

A Generalization of the Method of Conjugate Gradients for Solving Systems of Linear Algebraic Equations

It is the purpose of this note to establish a slight generalization of the method of conjugate gradients for solving linear equations. The generalization will then be used to provide a unified theory for two rather different algorithms which have been proposed for the case in which the matrix of the equations is not symmetric. The corresponding formulas for the determinant of the matrix will also be developed, although they are probably without much practical interest.

We shall use capital letters for matrices, Greek letters for vectors, and lower case Roman letters for scalars. All matrices will be real and of order $n \times n$, and all vectors will be real and of dimensionality n . We denote the transpose of a matrix or a vector by a prime.

Let S be a positive definite matrix. The method¹ of conjugate gradients is based upon the construction of two sequences of vectors $\gamma_0, \gamma_1, \dots$ and $\delta_0, \delta_1, \dots$, and a sequence of scalars a_0, a_1, \dots . The sequences are constructed by repeated substitution into the following formulas:

$$(1) \quad \gamma_0 = \delta_0,$$

$$(2) \quad a_N = \frac{\gamma_N' \gamma_N}{\delta_N' S \delta_N}, \quad N = 0, 1, \dots, n-1,$$

$$(3) \quad \gamma_{N+1} = \gamma_N - a_N S \delta_N, \quad N = 0, 1, \dots, n-1,$$

$$(4) \quad b_N = \frac{\gamma_{N+1}' \gamma_{N+1}}{\gamma_N' \gamma_N}, \quad N = 0, 1, \dots, n-2,$$

$$(5) \quad \delta_{N+1} = \gamma_{N+1} + b_N \delta_N, \quad N = 0, 1, \dots, n-2.$$

To avoid discussing exceptional cases, we adopt the convention that if $\delta_N = 0$ for some value of N , say $N = N_0$, then $\delta_{N_0+1}, \dots, \delta_{n-1}$ and $\gamma_{N_0+1}, \dots, \gamma_n$ are to be defined as being zero vectors.

HESTENES & STIEFEL¹ proved a long series of interesting results concerning the sequence of vectors generated by (1)–(5). In particular, they showed that the vectors $\gamma_0, \gamma_1, \dots, \gamma_{n-1}$ are orthogonal (or zero) and the vectors $\delta_0, \delta_1, \dots, \delta_{n-1}$ are S -orthogonal or conjugate with respect to S . Moreover, they show that the scalars a_N given by (2) are identical with the coefficients in the representation of γ_0 by a linear combination of the vectors $S\delta_N$. That is,

$$(6) \quad \gamma_0 = \sum_0^{n-1} a_k S \delta_k.$$

From this it can be deduced that the two quadratic forms $\gamma_N' S^{-1} \gamma_N$ and $(S^{-1} \gamma_N)' (S^{-1} \gamma_N)$ are reduced in value at each step.

Hestenes & Stiefel also mention without a specific proof that $\det S = (a_0 a_1 \cdots a_{n-1})^{-1}$, where $\det S$ stands for the determinant of S . This relation is valid only if γ_0 and S are such that $\delta_N \neq 0$, $N \leq n - 1$. A proof of this relation can be given as follows: Let Δ be the matrix whose columns are $\delta_0, \dots, \delta_{n-1}$, and let G be the matrix whose columns are $\gamma_0, \dots, \gamma_{n-1}$. Then by the S -orthogonality of the vectors δ_N , it follows that $\Delta' S \Delta = D$, where D is a principal diagonal matrix whose principal diagonal elements are $\delta_0' S \delta_0, \delta_1' S \delta_1, \dots, \delta_{n-1}' S \delta_{n-1}$. But from (5) it is seen that δ_N is the sum of γ_N and a linear combination of the vectors $\gamma_0, \gamma_1, \dots, \gamma_{N-1}$. Therefore by an elementary property of determinants, $\det \Delta = \det G$. Since the vectors γ_N are orthogonal, $G'G$ is a principal diagonal matrix whose diagonal elements are $\gamma_0' \gamma_0, \gamma_1' \gamma_1, \dots, \gamma_{n-1}' \gamma_{n-1}$. Putting these facts together, we get

$$(7) \quad \det S = \frac{\det D}{\det \Delta' \det \Delta} = \frac{\det D}{\det \Delta' \Delta} = \frac{\det D}{\det G' G}$$

$$= \prod_{k=0}^{n-1} \frac{\delta_k' S \delta_k}{\gamma_k' \gamma_k} = \frac{1}{a_0 a_1 \cdots a_{n-1}}.$$

We now consider the applicability of the construction given by (1)–(5) to the problem $A\xi = \eta$, where A is an arbitrary non-singular matrix, η is a given vector, and ξ is an unknown vector.

Choose an initial estimate ξ_0 of the solution $A^{-1}\eta$ of $A\xi = \eta$. Choose any positive definite S . Let B be an arbitrary non-singular matrix. Construct the sequences given by the algorithm with $\gamma_0 = \delta_0 = B(\eta - A\xi_0)$. Then by (6)

$$B(\eta - A\xi_0) = \sum_0^{n-1} a_k S \delta_k,$$

or

$$A^{-1}\eta = \xi_0 + \sum_0^{n-1} a_k A^{-1} B^{-1} S \delta_k.$$

To apply this result, it is advantageous to express S in the form

$$S = BATA'B',$$

where T is a positive definite matrix. Then the solution of $A^{-1}\eta$ is given by

$$A^{-1}\eta = \xi_0 + \sum_{k=0}^{n-1} a_k T A' B' \delta_k.$$

The solution can be expressed iteratively, if desired, by the recursion relation

$$(9) \quad \xi_{N+1} = \xi_N + a_N T A' B' \delta_N, \quad N = 0, \dots, n - 1.$$

From (7), we see that the absolute value of the determinant of A is given by

$$|\det A| = \frac{1}{(a_0 a_1 \cdots a_{n-1} \det T)^{\frac{1}{2}} |\det B|}.$$

For some purposes it might be convenient to rewrite the relevant parts of the algorithms (1)–(6) in terms of the residual vectors $\xi_N = \eta - A\xi_N$. To do this we notice that (9) and (3) imply that

$$B(\eta - A\xi_N) = B\rho_N = B\rho_0 - \sum_{k=0}^{N-1} a_k BATA'B'\delta_k = \gamma_N.$$

Thus we have only to replace γ_N by $B\rho_N$ in (1)–(5) to obtain the desired modification. The new relations are

$$(1)' \quad \delta_0 = B\rho_0,$$

$$(2)' \quad a_N = \frac{(B\rho_N)'B\rho_N}{\delta_N' BATA'B'\delta_N}, \quad N = 0, \dots, n - 1,$$

$$(3)' \quad \rho_{N+1} = \rho_N - a_N ATA'B'\delta_N, \quad N = 0, \dots, n - 1,$$

$$(4)' \quad b_N = \frac{(B\rho_{N+1})'B\rho_{N+1}}{(B\rho_N)'B\rho_N}, \quad N = 0, \dots, n - 2,$$

$$(5)' \quad \delta_{N+1} = B\rho_{N+1} + b_N\delta_N, \quad N = 0, \dots, n - 2.$$

These formulas, together with (9), represent the generalization of the conjugate gradient algorithm for the problem $A\xi = \eta$ which was promised in the first paragraph.

We shall now specialize T and B so as to obtain algorithms useful in practice.

Case 1. $B = I, T = A^{-1}, A$ positive definite. The algorithm (1)'–(5)', (9), assumes the standard fundamental form given by Hestenes & Stiefel.¹

Case 2. $B = A', T = (A'A)^{-1}$. The formulas (1)'–(5)' become:

$$\delta_0 = A'\rho_0,$$

$$a_N = \frac{(A'\rho_N)'A'\rho_N}{(A\rho_N)'(A\rho_N)}, \quad N = 0, \dots, n - 1,$$

$$\rho_{N+1} = \rho_N - a_N A\delta_N, \quad N = 0, \dots, n - 1,$$

$$b_N = \frac{(A'\rho_{N+1})'A'\rho_{N+1}}{(A'\rho_N)'A'\rho_N}, \quad N = 0, \dots, n - 2,$$

$$\delta_{N+1} = A'\rho_{N+1} + b_N\delta_N, \quad N = 0, \dots, n - 2.$$

The iterations (9) for the solution vector $\xi = A^{-1}\eta$ are given by

$$\xi_{N+1} = \xi_N + a_N\delta_N, \quad N = 0, \dots, n - 1,$$

and

$$|\det A| = (a_0 a_1 \dots a_{n-1})^{-1}.$$

This is the method of solution for the case of an arbitrary non-singular A proposed by Hestenes & Stiefel.¹ The algorithm constructs a set of $A'A$ -orthogonal vectors δ_N .

Case 3. $B = I$, $T = I$. The formulas (1)'–(5)' or (1)–(5) (there is no difference in this case) become:

$$\delta_0 = \rho_0,$$

$$a_N = \frac{\rho_N' \rho_N}{\delta_N' A A' \delta_N}, \quad N = 0, \dots, n-1,$$

$$\rho_{N+1} = \rho_N - a_N A A' \delta_N, \quad N = 0, \dots, n-1,$$

$$b_N = \frac{\rho_{N+1}' \rho_{N+1}}{\rho_N' \rho_N}, \quad N = 0, \dots, n-2,$$

$$\delta_{N+1} = \rho_{N+1} + b_N \delta_N, \quad N = 0, \dots, n-2.$$

The iterations for the solution vector $\xi = A^{-1}\eta$ are given by

$$\xi_{N+1} = \xi_N + a_N A' \delta_N, \quad N = 0, \dots, n-1.$$

The formula for the determinant is again

$$|\det A| = (a_0 a_1 \dots a_{n-1})^{-1}.$$

The algorithm constructs a set of AA' -orthogonal vectors δ_N instead of a set of $A'A$ -orthogonal vectors, as in Case 2.

This method of solution for the case of an arbitrary non-singular A was proposed by CRAIG.² The author learned of this method through a communication from Dr. Craig dated May 22, 1953, in which an algorithm essentially equivalent to the one given above was stated, but with ρ_N calculated at each step from the relation $\rho_N = \eta - A\xi_N$. It has an obvious advantage over the Hestenes-Stiefel proposal in that the number of arithmetic operations required is slightly smaller.

To be specific, if we measure the amount of computing work only by the required number of multiplications, assuming no zero elements, and counting a division as two multiplications, the count for the first step of the Craig method is $3n^2 + 5n + 4$, whereas that for the Hestenes-Stiefel method is $4n^2 + 6n + 4$. The detailed breakdown is as follows. First, for the Craig method, ρ_0 is n^2 multiplications, $A'\delta_0$ is then n^2 , a_0 is $2n + 2$, ξ_1 is n , ρ_1 is n^2 , $\rho_1'\rho_1$ is n , b_0 is 2 (we already have computed $\rho_0'\rho_0$ for a_0), δ_1 is n ; total $3n^2 + 5n + 4$. With the Hestenes-Stiefel method, ρ_0 is n^2 , δ_0 is n^2 , $A\delta_0$ is n^2 , a_0 is $2n + 2$, ξ_1 is n , ρ_1 is n (it is true that we already have computed $A\delta_0$, but not $a_0A\rho_0$ nor $Aa_0\delta_0$), $A'\rho_1$ is n^2 , $(A'\rho_1)'A'\rho_1$ is n , b_1 is 2, δ_1 is n ; total $4n^2 + 6n + 4$. We are obviously assuming full storage of all vectors computed in the previous step at each point in the computation. The detailed breakdown is given here because the author has encountered a number of apparent discrepancies in similar counts published elsewhere (not necessarily relating to the methods discussed here); these discrepancies doubtless arise largely from different orderings of operations and different assumptions as to what will be stored after computation.

After the first step there is little to choose between the two methods from the standpoint of the number of multiplications theoretically required. In the Craig method, each step after the first and before the last takes $2n^2 + 4n + 4$.

In the Hestenes-Stiefel method, the corresponding count is $2n^2 + 5n + 4$. In the latter method, if ρ_n in the last step were to be computed from the recursion relation, the method would win back the n^2 multiplications which it lost to the Craig method in the first step; but it would hardly be reasonable to calculate the last residual in this way.

We note that since the algorithm in either case will very probably last for n steps unless a peculiarly fortunate choice of ξ_0 was made, the total number of multiplications to be expected is of the order of $2n^3$. The eventual justification for the use of conjugate gradient methods in preference to the most economical direct methods (which theoretically require only $n^3/3 + O(n^2)$ multiplications) must rest on definitive *a priori* studies of round-off error, and on the peculiarities of individual computing machines, and on special characteristics of the matrix A . But it seems quite probable at this writing that no final recommendations will ever be formulated as to the "best" ways of solving linear equations.

There is a further practical use for the algorithm (1)'-(5)' other than the applications represented by the specializations described above. It may be possible under some circumstances to find a matrix B such that pre-multiplication of the equation $A\xi = \eta$ by B will improve the "condition"³ of A . Then (1)'-(5)', with T chosen conveniently, say $T = (A'A)^{-1}$, might give an advantageous algorithm to use. (With $T = (A'A)^{-1}$ the algorithm would be the same as that of Case 2 but with A replaced by B' .)

In conclusion, we note that Hestenes & Stiefel¹ assert (p. 424-425) that the most general cg-algorithm for any linear system is obtained from (1)'-(5)' by assuming that BA is positive definite and $T = (A'B')^{-1}$. Nevertheless without this specialization, the generality of (1)'-(5)' seems to exceed the limits specified by Hestenes & Stiefel, because here B is an arbitrary non-singular matrix. The disagreement, if it may be called that, seems to lie in the fact that their characterization of generality is derived from a certain generalization of the method of conjugate directions, and this generalization does not include all the possibilities represented by the algorithm given here.

J. H. CURTISS

New York Univ.
New York, N. Y.

¹ M. R. HESTENES & E. STIEFEL, "Methods of conjugate gradients for solving linear systems," NBS *Jn. of Research*, v. 49, 1952, p. 409-436.

² E. J. CRAIG, *Iteration Procedures for Simultaneous Equations*. Doctoral dissertation, Mass. Inst. Tech., Cambridge, Mass., 1954.

³ J. TODD, "The condition of certain matrices, I," *Quart. Jn. of Mech. and Appl. Math.*, v. 2, 1949, p. 469-472.

On the Estimation of Quadrature Errors for Analytic Functions

1. Introduction. The present paper sets forth a complex variable method for the estimation of errors which arise when approximate rules of quadrature are applied to analytic functions. In contrast to the usual real variable methods, this method does not involve the use of the higher derivatives of the function, but uses only a knowledge of the size of the function in the complex plane. It is therefore of practical value when dealing with situations