CHAPTER

7

Conjugate Gradient Acceleration

7.1 INTRODUCTION

In the early 1950s Hestenes and Stiefel [1952] presented a new iterative method for solving systems of linear algebraic equations. This new method was known as the "conjugate gradient method"; we shall refer to it as the *CG method*. The CG method, though an iterative method, converges to the true solution of the linear system in a finite number of iterations in the absence of rounding errors. Because of this and many other interesting properties, the CG method attracted considerable attention in the numerical analysis community when it was first presented. However, for various reasons the method was not widely used, and little was heard about it for many years. As noted by Concus *et al.* [1976b], there was hardly any mention of the CG method in the proceedings of a *Conference on Sparse Matrices and Their Applications* held in 1971 (see Rose and Willoughby [1972]).

Beginning in the mid-1960s there was a strong resurgence of interest in the CG method. A number of papers appeared, including those by Daniel [1965, 1967], J. K. Reid [1971, 1972], Bartels and Daniel [1974], Axelsson [1974], O'Leary [1975], Chandra *et al.* [1977], Concus *et al.* [1976b], and many others.

Although not generally recognized until fairly recently, the conjugate gradient method is not just one method, but a whole family of methods (Hestenes [1956]). Each such method can be regarded as an acceleration process for a particular (basic) linear stationary iterative method of first degree. (The classical CG method can be regarded as an acceleration procedure based on the RF method of Section 2.3). Moreover, as shown by Engeli *et al.* [1959], the CG method can be represented in a three-term form which resembles Chebyshev acceleration applied to the RF method. One can also develop a similar three-term form for conjugate gradient acceleration applied to more general basic methods which resembles Chebyshev acceleration applied to those methods. It can also be shown that conjugate gradient acceleration of a given iterative method converges with respect to a certain error measurement procedure at least as fast as the corresponding Chebyshev procedure. Furthermore, no parameter estimates are required in the implementation of conjugate gradient acceleration. Because of these and other attractive properties, conjugate gradient acceleration has been used extensively in recent years.

We describe the classical CG method in Section 7.2. In Section 7.3 we describe the equivalent three-term form. In Section 7.4 we describe conjugate gradient acceleration of a class of basic iterative methods. In Section 7.5 we describe procedures for deciding when to terminate the iterative process. Computational procedures are given in Section 7.6. In Section 7.7 numerical results based on simulation experiments are given.

7.2 THE CONJUGATE GRADIENT METHOD

We now describe the classical conjugate gradient method (CG method) of Hestenes and Stiefel [1952] as applied to the linear system $Au = b$ given by (2-1.1). We assume that the $N \times N$ matrix A is SPD.

The CG method can be regarded as a modification of the method of steepest descent. To derive the method of steepest descent we consider the quadratic form

$$F(u) = \tfrac{1}{2}(u, Au) - (b, u). \tag{7-2.1}$$

Since

$$F(u) = F(\bar{u}) + \tfrac{1}{2}((u - \bar{u}), A(u - \bar{u})), \tag{7-2.2}$$

where $\bar{u} = A^{-1}b$ is the solution of (2-1.1), and since A is SPD, it follows that the problem of solving $Au = b$ is equivalent to the problem of minimizing $F(u)$. Moreover, the gradient of $F(u)$ is given by

$$\text{grad } F(u) = b - Au. \tag{7-2.3}$$

The direction of the vector grad $F(u)$ is the direction for which the functional $F(u)$ at the point u has the greatest instantaneous rate of change. If $u^{(n)}$ is some approximation to $\bar{u}$, then in the method of steepest descent we obtain an improved approximation $u^{(n+1)}$ by moving in the direction of grad $F(u^{(n)})$ to a point where $F(u^{(n+1)})$ is minimal, i.e., $u^{(n+1)} = u^{(n)} + \lambda_n \operatorname{grad} F(u^{(n)})$, where λ_n is chosen to minimize $F(u^{(n+1)})$. Using (7-2.1), we easily calculate that $\lambda_n = (r^{(n)}, r^{(n)})/(r^{(n)}, Ar^{(n)})$, where $r^{(n)} \equiv b - Au^{(n)}$. Since, from (7-2.3), grad $F(u^{(n)}) = r^{(n)}$, we can express the method of steepest descent in the form

$$\begin{aligned} &u^{(0)} \text{ is arbitrary,} \\ &u^{(n+1)} = u^{(n)} + \lambda_n r^{(n)} \qquad \text{for} \quad n = 0, 1, \ldots, \\ &r^{(n)} = b - Au^{(n)} \\ &\lambda_n = \frac{(r^{(n)}, r^{(n)})}{(r^{(n)}, Ar^{(n)})}. \end{aligned} \tag{7-2.4}$$

For ill-conditioned matrices A, the convergence rate of the method of steepest descent can be very slow (see, e.g., Luenberger [1973]). However, by choosing our direction vectors differently, we obtain the CG method, which, as we shall see shortly, gives the solution in at most N iterations in the absence of rounding errors.

Let $u^{(0)}$ be arbitrary and let successive approximations to the solution $\bar{u}$ be given by $u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}$, where $p^{(n)}$ is a "direction vector." For the CG method, we let $p^{(0)} = r^{(0)}$ and $p^{(n)} = r^{(n)} + \alpha_n p^{(n-1)}$ for $n \geq 1$, where α_n is chosen so that $p^{(n)}$ is A-conjugate to $p^{(n-1)}$, i.e., $(p^{(n)}, Ap^{(n-1)}) = 0$. Evidently, $\alpha_n = -(r^{(n)}, Ap^{(n-1)})/(p^{(n-1)}, Ap^{(n-1)})$. As before, choosing λ_n to minimize $F(u^{(n+1)})$, we obtain $\lambda_n = (p^{(n)}, r^{(n)})/(p^{(n)}, Ap^{(n)})$. The formulas for the CG method are given by

$$\begin{aligned} &u^{(0)} \text{ is arbitrary,} \\ &u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}, && n = 0, 1, \ldots, \\ &p^{(n)} = \begin{cases} r^{(n)}, & \text{if } n = 0, \\ r^{(n)} + \alpha_n p^{(n-1)}, & n = 1, 2, \ldots, \end{cases} \\ &\alpha_n = -\frac{(r^{(n)}, Ap^{(n-1)})}{(p^{(n-1)}, Ap^{(n-1)})}, && n = 1, 2, \ldots, \\ &r^{(n)} = b - Au^{(n)}, && n = 0, 1, \ldots, \\ &\lambda_n = \frac{(p^{(n)}, r^{(n)})}{(p^{(n)}, Ap^{(n)})}, && n = 0, 1, \ldots. \end{aligned} \tag{7-2.5}$$

It can be shown that α_n, λ_n, and $r^{(n)}$ can be given equivalently by

$$\alpha_n = \frac{(r^{(n)}, r^{(n)})}{(r^{(n-1)}, r^{(n-1)})}, \qquad n = 1, 2, \ldots,$$

$$\lambda_n = \frac{(r^{(n)}, r^{(n)})}{(p^{(n)}, Ap^{(n)})}, \qquad n = 0, 1, \ldots, \tag{7-2.6}$$

$$r^{(n)} = r^{(n-1)} - \lambda_{n-1} Ap^{(n-1)}, \qquad n = 1, 2, \ldots.$$

Hestenes and Stiefel [1952] show that the residuals $r^{(0)}, r^{(1)}, \ldots$ and the direction vectors $p^{(0)}, p^{(1)}, \ldots$ generated by (7-2.5) satisfy the relations

$$\begin{aligned} &(r^{(i)}, r^{(j)}) = 0 \quad \text{for} \quad i \neq j, \\ &(p^{(i)}, Ap^{(j)}) = 0 \quad \text{for} \quad i \neq j, \\ &(r^{(i)}, Ap^{(j)}) = 0 \quad \text{for} \quad i \neq j \quad \text{and} \quad i \neq j + 1. \end{aligned} \tag{7-2.7}$$

Thus the residual vectors $r^{(0)}, r^{(1)}, \ldots$ are mutually orthogonal and the direction vectors $p^{(0)}, p^{(1)}, \ldots$ are mutually A-conjugate. From the first relation in (7-2.7), it follows that $r^{(s)} = 0$ for some $s \leq N$. Thus the method (7-2.5) converges, in the absence of rounding errors, in at most N iterations. Hestenes and Stiefel [1952] also show that the error vector $\varepsilon^{(n)} \equiv u^{(n)} - \bar{u}$ associated with the CG method satisfies

$$\|\varepsilon^{(n+1)}\|_2 < \|\varepsilon^{(n)}\|_2 \tag{7-2.8}$$

whenever $\varepsilon^{(n)} \neq 0$. In the next section, we shall have more to say concerning the average rate of convergence for the CG method.

It can be shown (see, e.g., Beckman [1960]) that the direction vector $p^{(n)}$ is a scalar multiple of the projection of the gradient vector $r^{(n)} = \text{grad } F(u^{(n)})$ in the linear space spanned by $p^{(n)}, p^{(n+1)}, \ldots, p^{(N-1)}$. This fact coupled with the fact that the direction vectors $p^{(0)}, p^{(1)}, \ldots$ are mutually A-conjugate accounts for the name "conjugate gradient method."

The CG method is a special case of the more general conjugate direction (CD) method. In the CD method, the vectors $p^{(0)}, p^{(1)}, \ldots, p^{(N-1)}$ are selected to be nonzero and mutually A-conjugate but have no further restrictions. To describe the basic idea involved, suppose that the set $\{p^{(n)}\}_{n=0}^{n=N-1}$ of nonzero mutually A-conjugate vectors is given. Since A is SPD, it is easy to show that the set $\{p^{(n)}\}_{n=0}^{n=N-1}$ is also linearly independent. Thus there exist constants $c_0, c_1, \ldots, c_{N-1}$ such that

$$\bar{u} = u^{(0)} + c_0 p^{(0)} + c_1 p^{(1)} + \cdots + c_{N-1} p^{(N-1)}, \tag{7-2.9}$$

where $u^{(0)}$ is some initial approximation to $\bar{u}$. Multiplying (7-2.9) by A and taking inner products with $p^{(n)}$, we obtain

$$c_n = \frac{(p^{(n)}, b - Au^{(0)})}{(p^{(n)}, Ap^{(n)})}, \qquad n = 0, \ldots, N-1. \tag{7-2.10}$$

Note that the constants c_n are easily calculable.†

The CD method is given by the formulas

$$\begin{aligned} &u^{(0)} \text{ is arbitrary,} \\ &r^{(n)} = b - Au^{(n)}, && n = 0, 1, \ldots, \\ &\lambda_n = \frac{(p^{(n)}, r^{(n)})}{(p^{(n)}, Ap^{(n)})}, && n = 0, 1, \ldots, \\ &u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}, && n = 0, 1, \ldots. \end{aligned} \tag{7-2.11}$$

It can be shown (see, e.g., Luenberger [1973]) that the λ_n in (7-2.11) are equal to the c_n from (7-2.10) and that the iterates $u^{(n+1)}$ in (7-2.11) can be expressed in the form

$$u^{(n+1)} = u^{(0)} + c_0 p^{(0)} + c_1 p^{(1)} + \cdots + c_n p^{(n)}, \tag{7-2.12}$$

where the c_i, $i = 0, \ldots, n$, are those given in (7-2.10). From (7-2.12) and (7-2.9), it follows that $u^{(n)} = \bar{u}$ for some $n \leq N$. Thus the CD method also enjoys the property that convergence is achieved, in the absence of rounding errors, in at most N iterations.

The CD method is not well defined in that no prescription is given for the computation of the direction vectors $p^{(0)}, p^{(1)}, \ldots$. Various formulas can be given, with each leading to a special method. We can generate an A-conjugate set of vectors from any set $\{v^{(n)}\}_{n=0}^{n=N-1}$ of linearly independent vectors, using the Gram–Schmidt orthogonalization procedure. Hestenes and Stiefel [1952] show that the CD method is equivalent to the Gauss elimination method when the set $\{v^{(n)}\}$ is chosen to be the unit basis vectors, i.e., when $v^{(0)} = [1, 0, \ldots, 0]^T$, $v^{(1)} = [0, 1, 0, \ldots, 0]^T$, etc. For the CG method, the set $\{v^{(n)}\}$ is chosen to be the residual vectors, i.e., $v^{(n)} = r^{(n)}$. The residual and direction vectors for the CG method are not defined beforehand but are determined sequentially in the order $r^{(0)}$, $p^{(0)}$, $r^{(1)}$, $p^{(1)}, \ldots$ as the iterations progress.

† If the set $\{p^{(n)}\}_{n=0}^{n=N-1}$ were orthogonal and not A-conjugate, then $c_n = (p^{(n)}, (\bar{u} - u^{(0)}))/(p^{(n)}, p^{(n)})$. This expression for the c_n is of little help since $\bar{u}$ is not known.

7.3 THE THREE-TERM FORM OF THE CONJUGATE GRADIENT METHOD

Engeli *et al.* [1959] considered the following three-term form of the CG method:

$$u^{(n+1)} = \rho_{n+1}\{\gamma_{n+1}r^{(n)} + u^{(n)}\} + (1 - \rho_{n+1})u^{(n-1)}, \tag{7-3.1}$$

where

$$\gamma_{n+1} = \frac{(r^{(n)}, r^{(n)})}{(r^{(n)}, Ar^{(n)})} \tag{7-3.2}$$

and

$$\rho_{n+1} = \left[1 - \frac{\gamma_{n+1}}{\gamma_n}\frac{(r^{(n)}, r^{(n)})}{(r^{(n-1)}, r^{(n-1)})}\frac{1}{\rho_n}\right]^{-1}, \qquad \text{if} \quad n \geq 1 \quad (\rho_1 = 1). \tag{7-3.3}$$

The above formulas can be obtained from (7-2.5) by eliminating $p^{(n)}$ and $p^{(n-1)}$ from the pair of equations $u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}$ and $u^{(n)} = u^{(n-1)} + \lambda_{n-1}p^{(n-1)}$. Thus we obtain (7-3.1) with

$$\rho_{n+1} = 1 + \alpha_n\lambda_n/\lambda_{n-1}, \qquad \text{if} \quad n \geq 1 \quad (\rho_1 = 1) \tag{7-3.4}$$

and

$$\gamma_{n+1} = \lambda_n/\rho_{n+1}. \tag{7-3.5}$$

Formulas (7-3.2) and (7-3.3) can then be derived directly from (7-2.5) (see, e.g., Reid [1971]).

An alternative derivation of (7-3.2) and (7-3.3) can be given as follows (see Concus *et al.* [1976b]). By (7-2.5) and (7-3.1), we have

$$r^{(n+1)} = \rho_{n+1}\{-\gamma_{n+1}Ar^{(n)} + r^{(n)}\} + (1 - \rho_{n+1})r^{(n-1)}. \tag{7-3.6}$$

We now use the fact that the residuals are mutually orthogonal. If we require that $(r^{(n+1)}, r^{(n)}) = 0$, we get (7-3.2), provided that $\rho_{n+1} \neq 0$. But by (7-3.4) and (7-2.6) it follows that

$$\rho_{n+1} \geq 1, \qquad n = 0, 1, \ldots. \tag{7-3.7}$$

If we require that $(r^{(n+1)}, r^{(n-1)}) = 0$ we get, assuming that $(r^{(n)}, r^{(n-1)}) = 0$,

$$\begin{aligned} 0 &= (r^{(n+1)}, r^{(n-1)}) \\ &= \rho_{n+1}\{-\gamma_{n+1}(Ar^{(n)}, r^{(n-1)})\} + (1 - \rho_{n+1})(r^{(n-1)}, r^{(n-1)}). \end{aligned} \tag{7-3.8}$$

Replacing n by $n-1$ in (7-3.6) and taking the inner product of both sides with $r^{(n)}$, we get

$$(r^{(n)}, r^{(n)}) = \rho_n(-\gamma_n(Ar^{(n-1)}, r^{(n)})) \tag{7-3.9}$$

or

$$\begin{aligned}(Ar^{(n-1)}, r^{(n)}) &= (r^{(n-1)}, Ar^{(n)}) = (Ar^{(n)}, r^{(n-1)}) \\ &= -(r^{(n)}, r^{(n)})/\rho_n\gamma_n. \end{aligned} \tag{7-3.10}$$

Substituting (7-3.10) into (7-3.8), we get (7-3.3).

Replacing $r^{(n)}$ by $b - Au^{(n)}$ in (7-3.1), we can express $u^{(n+1)}$ in the alternative form

$$u^{(n+1)} = \rho_{n+1}\{\gamma_{n+1}((I - A)u^{(n)} + b) + (1 - \gamma_{n+1})u^{(n)}\} + (1 - \rho_{n+1})u^{(n-1)}. \tag{7-3.11}$$

Since $(I - A)$ is the iteration matrix for the RF method (see Section 2.3), it follows from Theorem 3-2.1 that the iterates (7-3.11) correspond to a polynomial acceleration procedure applied to the RF method. Thus from (3-2.5), there exists a matrix polynomial $Q_n(G) = \alpha_{n,0}I + \alpha_{n,1}G + \cdots + \alpha_{n,n}G^n$ such that the error vector $\varepsilon^{(n)} \equiv u^{(n)} - \bar{u}$ associated with (7-3.11) can be expressed in the form

$$\varepsilon^{(n)} = Q_n(G)\varepsilon^{(0)} = Q_n(I - A)\varepsilon^{(0)}.$$

It can be shown (see, e.g., Young *et al.* [1980]) that the property that the residual vectors are mutually orthogonal characterizes the CG method among all polynomial acceleration procedures applied to the RF method.

It can also be shown (see, e.g., Young *et al.* [1980]) that the CG method minimizes the $A^{1/2}$-norm of the error vector among all polynomial acceleration methods applied to the RF method. Thus if $\varepsilon^{(n)}$ is the error vector associated with (7-3.11) and if $\tilde{\varepsilon}^{(n)}$, where $\tilde{\varepsilon}^{(0)} = \varepsilon^{(0)}$, is the error vector associated with any other polynomial method applied to the RF method, then $\|\varepsilon^{(n)}\|_{A^{1/2}} \leq \|\tilde{\varepsilon}^{(n)}\|_{A^{1/2}}$. In particular, if $\tilde{\varepsilon}^{(n)}$ corresponds to the Chebyshev acceleration method, it then follows from (3-2.12) and (4-2.20), since $A^{1/2}$ is a symmetrization matrix for the RF method, that

$$\|\varepsilon^{(n)}\|_{A^{1/2}} \leq \|\tilde{\varepsilon}^{(n)}\|_{A^{1/2}} \leq \frac{2\bar{r}^{n/2}}{1 + \bar{r}^n}\|\varepsilon^{(0)}\|_{A^{1/2}}, \tag{7-3.12}$$

where

$$\bar{r} = (1 - \sqrt{1 - \bar{\sigma}^2})/(1 + \sqrt{1 - \bar{\sigma}^2}) \tag{7-3.13}$$

and

$$\bar{\sigma} = [M(A) - m(A)]/[M(A) + m(A)] = [\kappa(A) - 1]/[\kappa(A) + 1]. \tag{7-3.14}$$

Here $\kappa(A)$ is the spectral condition number of A defined by (1-4.18).

7.4 CONJUGATE GRADIENT ACCELERATION

In the previous section we have shown how the CG method may be regarded as a polynomial acceleration procedure based on the RF method. We now show how the CG method may be modified so as to correspond to a polynomial acceleration procedure applied to more general basic iteration methods.

Let us consider the basic iterative method defined by

$$u^{(n+1)} = Gu^{(n)} + k, \tag{7-4.1}$$

where for some nonsingular splitting matrix Q we have

$$G = I - Q^{-1}A, \qquad k = Q^{-1}b. \tag{7-4.2}$$

We assume that the method is *symmetrizable* in the sense of Definition 2-2.1. Thus there exists a nonsingular symmetrization matrix W such that $W(I - G)W^{-1}$ is SPD. If A is SPD, then our discussion includes the RF, Jacobi, and SSOR methods, as well as any method in which the splitting matrix Q is SPD. (In the latter case we can let W be any matrix such that $W^{T}W = Q$.)

To derive a CG acceleration procedure based on the iterative method (7-4.1), we first construct a new linear system which has the same solution as the original system (2-1.1). The coefficient matrix of the new system will be SPD and generally will have a much smaller spectral condition number than that of the matrix A of the original linear system.

To derive the new system, we first consider the related linear system (2-2.2):

$$(I - G)u = k, \tag{7-4.3}$$

which, by complete consistency,† has the same solution as (2-1.1). We next multiply both sides of (7-4.3) by a symmetrization matrix W, obtaining

$$W(I - G)u = Wk. \tag{7-4.4}$$

The matrix $W(I - G)$ is not in general symmetric. However, by introducing a new vector $\hat{u} = Wu$, we can write (7-4.4) in the form

$$\hat{A}\hat{u} = \hat{b}, \tag{7-4.5}$$

where

$$\hat{A} = W(I - G)W^{-1}, \qquad \hat{u} = Wu, \qquad \hat{b} = Wk. \tag{7-4.6}$$

† See Section 2.2.

The system (7-4.5) is often called the preconditioned system† since, in general, the condition number of $\hat{A}$ is much less than that of A.

We remark that one can also obtain the preconditioned system from (2-1.1) as follows. Let Q be the splitting matrix corresponding to (7-4.1). We first multiply both sides of (2-1.1) by Q^{-1}, obtaining

$$Q^{-1}Au = Q^{-1}b,$$

which is the same as (7-4.3) by (7-4.2). Then we multiply both sides by W and replace u by $W^{-1}\hat{u}$. We then obtain (7-4.5), where

$$\hat{A} = WQ^{-1}AW^{-1}, \qquad \hat{u} = Wu, \qquad \hat{b} = WQ^{-1}b.$$

If we apply the CG method of Section 7.2 to the preconditioned system (7-4.5), we obtain, using (7-2.5) and (7-2.6), that

$$\begin{aligned}
&u^{(0)} \text{ is arbitrary,}\\
&u^{(n+1)} = u^{(n)} + \lambda_n p^{(n)}, \qquad n = 0, 1, \ldots,\\
&p^{(n)} = \begin{cases} \delta^{(0)}, & n = 0,\\ \delta^{(n)} + \alpha_n p^{(n-1)}, & n = 1, 2, \ldots,\end{cases}\\
&\alpha_n = -\frac{(W\delta^{(n)}, W(I-G)p^{(n-1)})}{(Wp^{(n-1)}, W(I-G)p^{(n-1)})} \qquad (7\text{-}4.7)\\
&\quad = \frac{(W\delta^{(n)}, W\delta^{(n)})}{(W\delta^{(n-1)}, W\delta^{(n-1)})}, \qquad n = 1, 2, \ldots,\\
&\lambda_n = \frac{(Wp^{(n)}, W\delta^{(n)})}{(Wp^{(n)}, W(I-G)p^{(n)})}\\
&\quad = \frac{(W\delta^{(n)}, W\delta^{(n)})}{(Wp^{(n)}, W(I-G)p^{(n)})}, \qquad n = 0, 1, 2, \ldots.
\end{aligned}$$

Here $\delta^{(n)}$ is the pseudoresidual vector

$$\delta^{(n)} = Gu^{(n)} + k - u^{(n)}. \qquad (7\text{-}4.8)$$

This extension of the CG method is equivalent to that given by Hestenes [1956]. (See also Daniel [1965], [1967].)

From the first relation given in (7-2.7), it follows that the pseudoresidual vectors $\delta^{(0)}, \delta^{(1)}, \ldots$, defined in (7-4.8), are mutually W-orthogonal in the sense that

$$(W\delta^{(i)}, W\delta^{(j)}) = 0, \qquad i \neq j. \qquad (7\text{-}4.9)$$

† Preconditioning was used by Evans [1967], Axelsson [1974], and others. The preconditioning used here is slightly more general than that of Evans and Axelsson but reduces to theirs if Q is SPD and if $W = Q^{1/2}$ or if W is any matrix such that $W^TW = Q$.

The second relation in (7-2.7) implies that the direction vectors $p^{(0)}, p^{(1)}, \ldots$ of (7-4.7) are $W^{\mathrm{T}}W(I - G)$-conjugate, i.e., that

$$(p^{(i)}, W^{\mathrm{T}}W(I - G)p^{(j)}) = 0, \qquad i \neq j. \tag{7-4.10}$$

Analogous to (7-3.1)–(7-3.3) we can obtain the following three-term form of the CG acceleration procedure (7-4.7):

$$u^{(n+1)} = \rho_{n+1}\{\gamma_{n+1}\delta^{(n)} + u^{(n)}\} + (1 - \rho_{n+1})u^{(n-1)}, \tag{7-4.11}$$

where

$$\gamma_{n+1} = \frac{(W\delta^{(n)}, W\delta^{(n)})}{(W\delta^{(n)}, W(I - G)\delta^{(n)})} = \left[1 - \frac{(W\delta^{(n)}, WG\delta^{(n)})}{(W\delta^{(n)}, W\delta^{(n)})}\right]^{-1}, \tag{7-4.12}$$

$$\rho_1 = 1, \quad \rho_{n+1} = \left[1 - \frac{\gamma_{n+1}}{\gamma_n}\frac{(W\delta^{(n)}, W\delta^{(n)})}{(W\delta^{(n-1)}, W\delta^{(n-1)})}\frac{1}{\rho_n}\right]^{-1}, \quad \text{if} \quad n \geq 1. \tag{7-4.13}$$

The above method represents a slight extension of the "generalized conjugate gradient procedure" presented by Concus *et al.* [1976]; see also Axelsson [1974]. It is equivalent to their method if Q is SPD and $W = Q^{1/2}$ or W is any matrix such that $W^{\mathrm{T}}W = Q$.

We refer to the methods obtained by applying CG acceleration to the RF, Jacobi, and SSOR methods as the RF–CG, J–CG, and SSOR–CG methods, respectively. Thus in our terminology, the classical CG method presented previously in Sections 7.2 and 7.3 is the same as the RF–CG method.

Concerning the choice between the two-term form (7-4.7) and the three-term form (7-4.11)–(7-4.13) of CG acceleration, results of Reid [1972] indicate that the two-term form is somewhat more efficient. On the other hand, the three-term form is the same as that used for Chebyshev acceleration. In any case, the difference between the two- and three-term forms does not appear to be very significant.

It can easily be shown that CG acceleration of the method (7-4.1) minimizes the $[W^{-1}W(I - G)]^{1/2}$-norm† of the error as compared with any polynomial acceleration procedure based on (7-4.1). This follows from (7-4.5) and (7-4.6). Indeed, we have

$$(\hat{\varepsilon}^{(n)}, \hat{A}\hat{\varepsilon}^{(n)}) = (W\varepsilon^{(n)}, W(I - G)\varepsilon^{(n)}) = (\varepsilon^{(n)}, W^{\mathrm{T}}W(I - G)\varepsilon^{(n)}), \tag{7-4.14}$$

where we let $\hat{\varepsilon}^{(n)} = W\varepsilon^{(n)}$. Thus the CG acceleration procedure applied to (7-4.3) minimizes $(\hat{\varepsilon}^{(n)}, \hat{A}\hat{\varepsilon}^{(n)})$, which is equal to the square of the $[W^{\mathrm{T}}W(I - G)]^{1/2}$-norm of the error $\varepsilon^{(n)}$. If A and Q are SPD and if $W^{\mathrm{T}}W = Q$, then we minimize the $A^{1/2}$-norm of the error as in the CG method.

† In order for the square root to be well defined, $W^{-1}W(I - G)$ must be SPD; but this follows from the facts that $W^{\mathrm{T}}W(I - G) = W^{\mathrm{T}}W(I - G)W^{-1}W$ and that $W(I - G)W^{-1}$ is SPD.

As in (7-3.12), we can obtain a bound for the error vector associated with the CG procedure (7-4.7), using known results for the Chebyshev procedure. We first note that $[W^TW(I - G)]^{1/2}$ is a symmetrization matrix for the basic method (7-4.1). To show this, since $W^TW(I - G)$ is SPD, we can write

$$\begin{aligned}[W^TW(I - G)]^{1/2}(I - G)[W^TW(I - G)]^{-1/2} \\ = [W^TW(I - G)]^{1/2}(W^TW)^{-1}[W^TW(I - G)]^{1/2} \\ = \{[W^TW(I - G)]^{1/2}W^{-1}\}\{[W^TW(I - G)]^{1/2}W^{-1}\}^T, \quad (7\text{-}4.15)\end{aligned}$$

from which the desired result follows. Now let $\varepsilon^{(n)}$ and $\tilde{\varepsilon}^{(n)}$, respectively, be the error vectors associated with the CG and Chebyshev acceleration methods applied to (7-4.1). We assume $\tilde{\varepsilon}^{(0)} = \varepsilon^{(0)}$. Using the error minimization property of CG acceleration and using (3-2.12) and (4-2.20) for Chebyshev acceleration, we obtain

$$\|\varepsilon^{(n)}\|_{[W^TW(I-G)]^{1/2}} \leq \|\tilde{\varepsilon}^{(n)}\|_{[W^TW(I-G)]^{1/2}} \leq \frac{2\bar{r}^{n/2}}{1 + \bar{r}^n}\|\varepsilon^{(0)}\|_{[W^TW(I-G)]^{1/2}}, \quad (7\text{-}4.16)$$

where $\bar{r}$ is given by (4-2.19). From (7-4.16) and (2-2.8), it follows that the average rate of convergence for the CG acceleration method, when measured in the $[W^TW(I - G)]^{1/2}$ norm, is at least as large as that for the corresponding Chebyshev procedure.

7.5 STOPPING PROCEDURES

We now describe a procedure for deciding when the CG iterative procedure of Section 7.4 should be terminated. Ideally, we should like to stop the iterative process and accept $u^{(n)}$ as a satisfactory approximation to the true solution $\bar{u}$ whenever $u^{(n)}$ satisfies the inequality

$$E_T \equiv \|u^{(n)} - \bar{u}\|_W / \|\bar{u}\|_W \leq \zeta, \quad (7\text{-}5.1)$$

where ζ is the stopping criterion number. As in previous chapters, we shall express the unknown quantity $u^{(n)} - \bar{u}$ in terms of $\delta^{(n)}$.

By (5-2.5), it follows that $\|u^{(n)} - \bar{u}\|_W \leq (1 - M(G))^{-1}\|\delta^{(n)}\|_W$, and hence we have

$$\frac{\|u^{(n)} - \bar{u}\|_W}{\|\bar{u}\|_W} \leq \frac{1}{1 - M(G)}\frac{\|\delta^{(n)}\|_W}{\|\bar{u}\|_W}. \quad (7\text{-}5.2)$$