance along {h} to minimum point
c        -----------------------------------------------------
         rp=0.0
         do 320 j=1,numeq
            rp = rp + r(j)*p(j)

            Ap(j)=0.0
            do 310 k=1,numeq
               Ap(j) = Ap(j) + A(j,k)*p(k)
310         continue
320      continue

         pAp = 0.0
         do 330 j=1,numeq
            pAp = pAp + p(j)*Ap(j)
330      continue

         s = (rp)/(pAp)

c        -------------------------------------------
c        Update position vector and residual vector
c            {x} = {x} + s*{Ap}
c            {r} = {r} - s*{Ap}
c        -------------------------------------------
         do 350  j=1,numeq
            x(j)=x(j)+s*p(j)
            r(j)=r(j)-s*Ap(j)
350      continue

c        -------------------------------
c        Determine magnitude of residual
c        -------------------------------
         rmag=0.0
         do 450 j=1,numeq
            rmag=rmag+r(j)**2
450      continue
         rmag=sqrt(rmag)

         if(rmag.lt.1.0e-05) then
c           -----------------------
c           convergence is achieved.
c           print results and stop
c           -----------------------
            write(llpt,5)  rmag
            write(llpt,6)
            write(llpt,3)
            write(llpt,1) (x(ii),ii=1,numeq)
```

```
          write(*,1)    (x(ii),ii=1,numeq)
          stop
        endif

        write(llpt,5)  rmag,s

500   continue
      write(llpt,2)
      write(llpt,3)
      write(llpt,1) (x(ii),ii=1,numeq)
      write(*,1)    (x(ii),ii=1,numeq)
      stop

c     ----------------------------------------
c     F O R M A T   S T A T E M E N T S
c     ----------------------------------------
1     format(1p5e11.3)
2     format(///
     *      /'         ******************************'
     *      /'                DID NOT CONVERGE'
     *      /'             in specified number of '
     *      /'                   iterations'
     *      /'         ******************************'
     *     ///)
3     format(///
     *      /'                   FINAL VECTOR'
     *      /55('-'))
4     format(///
     *      /'         Magnitude '
     *      /'             of    '
     *      /'          Gradient        Alpha'
     *      /' ---------------------------------------')
5     format(3x,1p2e15.4)
6     format(///
     *      /'         ******************************'
     *      /'              CONVERGENCE ACHIEVED'
     *      /'         ******************************'
     *     ///)

      end
```

```
                                           subroutine OpenUp
c     ----------------------------------------------------------
      parameter (INNI=50)
      COMMON/CsdmR/A(INNI,INNI),x(INNI),f(INNI),r(INNI),p(INNI),
     *             Ap(INNI)
      COMMON/CsdmI/numeq,mni,llpt,lcrd
c     ----------------------------------------------------------
      lcrd=1
      llpt=2

      open(  unit   =  lcrd,
     *       file   = 'DATAI',
     *       form   = 'formatted',
     *       status = 'old')
      open(  unit   =  llpt,
     *       file   = 'DATAO',
     *       form   = 'formatted',
     *       status = 'unknown')

      return
      end



                                             subroutine data
c     ----------------------------------------------------------
      parameter (INNI=50)
      COMMON/CsdmR/A(INNI,INNI),x(INNI),f(INNI),r(INNI),p(INNI),
     *             Ap(INNI)
      COMMON/CsdmI/numeq,mni,llpt,lcrd
c     ----------------------------------------------------------

      read(lcrd,*) numeq,mni
      do 110 i=1,numeq
         read(lcrd,*)((A(i,j),j=1,numeq))
110   continue

      read(lcrd,*)(f(i),i=1,numeq)

      do 220 i=1,numeq
         x(i)=0.0
220   continue

      return
      end
```