Sufficient Conditions for a Symmetric Definite Multigrid V-Cycle for Preconditioned Conjugate Gradient

Jeffrey Lee Hellrung, Jr.

July, 2011

Abstract

We give sufficient yet very general conditions on the relaxer in a multigrid v-cycle to ensure symmetry and definiteness of the v-cycle when viewed as a linear operator on the right-hand side. Symmetry and definiteness of the v-cycle are necessary when it is used as a preconditioner for conjugate gradient. We additionally show that a wide variety of common relaxers indeed satisfy these sufficient conditions.

1 Introduction

Let us consider the solution to the linear system $Ax = b$ for symmetric A with initial guess $x = x_0$. Let $\widetilde{A}$ denote a symmetric matrix. Typically, $\widetilde{A}$ will be an approximate inverse to A; for example, in a multigrid context, $\widetilde{A}$ may be an approximate or exact coarse grid inverse conjugated with restriction and prolongation operators. We will make additional assumptions on A and $\widetilde{A}$ later as needed.

A *relaxer* is given by a pair of matrices (S, T) such that the *action* of the relaxer is effected by $x \mapsto Sx+Tb$. Examples of such relaxers are those derived from a splitting $A = P-Q$ with P invertible in which $S = P^{-1}Q$ and $T = P^{-1}$; we will investigate this class of relaxers in more detail later. We borrow the term "relaxer" from multigrid theory, where the term "smoother" is also used, even though we will not be directly concerned with any actual relaxation (or smoothing) properties.

Given $\widetilde{A}$ and given a *pre-relaxer* (S_0, T_0) and *post-relaxer* (S_1, T_1) (together referred to as a *relaxer pair*), a complete *cycle* is given by

$$\begin{aligned} x^{(1)} &\leftarrow S_0 x_{\text{in}} + T_0 b; \quad \text{(pre-relaxation)} \\ x^{(2)} &\leftarrow x^{(1)} + \widetilde{A}\left(b - Ax^{(1)}\right); \quad (\widetilde{A}\text{-correction}) \\ x_{\text{out}} &\leftarrow S_1 x^{(2)} + T_1 b; \quad \text{(post-relaxation)} \end{aligned} \tag{1}$$

or, taken altogether,

$$\begin{aligned} x_{\text{out}} &\leftarrow S_1\left(I - \widetilde{A}A\right)S_0 x_{\text{in}} + \left(S_1\left(\widetilde{A} - \widetilde{A}AT_0 + T_0\right) + T_1\right)b && \text{(2a)} \\ &=: Lx_{\text{in}} + Mb, && \text{(2b)} \end{aligned}$$

where I denotes the identity matrix. The *iteration matrix* of the cycle (1) is thus M (via setting $x_{\text{in}} = 0$). For brevity of exposition, we say that a relaxer pair $\{(S_i, T_i)\}_{i=0,1}$ and the corresponding cycle (1) are symmetric/(semi-) definite precisely when the corresponding iteration matrix M (2) is symmetric/(semi-) definite.

2 Symmetry

We first investigate the conditions under which the cycle (1) is symmetric. The next theorem gives relatively simple such sufficient conditions.

Theorem 1. *The iteration matrix M (2) of the cycle* (1) *is symmetric if*

(S0) $S_1 + T_0^t A = I$*; and*

(S1) $T_0 + T_1 = (T_0 + T_1)^t$.

Proof. Condition (S0) allows us to write

$$\begin{aligned} M &= S_1 \left(\widetilde{A} - \widetilde{A} A T_0 + T_0\right) + T_1 \\ &= S_1 \widetilde{A} \left(I - A T_0\right) + S_1 T_0 + T_1 \\ &= S_1 \widetilde{A} S_1^t + \left(I - T_0^t A\right) T_0 + T_1 \\ &= S_1 \widetilde{A} S_1^t + \left(T_0 + T_1\right) - T_0^t A T_0, \end{aligned}$$

which is clearly symmetric given condition (S1) and the symmetry of A and $\widetilde{A}$. □

Conditions (S0) and (S1) in Theorem 1 are actually more general than necessary, as we are typically only interested in *consistent* relaxers, that is, those satisfying $S + TA = I$ (i.e., $Sx + Tb = x$ whenever x satisfies $Ax = b$). In this context, condition (S0) is equivalent to $T_0 = T_1^t$ whenever A is invertible. As this simple transpose relation additionally implies condition (S1), we will, from now on, assume $T_0 = T_1^t$ for all subsequent relaxer pairs $\{(S_i, T_i)\}_{i=0,1}$, and by Theorem 1 all such relaxer pairs are symmetric. Together with the consistency relation $S_i + T_i A = I$, we see that such symmetric relaxer pairs are entirely determined by the matrix $R := T_0$ via the relations

$$(S_0, T_0) := (I - RA, R)\,, \tag{4a}$$
$$(S_1, T_1) := \left(I - R^t A, R^t\right). \tag{4b}$$

Expressing L and M in terms of R then gives

$$L = \left(I - R^t A\right)\left(I - \widetilde{A} A\right)(I - RA)\,, \tag{5a}$$
$$M = \left(I - R^t A\right) \widetilde{A} \left(I - AR\right) + \left(R + R^t - R^t A R\right). \tag{5b}$$

3 Composition

We are also interested in symmetry-preserving compositions of symmetric relaxer pairs. We compose two relaxer pairs $\{(S_i^a, T_i^a)\}_{i=0,1}$ and $\left\{\left(S_i^b, T_i^b\right)\right\}_{i=0,1}$ (specified by R^a and R^b, respectively, via (4)) into a *composed cycle* as

$$\begin{aligned} x^{(1)} &\leftarrow S_0^a x_{\text{in}} + T_0^a b; \quad (a\text{-pre-relaxation}) \\ x^{(2)} &\leftarrow S_0^b x^{(1)} + T_0^b b; \quad (b\text{-pre-relaxation}) \\ x^{(3)} &\leftarrow x^{(2)} + \widetilde{A}\left(b - A x^{(2)}\right); \quad (\widetilde{A}\text{-correction}) \\ x^{(4)} &\leftarrow S_1^b x^{(3)} + T_1^b b; \quad (b\text{-post-relaxation}) \\ x_{\text{out}} &\leftarrow S_1^a x^{(4)} + T_1^a b; \quad (a\text{-post-relaxation}) \end{aligned} \tag{6}$$

giving a *composed relaxer pair*

$$(S_0, T_0) := \left(S_0^b S_0^a, S_0^b T_0^a + T_0^b\right), \tag{7a}$$
$$(S_1, T_1) := \left(S_1^a S_1^b, S_1^a T_1^b + T_1^a\right). \tag{7b}$$

This composed relaxer pair can be specified by

$$R := R^a + R^b - R^b A R^a, \tag{8}$$

i.e., one can quickly verify that with the above R, $\{(S_i, T_i)\}_{i=0,1}$ from (7) are given by the relations (4). Theorem 1 thus implies that the composed cycle (6) is also symmetric, with no additional assumptions necessary on the component relaxer pairs.

In summary, composing any number of symmetric relaxer pairs (including iterating a single pre-relaxer and post-relaxer a fixed number of times) preserves symmetry in the iteration matrix as long as one simply reverses the application order of the post-relaxers relative to the corresponding pre-relaxers.

4 Definiteness

We now turn to the issue of definiteness of the cycle (1). Here, we use the term "definite" to mean (consistently) either "positive definite" or "negative definite", as the following discussion applies to either case.

Assuming that $\widetilde{A}$ is semi-definite, it is clear from (5) that M will be (semi-)definite if $W := R + R^t - R^t A R$ is (semi-)definite. This is actually not so hard to verify for common relaxer pairs given additional assumptions on A, as we will see later. However, it is less trivial to deduce the definiteness of the composed cycle (6), which the following theorem addresses.

Theorem 2. *Suppose $W^z := R^z + (R^z)^t - (R^z)^t A R^z$ is semi-definite, $z = a, b$. Then $W := R + R^t - R^t A R$, with R as defined in* (8)*, is also semi-definite.*

Proof. Some algebraic manipulations, after substituting $R = R^a + R^b - R^b A R^a$, give

$$W = W^a + \left(I - (R^a)^t A\right) W^b \left(I - A R^a\right), \tag{9}$$

which is clearly semi-definite given that W^a and W^b are semi-definite. □

In other words, the composed cycle (6) is semi-definite as long as the component relaxer pairs are themselves semi-definite. Thus, semi-definite relaxer pairs are closed under composition.

Upon closer inspection of (9), we see that W could easily be *strictly* definite within a variety of situations. For example, if W^a is strictly definite; or if W^b is strictly definite and the null spaces of W^a and $I - AR^a$ trivially intersect; then W will be strictly definite as well.

5 Examples

5.1 Trivial Relaxer

The *trivial relaxer* is given by $(S, T) = (I, 0)$, and taking $R = 0$ in (4) yields a trivial pre-relaxer and trivial post-relaxer. Regarding definiteness, clearly $W = R + R^t - R^t A R = 0$ is semi-definite.

5.2 Exact Relaxer

At the other end of the spectrum opposite the trivial relaxer is the *exact relaxer* given by $(S, T) = \left(0, A^{-1}\right)$. Taking $R = A^{-1}$ in (4) yields an exact pre-relaxer and post-relaxer. Regarding definiteness, clearly $W = R + R^t - R^t A R = A^{-1}$ is definite whenever A is definite. The exact relaxer is obviously not very interesting in isolation, but it may be used as a component to a block relaxer (below), yielding a *box relaxer*.

5.3 Splitting $A = P - Q$

As mentioned earlier, an example of a relaxer is derived from a splitting $A = P - Q$ with P invertible in which $S = P^{-1}Q$ and $T = P^{-1}$. Taking $R = P^{-1}$ gives the symmetric relaxer pair

$$
\begin{aligned}
(S_0, T_0) &= (I - RA, R) = \left(P^{-1}Q, P^{-1}\right), \\
(S_1, T_1) &= \left(I - R^t A, R^t\right) = \left(P^{-t}Q^t, P^{-t}\right).
\end{aligned}
$$

Thus, if (S_0, T_0) is derived from a splitting $A = P - Q$, (S_1, T_1) is derived from the transposed splitting $A = P^t - Q^t$. Regarding definiteness, we require

$$
W = R + R^t - R^t AR = P^{-1} + P^{-t} - P^{-t}AP^{-1} = P^{-t}\left(P + P^t - A\right)P^{-t}
$$

to be (semi-)definite, which is equivalent to $W_P := P + P^t - A$ being (semi-)definite. This latter expression is often easier to analyze.

5.3.1 Jacobi

The Jacobi method uses $P = D$ and $Q = D - A$, where D is the diagonal component of A. Thus, $W_P = P + P^t - A = 2D - A$ is definite if, for example, A is strictly or irreducibly diagonally dominant and all its diagonal entries are like-signed.

The weighted Jacobi method uses $P = \frac{1}{\omega}D$ and $Q = \frac{1}{\omega}D - A$, where $0 < \omega \leq 1$ ($\omega = 1$ reduces to the simple Jacobi method described above). Thus, $W_P = P + P^t - A = \frac{2}{\omega}D - A$ is definite if, for example, again, A is strictly or irreducibly diagonally dominant and all its diagonal entries are like-signed.

5.3.2 Gauss-Seidel

The Gauss-Seidel method uses $P = D - L$ and $Q = L^t$, where D is the diagonal component of A and L is the negative of the strictly lower triangular component of A (hence L^t is the negative of the strictly upper triangular component of A). Thus, $W_P = P + P^t - A = D$ is definite precisely when all diagonal entries of A are like-signed.

The weighted Gauss-Seidel method, also known as the SOR (Successive Over-Relaxation) method, uses $P = \frac{1}{\omega}D - L$ and $Q = \frac{1-\omega}{\omega}D + L^t$, where $0 < \omega < 2$ ($\omega = 1$ reduces to the simple Gauss-Seidel method described above). Thus, $W_P = P + P^t - A = \frac{2-\omega}{\omega}D$ is definite precisely when, again, all diagonal entries of A are like-signed.

5.3.3 SSOR

The SSOR (Symmetric Successive Over-Relaxation) method uses

$$
\begin{aligned}
P &= \frac{\omega}{2}\left(\frac{1}{\omega}D - L\right)D^{-1}\left(\frac{1}{\omega}D - L^t\right), \\
Q &= P - A,
\end{aligned}
$$

where D is the diagonal component of A; L is the negative of the strictly lower triangular component of A; and $0 < \omega \leq 1$. Thus,

$$
W_P = P + P^t - A = \frac{1-\omega}{\omega}D + \frac{1}{\omega}LD^{-1}L^t
$$

is definite precisely when, as for the Gauss-Seidel method, all diagonal entries of A are like-signed.

5.4 Kaczmarz Relaxer

The Kaczmarz method is given by

$$\begin{aligned}
x \mapsto x &+ \omega \frac{y^t (b - Ax)}{(Ay)^t (Ay)} Ay \\
&= \left(I - \frac{\omega}{y^t A^2 y} Ayy^t A \right) x + \left(\frac{\omega}{y^t A^2 y} Ayy^t \right) b \\
&= Sx + Tb
\end{aligned}$$

for $\omega > 0$ and some vector y (typically, y is chosen as one of the standard basis vectors, thus selecting out a row/column of A). Thus, taking

$$R = \frac{\omega}{y^t A^2 y} Ayy^t$$

yields a symmetric relaxer pair corresponding to the Kaczmarz method, but it is not clear if and when this is definite.

5.5 Block Relaxer

Let us consider a *block relaxer*, in which the scalar variables are partitioned and separate relaxers are applied to each partition independently. To be precise, let our solution vector $x \in \mathbf{F}^{m+n}$ and let

$$\begin{aligned}
J_m &= \begin{pmatrix} I_{mm} & 0_{mn} \end{pmatrix} \in \mathbf{F}^{m \times (m+n)}, & K_m &= J_m^t J_m = \begin{pmatrix} I_{mm} & 0_{mn} \\ 0_{nm} & 0_{nn} \end{pmatrix} \in \mathbf{F}^{(m+n) \times (m+n)}, \\
J_n &= \begin{pmatrix} 0_{nm} & I_{nn} \end{pmatrix} \in \mathbf{F}^{n \times (m+n)}, & K_n &= J_n^t J_n = \begin{pmatrix} 0_{mm} & 0_{mn} \\ 0_{nm} & I_{nn} \end{pmatrix} \in \mathbf{F}^{(m+n) \times (m+n)},
\end{aligned}$$

where I_{rc} denotes a $r \times c$ identity matrix and 0_{rc} denotes a $r \times c$ zero matrix. Now consider component relaxer pairs $\{(S_i^m, T_i^m)\}_{i=0,1}$ and $\{(S_i^n, T_i^n)\}_{i=0,1}$ (specified by R^m and R^n, respectively, via (4)) separately applied to $J_m A J_m^t$ and $J_n A J_n^t$. Taking $R = J_m^t R^m J_m + J_n^t R^n J_n$ gives the relaxer pair

$$\begin{aligned}
S_0 &= I - RA = I - \left(J_m^t R^m J_m + J_n^t R^n J_n \right) A \\
&= J_m^t \left(S_0^m J_m + T_0^m J_m A K_n \right) + J_n^t \left(S_0^n J_n + T_0^n J_n A K_m \right), \\
T_0 &= R = J_m^t R^m J_m + J_n^t R^n J_n \\
&= J_m^t T_0^m J_m + J_n^t T_0^n J_n, \\
S_1 &= I - R^t A = I - \left(J_m^t (R^m)^t J_m + J_n^t (R^n)^t J_n \right) A \\
&= J_m^t \left(S_1^m J_m + T_1^m J_m A K_n \right) + J_n^t \left(S_1^n J_n + T_1^n J_n A K_m \right), \\
T_1 &= R^t = J_m^t (R^m)^t J_m + J_n^t (R^n)^t J_n \\
&= J_m^t T_1^m J_m + J_n^t T_1^n J_n.
\end{aligned}$$

Regarding definiteness, assume that $W^z = R^z + (R^z)^t - (R^z)^t J_z A J_z^t R^z$ is semi-definite, $z = m, n$. Unfortunately,

$$\begin{aligned}
W &= R + R^t - R^t AR \\
&= J_m^t W^m J_m + J_n^t W^n J_n - \left(J_m^t (R^m)^t J_m A J_n^t R^n J_n + J_n^t (R^n)^t J_n A J_m^t R^m J_m \right).
\end{aligned}$$

will generally not also be semi-definite without some additional assumptions on R^m, R^n, and/or A. However, a common case where W *is* semi-definite is $R^n = 0$, corresponding to $\{(S_i^n, T_i^n)\}_{i=0,1}$ being trivial relaxers. In this case, $\{(S_i, T_i)\}_{i=0,1}$ effectively amount to an identity operation with respect to the last n variables of x.