June 1982

Numerical Analysis Project
Manuscript NA-82-03

GENERALIZED ITERATIVE METHODS FOR SEMIDEFINITE LINEAR SYSTEMS

Robert Schreiber

Numerical Analysis Project
Computer Science Department
Stanford University
Stanford, California 94305

GENERALIZED ITERATIVE METHODS FOR

SEMIDEFINITE LINEAR SYSTEMS

Robert Schreiber

1. Introduction

In this paper, we consider iterative solution procedures for solving singular linear systems

$$(1) \qquad Ax = b, \; b \in \text{Range } (A)$$

where A is an n by n, Hermitian, positive semidefinite (hereafter HPSD) matrix. Our aim is to consider variants of the block Jacobi, SOR, and SSOR iterations. The fundamental paper of Keller ([1965]) considers methods based on splittings

$$A = B - C$$

with B a nonsingular matrix. Here we allow B to be singular.

This paper concerns block iterative methods. We suppose that A has the k by k block structure:

$$(2) \qquad A = \begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}$$

We call the matrix $D \equiv \text{diag}\,(A_{11}, \ldots, A_{kk})$ the block-diagonal of A.

For any subspace S of $\mathbb{C}^n$, $S^{\perp}$ denotes its orthogonal complement. For any matrix A we let N(A) be its null space, R(A) its range, A* its conjugate transpose, and A^+ its generalized inverse. Recall that

$$N(A^+) = N(A^*) = R(A)^{\perp}$$

Also, recall that AA^+ is the orthogonal projection onto R(A).

We shall consider iterations of the form

$$(3) \qquad x^{n+1} = x^n + H(b-Ax^n) \qquad n = 0, 1, \ldots\ldots$$

where $N(H) \cap R(A) = \{0\}$.

Letting $T = I - HA$ we have, for any solution x of (1),

$$(x^{n+1} - x) = T(x^n - x).$$

Thus, we are concerned with the matrix $Q = \lim_{n\to\infty} T^n$.

Definition: The square matrix S is an R-matrix if

$$\text{rank } (S^2) = \text{rank } (S).$$

If S is an R-matrix, then S is nonsingular on its own range, and

$$\mathbb{C}^n = R(S) \oplus N(S).$$

Theorem 1: **[Kutznetsov [1975]]:** Q exists if and only if

(i) $\hat{\rho}(T) \equiv \sup_{\substack{\lambda\in\sigma(T)\\ \lambda\neq 1}} |\lambda| < 1$

(ii) HA is an R-matrix

where a(T) is the set of T's eigenvalues. In this case, Q is the projection onto N(A) parallel to R(HA).

When H and A satisfy the hypotheses of the theorem, we say that the method (3), or the matrix T, is convergent for A.

2. Main Results

We shall now obtain conditions on a possibly singular matrix B that guarantee convergence for A of the matrix $I - B^+A$. We then apply these results to analyze block Jacobi overrelaxation, SOR, and SSOR iteration for matrices whose diagonal blocks may be singular.

Lemma 0: Let A be HPSD. Then $(x, Ax) = 0$ if and only if $Ax = 0$.

Proof: Sufficiency is trivial. For necessity, expand x in the eigenvectors of A.

We collect here several properties of partitioned HPSD matrices. Dahlquist [1979] and Albert [1969] obtain like results.

Lemma 1: Let the HPSD matrix A be partitioned as in (2). Let D be its block-diagonal and let E = D-A. Then:

(i) $N(D) = N(A_{11}) \oplus N(A_{22}) \oplus \ldots\ldots \oplus N(A_{kk})$;

(ii) D is HPSD, as are each of itsdiagonal blocks;

(iii) if, for some $1 \leq j \leq k$, $A_{jj}x_j = 0$, then $A_{ij}x_j = 0$ for all $1 \leq i \leq k$;

(iv) if $E = L + L^*$ where L is strictly lower triangular, then

$$Dx = 0 \text{ only if } Lx = L^*x = 0; \text{ i.e.,}$$

(4) (a) $N(D) \subset N(L)$

(b) $N(D) \subset N(L^*)$

(c) $N(D) \subset N(E)$

(d) $N(D) \subset N(A)$.

(v) For each $\alpha \in [0,1)$, let A(a) be defined by

$$A(\alpha) \equiv D - \alpha E$$

Then A(a) is HPSD and

(5) $$N(A(\alpha)) \subset N(D).$$

Proof: (i) and (ii) are obvious. To prove (iii), suppose $A_{jj}z = 0$, while $A_{ij}z = c \neq 0$.

Let x(6) be partitioned conformably with (2),

$$x(\delta) \equiv (0, \ldots, 0, \underset{i^{th} \text{ position}}{c^*}, 0, \ldots, 0, \underset{j^{th} \text{ position}}{-\delta z^*}, 0, \ldots, 0)^*$$

Then for δ sufficiently large

$$x(\delta)^* A x(\delta) = c^* A_{ii} c - 2\delta c^* c < 0,$$

a contradiction. (iv) is a trivial consequence of (i) and (iii). To prove (v), assume that for some x,

$$0 > x^* A(\alpha) x = x^* D x - \alpha x^* E x.$$

By (ii),

$$\alpha x^*Ex > x^*Dx > 0,$$

and, therefore,

$$x^*Ex > x^*Dx,\ 0 > x^*Ax,$$

a contradiction. This shows that A(a) is HPSD. If A(a)x = 0, then

$$0 < x^*Dx = \alpha x^*Ex \le x^*Ex$$

with strict inequality and a contradiction unless $0 = x^*Dx = x^*Ex$. By Lemma 0, $Dx = 0$ then, so we have the inclusion (5).

Lemma 2: If A is Hermitian and B is any matrix such that

(6) $B+B^*$ is HPSD

and

(7) $(B+B^*)x = 0$ only if $Ax = 0$

then BA is an R-matrix. If in addition

(8) $(B+B^*)x = 0$ only if $Bx = 0$

then B^+A is an R-matrix.

Proof: Suppose $(Bx,x) = 0$. Then

$$0 = (Bx,\ x) = (B^*x,x) = ((B+B^*)x,x).$$

By Lemma 0, $(B+B^*)x = 0$, so by (7) $Ax = 0$. Thus $N(B) \subset N(A)$ and $N(B^*) \subset N(A)$ by the same reasoning. Hence B is nonsingular on R(A) and rank $(BA)^2$ = rank (BA) unless $ABz = 0$ for some nonzero $z \in R(A)$.. But for any such z,

$$(Bz,\ z) = 0$$

since $N(A) \perp R(A)$, so $z \in N(A) \cap R(A) = \{0\}$. This shows that BA is an R-matrix. For B^+A, B^+ is also nonsingular on R(A) since $N(B^+) = N(B^*) \subset N(A)$. Now suppose $B^+z \in N(A)$ for $z \in R(A)$. Then, letting $u = B^+z$, we have $Bu = BB^+z = z$, since

$$R(A) = N(A)^{\perp} \subset N(B^*)^{\perp} = R(B).$$

Thus,

$$0 = (u,z) = (Bu,u),$$

so that $u \in N(B+B^*)$. By hypothesis, then, $z = Bu = 0$. QED

The next lemma provides sufficient conditions for satisfaction of the first hypothesis of Kuznetsov's theorem. Our proof parallels Keller's for the case of a nonsingular matrix B (Keller [1965]).

Lemma 3: Let A be HPSD and let $T = I-B^{+}A$, 'where B is such that

$$(9) \qquad N(B+B^*) \subset N(A)$$

the matrix P defined by

$$(10) \qquad P \equiv B+B^*-A$$

is HPSD, and

$$(11) \qquad N(P) \subset N(B).$$

Then $\hat{\rho}(T) < 1$.

Proof: Using (9) we can show, as in the proof of Lemma 2, that

$$(12) \qquad N(B) \subset N(A)$$

and

$$(13) \qquad N(B^*) \subset N(A).$$

Thus, B^*, and hence B^{+} is nonsingular on $R(A)$. Thus $Tx = x$ if and only if $Ax = 0$. Now, let u be an eigenvector of T corresponding to the eigenvalue $\lambda \neq 1$. Thus,

$$(1 - \lambda)u = B^{+}Au;$$

left-multiply by B and take the inner product with u to obtain

$$\frac{(Bu,u)}{(BB^{+}Au,u)} = \frac{1}{1-\lambda}$$

Now $R(A) \subset R(B)$ since as we have seen, $R(B)' = N(B^*) \subset N(A)$; thus $BB+A = A$.

Thus, with $\lambda = \alpha + i\beta$,

$$\frac{2(1-\alpha)}{(1-\alpha)^2 + \beta^2} = 2\,\mathrm{Re}\left[\frac{1}{1-\lambda}\right] = 1 + \frac{(Pu,u)}{(Au,u)}$$

By (11) and (12), $(Pu,u) > 0$ if $u \notin N(A)$. Thus the last expression on the right is positive. The inequality obtained by dropping it yields

$$|\lambda|^2 = \alpha^2 + \beta^2 < 1. \qquad \text{QED.}$$

We now obtain necessary and sufficient conditions for convergence when B is HPSD, as is the case for Jacobi-like methods.

Theorem 2. Let B be an HPSD matrix such that $N(B) \subset N(A)$. Let $C = B-A$ and let $T = I-B^{+}A$. T is convergent for A if and only if

(i) B+C is HPSD

and

(ii) $N(B+C) \subset N(A)$.

Proof: Sufficiency follows from Theorem 1, since the hypotheses of Lemmas 2 and 3 are easily verified. For necessity, note first that since $N(B) \subset N(A)$, if $Bx = 0$, then $Cx = Bx-Ax = 0-0 = 0$, so $N(B) \subset N(C)$ also. Thus R(B) is invariant under all of A, B, and C. For (i) suppose that B+C is indefinite. An $x \in R(B)$ can be found for which

$$((B+C)x,x) < 0,$$

so that

$$(Cx,x)/(Bx,x) < -1. \tag{14}$$

Consider the generalized eigenvalue problem $Cx = \lambda Bx$ for $x \in R(B)$, a problem which makes sense since B is nonsingular on R(B) and R(B) is C-invariant. By (14), an eigenvalue $\lambda < -1$ exists. Let x be the eigenvector. Then

$$\begin{aligned} Tx &= x - B^{+}(B-C)x \\ &= B^{+}Cx \\ &= \lambda x, \end{aligned}$$

so that $\hat{\rho}(T) > 1$. For (ii), suppose $(B+C)x = 0$ while $Ax \neq 0$. Take $x \in R(B)$ by removing its orthogonal projection on N(B) if necessary---x remains nonzero since if x had no component in R(B), Ax would have been zero---the resulting x still is a null vector of B+C and Ax is not changed. Now

$$-Bx = Cx,$$

and since $B^+Bx = x$,

$$\begin{aligned} -x &= B^+Cx \\ &= x - B^+Bx + B^+Cx \\ &= Tx, \end{aligned}$$

so $\hat{\rho}(T) \geq 1$. **QED**

As an example, we consider the block-Jacobi overrelaxation (BJOR) method, based on the choice $B = \omega D$ where D is the block-diagonal of A.

Corollary: The BJOR method is convergent for A if and only if $2\omega D-A$ is HPSD and $N(2\omega D-A) \subset N(A)$.

By choosing ω sufficiently large, these conditions are necessarily satisfied.

Next, let $A = D-L-L^*$ where L is strictly lower triangular, and consider the (symmetric) block-SSOR method, defined for $\omega \neq 0$ or 2 by

$$B = \left(\frac{2-\omega}{\omega}\right)^{-1} \left(\frac{1}{\omega} D-L\right) D^+ \left(\frac{1}{\omega} D-L^*\right]$$

$$C = \left(\frac{2-\omega}{\omega}\right)^{-1} \left(\frac{1-\omega}{\omega} D+L\right) D^+ \left(\frac{1-\omega}{\omega} D+L^*\right)$$

Corollary: The block SSOR method converges for A if and only if $0 < \omega < 2$.

Proof: A straightforward computation, making use of Lemma 1(iv), shows that

$$(15) \qquad B+C = \left(\frac{2-\omega}{\omega}\right)^{-1} \left[\frac{1+(1-\omega)^2}{\omega} D - (L+L^*) + 2LD^+L^*\right],$$

from which the hypotheses of Theorem 2 can be verified. For ω outside [0,2], B+C must be negative semidefinite, as (15) shows.

We call the case $\omega = 1$ in the BJOR the block-Jacobi method. We consider the case of block Z-cyclic matrices.

Theorem 3: If A-D is block Z-cyclic, then the block-Jacobi method is convergent for A if and only if N(D) = N(A).

Proof: Every eigenpair of T is an eigenpair of (E,D), for if

$$Tu = \lambda u$$

then since $DD^{+}E = E$,

$$Tu = u - D^{+}(D-E)u = \lambda u,$$

so that

$$Eu = \lambda Du.$$

If $u \in N(A)$ then $\lambda = 1$; otherwise $Du \neq 0$ and $0 < (Au,u) = (Du,u) - (Eu,u) = (1-\lambda)(Du,u)$ so that $\lambda < 1$. Since E is Z-cyclic, $-\lambda$ is also an eigenvalue, so $\lambda > -1$ (see Varga [1962].) If N(A) = N(D), we have convergence. But if N(A) - N(D) is nonempty, we have

$$Eu = Du$$

for some $u \in R(D)$, so 1 and -1 are both eigenvalues.

QED

We now consider the block-SOR splitting

$$B = \omega^{-1}D - L$$

$$C = \omega^{-1}(1-\omega)D + L^{*}$$

Theorem 4: Block-SOR is convergent for an HPSD matrix A if and only if $0 < \omega < 2$.

Proof: Let $0 < \omega < 2$. According to Lemma 1(v), since

$$B + B^{*} = 2\omega^{-1}A(\omega/2)$$

we have that $B+B^{*}$ is HPSD and $N(B+B^{*}) \subset N(D) \subset N(A)$. Moreover, by Lemma 1(iv), $N(D) \subset N(L)$, so $N(B+B^{*}) \subset N(B)$. The matrix P of (10),

$$P = B^{*} + C = \omega^{-1}(2-\omega)D,$$

is HPSD and its null space is contained in N(B), as shown above. Thus Lemmas 2 and 3, and hence Theorem 1, apply.

Our proof that convergence requires $0 < \omega < 2$ mimics the proof of Lemma 3. First we dispose of the case $\omega = 0$. Actually for $\omega = 0$, our definition of the method is nonsense. But the "blockwise" definition.

$$\tilde{x}_j^n = A_{jj}^+\left(b_j - \sum_{k<j} A_{jk}x_k^{n+1} - \sum_{k>j} A_{jk}x^n\right)$$

$$x_j^{n+1} = x_j^n + \omega(\tilde{x}_j^n - x_j^n)$$

makes perfect sense. In fact, for $\omega = 0$, $T = 1$ and $B = 0$. T is convergent for A if and only if **A is** the zero matrix. For ω outside of $[0,2)$, we shall show that $\hat{\rho}(T) \geq 1$ unless $A \equiv 0$. First we show that $N(B^+) \subset N(A)$. Let $B^+x = 0$. Then $B^*x = 0$. B^* is block-upper triangular and its diagonal blocks are **nonzero** multiples of those of A. Partition x as $(x_1,\ldots,x_k)^*$ conformably with A. Then $A_{kk}x_k = 0$. **By Lemma 1 (iii)**, $A_{ik}x_k = 0$ for $1 \leq i \leq k-1$; these are the blocks in the kth block-column of B^*. Hence, $0 = B^*x = B^*(x_1,\ldots,x_{k-1},0)^*$. We can repeat this argument to show, eventually, that $Dx = Ax = 0$, as required.

We now proceed as in the proof of Lemma 3 to show that if $Tx = \lambda x$ and $\lambda \neq 1$ then

$$2\mathrm{Re}\left[\frac{1}{1-\lambda}\right] = 1 + \frac{(Px,x)}{(Ax,x)}$$

P is a negative scalar multiple of D if $\omega \notin [0,2]$ and is zero for $\omega = 2$. In the former case, since $x \notin N(A)$, $(Px,x) < 0$ and this implies that $\hat{\rho}(T) > |\lambda| > 1$. In the later, we have $\hat{\rho}(T) = |\lambda| = 1$.

QED

Concerning necessary and sufficient conditions for a general splitting $A = B-C$, we have only partial results. Sufficient conditions are provided by Lemmas 2 and 3. When all conditions except (10) are satisfied, we have that if B^*+C is negative semidefinite then T is not convergent for A unless $A \equiv 0$---this was shown in the preceding proof. When B^*+C is indefinite, we cannot say. For example, when

$$A = A(\alpha) = \begin{bmatrix} 1 & 1 & \alpha \\ 1 & 1 & \alpha \\ \alpha & \alpha & 2 \end{bmatrix}$$

and $B = D = \mathrm{diag}\,(1,1,2)$, then for $|\alpha| \leq \sqrt{2}$ A is HPSD (its nonzero eigenvalues are $2 \pm \sqrt{2}\,\alpha$); unless $\alpha = 0$, B^*+C is indefinite: since its trace is 4 and its determinant is $-4\alpha^2 < 0$, it has exactly one negative and two positive eigenvalues. Finally, $T(\alpha) = I - D^{-1}A(\alpha)$ has the eigenvalues $\{1, \left(\pm(1-4\alpha^2)^{\frac{1}{2}} - 1\right)/2\}$ so that

$$\hat{\rho}\,(T(\alpha)) \begin{cases} < 1 \text{ for } |\alpha| < 1 \\ = 1 \text{ for } \alpha = 1 \\ > 1 \text{ for } |\alpha| > 1 \end{cases}$$

References

A. Albert [1975].
Conditions for positive and nonnegative definiteness in terms of pseudoinverses.
SIAM J. Appl. Math., 17, 434-440.

Germund Dahlquist [1979].
Some contractivity questions for one-leg and linear multistep methods.
Report TRITA-NA-7905, Royal Institute of Technology, Stockholm, Sweden.

Herbert B. Keller [1965].
On the solution of singular and semidefinite linear systems by iteration.
SIAM J. Numer. Anal., 2, 281-290.

Y. A. Kuznetsov [1975].
Iterative methods for solution of non-compatible systems of linear equations.
In R. Glowinski and J. L. Lions, eds., Computer Methods in Applied Sciences and Engineering, 41-55, Springer-Verlag.

Richard S. Varga [1962].
Matrix Iterative Analysis.
Prentice-Hall.