JOURNAL OF APPROXIMATION THEORY 17, 44–56 (1976)

Graphical Techniques and 3-Part Splittings for Linear Systems

J. DE PILLIS

Istituto Matematico, Florence, Italy, and Department of Mathematics, University of California, Riverside, California 92507

Communicated by Oved Shisha

Received July 29, 1974

1. BACKGROUND

With A an invertible bounded linear operator on Hilbert spaze $\mathscr{H}$ (an invertible matrix on n-dimensional spaze $\mathscr{H}_n$) and y_0 fixed in $\mathscr{H}$, we seek the solution vector $x \in \mathscr{H}$ for the linear system

$$Ax = y_0 . \tag{1.1}$$

If A^{-1} is not immediately accessible, we can, at least, extract an invertible term A_1, and from the 2-part splitting, $A = A_1 + A_2'$, we define the so-called 2-part sequence $\{x_n'\}$ recursively by

$$A_1 x_{n+1}' + A_2' x_n' = y_0 , \qquad n = 0, 1, 2,..., \tag{1.2}$$

for arbitrary but fixed initial $x_0' \in \mathscr{H}$. Similarly, we define the 3-part sequence $\{x_n\}$, resulting from the 3-part splitting $A = A_1 + A_2 + A_3$, by the equations

$$A_1 x_{n+2} + A_2 x_{n+1} + A_3 x_n = y_0 , \qquad n = 0, 1, 2,..., \tag{1.3}$$

for arbitrary, but fixed initial couple x_0, $x_1 \in \mathscr{H}$. (Note that the first term, A_1, of the splittings in (1.2) and (1.3) are the same, so that of necessity, we have $A_2' = A_2 + A_3$.) Clearly, if the sequences $\{x_n'\}$ of (1.2), or $\{x_n\}$ of (1.3) converge at all, then the convergence must be to the solution vector x of (1.1). In a recent paper [1] it is shown that for certain complex analytic $\phi(\cdot)$ defined in $\sigma(A_1^{-1}A_2')$, the spectrum of $A_1^{-1}A_2' = B$, we may choose A_3 in (1.3) of the form

$$A_3 = A_1 \phi(B)(I + \phi(B))^{-1}(B - \phi(B)), \tag{1.4}$$

where $\phi(\cdot)$ is the corresponding analytic function acting on the operator (matrix) $\mathscr{B}$. Conditions are given [1, Theorem 6.3] so that for certain constraints on $\sigma(A_1^{-1}A_2')$, the 3-part splitting of (1.3) provides faster convergence than the 2-part splitting of (1.2). Of course we must raise the questions: (a) How much faster (and by what measure) does $\{x_n\}$ converge, and (b) is it worth the extra effort to compute A_3 of (1.4), i.e., is computation of $(I + \phi(A_1^{-1}A_2'))^{-1}$ reasonably easy?

2. The Objectives of this Paper

In deciding whether passage from a given 2-part splitting, $A = A_1 + A_2'$, to a 3-part splitting, is feasible, we answer the following questions:

(1) How far outside the unit circle can $\sigma(A^{-1}A_2')$ lie in order that a 3-part splitting (1.3) will produce a convergent sequence $\{x_n\}$ for all initial x_0?

(2) Is there some graphical (ruler-and-compass type) construction which, relative to the elements of $\sigma(A_1^{-1}A_2')$, gives us a way of finding an analytic $\phi(\cdot)$ so as to construct A_3 in our 3-part splitting (cf. (1.4))?

(3) What conditions will allow the simplest possible case (viz $\phi(\cdot) =$ constant) for construction of $\phi(\cdot)$ (hence, of A_2) in (1.4)?

(4) How much faster will the 3-part sequence $\{x_n\}$ converge, relative to the 3-part sequence $\{x_n'\}$?

Question (1) is answered in Theorem 3.1, although a sketch of the proof appears in [1, cf. (6.12)], in which we see that $\sigma(A_1^{-1}A_2')$ may not lie anywhere outside the cardiod $\mathscr{C} = \{z: 2z[\mathrm{Re}(z) + 1] - 1, |z| = 1\}$, for any $\phi(\cdot)$ of (1.4) resulting in a *convergent* sequence $\{x_n\}$ of (1.3).

Question (2) is answered in Theorem 3.2, in which a graphical algorithm is presented for a construction of analytic $\phi(\cdot)$ for (1.4) in the following sense: From an individual element λ in $\sigma(A_1^{-1}A_2')$, we construct the value $\phi(\lambda)$.

Question (3) (which asks when $\phi(\cdot)$ might be constant) is addressed by Theorem 3.3 for the case $\sigma(A_1^{-1}A_2')$ is real. For example, we show that if $\sigma(A_1^{-1}A_2') \subset [-s^2, s^2 + 2s]$ for some s, $0 < s < 1$, where $\rho(A_1^{-1}A_2') = s^2 + 2s$ (this includes $A_1^{-1}A_2'$ positive semidefinite), then the constant analytic $\phi(A_1^{-1}A_2') = sI$ yields a 3-part sequence $\{x_m\}$ whose average reduction factor (definitions follow) is eventually about $1/(s + 2)$ times the average reduction factor of the 2-part sequence $\{x_m'\}$. In other words, if $R(x_m')$, the rate of convergence of $\{x_m'\}$, is defined as $-\ln \rho(A_1^{-1}A_2')$, then $R(x_m)$ will be $R(x_m') + \ln(2 + s)$. An interesting consequence of this will be that if $A_1^{-1}A_2$ is positive semidefinite, with maximal eigenvalue λ_0, where $\lambda_0^{k-1} > \frac{1}{2}$ for $k > 0$, then the prescribed 3-part splitting will always increase the convergence rate by a factor of at least k.

3. The Main Results

Given any iteratively defined sequence $\{x_0, x_1, \ldots, x_m, \ldots\}$ converging to the solution vector x for the linear system $Ax = y_0$, we measure its speed of convergence by $\sigma(m)$, its *average reduction factor* (after m iterations):

$$\sigma(m) = (\| x_m - x \| / \| x_0 - x \|)^{1/m} \tag{3.1}$$

(cf. [2, p. 62]).

We distinguish the average reduction factor of the 2-part (primed) sequence $\{x_0, x_1', x_2', \ldots, x_m', \ldots\}$ of (1.2), and the 3-part (unprimed) sequence $\{x_0, x_1, x_2, \ldots, x_m, \ldots\}$ of (1.3) by the symbols $\sigma'(m)$ and $\sigma(m)$, respectively. The comparison of $\sigma'(m)$ with $\sigma(m)$ will concern us. We know that the spectral radius $\rho(A_1^{-1}A_2')$ is an "eventual" upper bound for $\sigma'(m)$ [2, p. 62], where $A = A_1 + A_2'$. By eventual upper bound, we mean that $\sigma'(m)$ is actually shown to be bounded above by scalars a_m, say, and these scalars a_m eventually converge (downward) to $\rho(A_1^{-1}A_2')$ for m sufficiently large. Henceforth, we shall indicate this by the symbol $\sigma(m) \approx \rho(A_1^{-1}A_2')$. Now in [1, Theorem 6.3], the following comparison is established: For the 2-part sequence $\{x_m'\}$ defined by (1.2), and the 3-part sequence $\{x_m\}$ defined by an analytic function $\phi(\cdot)$ on $\sigma(A_1^{-1}A_2')$ in (1.3), where $x_0' = x_0 = x_1$, we have (in the eventual sense mentioned above)

$$\sigma'(m) \approx \rho(A_1^{-1}A_2') \tag{3.2}$$

[2, p. 62], while

$$\begin{aligned} \sigma(m) &\approx \max\{| \phi(z)|, |(z - \phi(z))/(1 + \phi(z))| : z \in \sigma(A_1^{-1}A_2')\}. \\ &= r. \end{aligned} \tag{3.3}$$

It is also shown that convergence of 3-part sequences is assured when r, the right-hand side of (3.3), is less than unity. This allows us to consider situations where a 2-part sequence $\{x_m'\}$ diverges ($\rho(A_1^{-1}A_2') > 1$), yet an analytic $\phi(\cdot)$ on $\sigma(A_1^{-1}A_2')$ can be found so that r, the right-hand side of (3.3) is less than one, i.e., so that the 3-part sequence $\{x_m\}$ converges. In any case, for such a $\phi(\cdot)$ to be found, $\sigma(A_1^{-1}A_2')$ must lie within a certain cardioid, described in the following theorem.

THEOREM 3.1. *Consider any complex function ϕ: $\mathbf{C} \to \mathbf{C}$ with the properties*

$$\begin{aligned} &\text{(i)} \quad | \phi(u)| < 1 \\ &\text{(ii)} \quad |(u - \phi(u))/(1 + \phi(u))| < 1, \qquad \textit{where} \qquad \phi(u) \neq -1. \end{aligned} \tag{3.1}$$

Then necessarily, the domain of ϕ lies in the interior of the cardioid

$$\mathscr{C} = \{2z[\mathrm{Re}(z) + 1] - 1 : z = e^{i\theta}\}.$$

Proof. Let u be an arbitrary point of the complex plane $\mathbf{C}$, and let a be the midpoint between u and -1. Now let L be the line through a perpendicular to the line through -1 and u (see Fig. 1.)

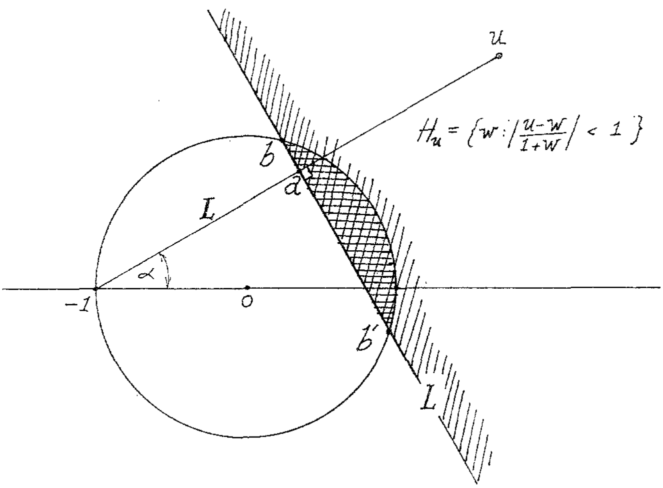

FIGURE 1.

Now observe (cf. Fig. 1) that the open half-plane, H_u, defined by line L, containing u, is the set of all complex w which are at least as close to u as they are to -1. That is,

$$H_u = \{w : |(u - w)/(1 + w)| < 1\}.$$

Thus, the allowable values of $\phi(u)$ subject to conditions (3.1(i)) and (3.1(ii)), must belong to both the open unit disc and to H_u. But this constraint (requiring that $H_u \cap$ unit disc $\neq \varnothing$) tells us something about u. As per Fig. 1, on any line segment L, the furthest that u may place itself from -1 is only to that point which forces the points of intersection, b and b', to coincide on the rim of the unit circle. This limiting position is illustrated in Fig. 2.

Observe that the distance between -1 and $b = b'$ is $2\cos(\alpha/2)$, from which it follows that for angle α, the distance $|1 + u|$ from -1 to u such that equality obtains for both (3.1)(i) and (3.1)(ii), is $4\cos^2(\alpha/2)$. In polar coordinates, then, equality for (3.1)(i) and (3.1)(ii) prevails only for those $u(\alpha)$ of the form $u(\alpha) = 4\cos^2(\alpha/2) - 1$, or $u(\alpha) = 2(\cos\alpha + 1) - 1$. In complex

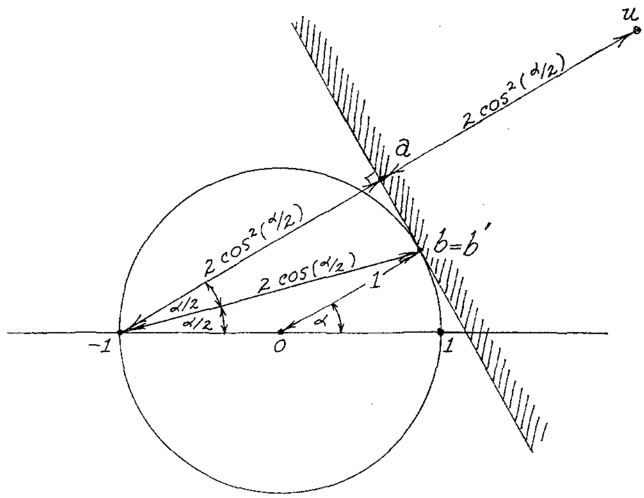

FIGURE 2.

form, then, $u(\alpha)$ is the cardioid $\mathscr{C}$ of complex $w = 2z[\mathrm{Re}(z) + 1] - 1$, as z runs over all unit vectors $e^{i\alpha}$. Finally, then, any u for which both (3.1)(i) and (3.1)(ii) obtain, must lie in the *interior* of the cardioid $\mathscr{C}$. Moreover, the image point, $\phi(u)$, must lie in the intersection of the open unit disc and the open half plane H_u shown in Fig. 1. This ends the proof.

NOTATION. In what follows, the symbols $D(a, b)$ and $C(a, b)$ will denote, respectively, the closed disc and the circle in the complex plane, each with center a, and radius $b \geqslant 0$.

Consider $\lambda \in \sigma(A_1^{-1}A_2')$. We now persent a graphical algorithm for constructing candidates for $\phi(\lambda)$, (for $\phi(\cdot)$ required in (1.4).) Moreover, the construction will indicate (for that particular λ) a value r equal to the right-hand side of (3.3), thereby giving us an upper bound on the average reduction factor $\sigma(m)$ for the 3-part sequence $\{x_m\}$.

THEOREM 3.2. *Suppose* $\lambda \in \sigma(A_1^{-1}A_2')$, *where* $A = A_1 + A_2'$. *Then if the value* $\phi(\lambda)$ *exists such that*

$$\begin{aligned} &\text{(i)} \quad |\phi(\lambda)| \leqslant r, \qquad \textit{and} \\ &\text{(ii)} \quad |(\lambda - \phi(\lambda))/(1 + \phi(\lambda))| \leqslant r, \end{aligned} \tag{3.4}$$

then it is necessary and sufficient that $\phi(\lambda)$ *lie in the shaded region of Fig.* 3. *Given* $\lambda \in \sigma(A_1^{-1}A_2')$, *the key reference points* λ' *and* d *(defining the disc* $D(\lambda', | d - \lambda' |)$ *of those complex* z *for which* $| \lambda - z |/| 1 + z | \leqslant r$*) are constructed by the following five-step algorithm*:

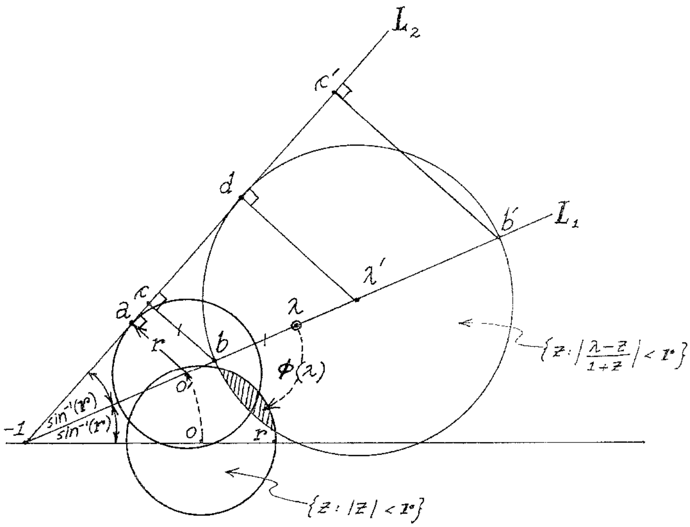

FIGURE 3.

(1) *Draw line L_1 through the points -1 and λ.*

(2) *On line L_1, place point $0'$, one unit from -1, and construct the circle $C(0', r)$.*

(3) *Draw tangent line L_2 through -1 and tangent to $C(0', r)$ at point a.*

(4) *Locate b and b' on line L_1 so that they are equidistant from line L_2 and from λ, i.e., so that $| b - c | = | b - \lambda |$ and $| b' - c' | = | b' - \lambda |$.*

(5) *Construct λ' and d to be the midpoints between b and b', and between c and c', respectively.*

Proof. We now justify the algorithm. We note that $C(\lambda', | \lambda' - d |)$ is a so-called circle of Apollonius, that is, the locus of points z whose distances from two points -1 and λ have the same ratio. In fact we see that for b and $b' \in C(\lambda' | \lambda', d |)$.

$$\frac{| b - \lambda |}{| b - (-1)|} = \frac{| b - c |}{| b - (-1)|} \qquad \text{(step (4))},$$
$$= r$$

(see triangle $(-1, b, c)$ of Fig. 3), and

$$\frac{| b' - \lambda |}{| b - (-1)|} = \frac{| b' - c' |}{| b - (-1)|} \qquad \text{(step (4))},$$
$$= r$$

(see triangle $(-1, b', c')$ of Fig. 3).

Thus, $D(\lambda' \mid \lambda' - d \mid)$, the disc defined by the circle of Appolonius described above, is the set of all z such that

$$\frac{|z - \lambda|}{|z - (-1)|} = \frac{|\lambda - z|}{|1 + z|} \leqslant r.$$

What this says is that if $\phi(\lambda)$ is to satisfy (3.4(ii)), $\phi(\lambda)$ must be one of the z's of $D(\lambda', |\lambda' - d|)$. (We remark that $|\lambda' - d| = |\lambda' - b| = |\lambda' - b'|$ in Fig. 3 since the polygon $bcc'b'$ is a trapezoid.) On the other hand, if $\phi(\lambda)$ is to satisfy (3.4(i)), $\phi(\lambda)$ must belong to $D(0, r)$, as well. In a word, if $\phi(\lambda)$ satisfies both (3.4(i)) and (3.4(ii)), then necessarily, $\phi(\lambda) \in D(0, r) \cap D(\lambda', |\lambda' - d|)$, which justifies the five-step construction algorithm for λ' and d of Fig. 3.

Remark. The crucial intersection of Fig. 3, describing the range of $\phi(\cdot)$, is determined by the discs $D(0, r)$ and $D(\lambda', |\lambda' - d|)$, each with *common* radius r. We might have proceeded more generally by constructing $D(\lambda', |\lambda' - d|)$ with radius r, and then constructing $D(0, s)$, large or small enough to provide nonempty intersection $D(0, s) \cap D(\lambda', |\lambda' - d|)$. In this case, we would have, that eventually,

$$\sigma(m) \approx \max\{r, s\}.$$

Our previous construction provides an r and a $\phi(\lambda)$ satisfying the inequalities of (3.4) for a *single* λ in $\sigma(A_1^{-1}A_2')$. We can ask how the domain of $\phi(\cdot)$ can be extended beyond the singleton $\{\lambda\}$. It is easy to describe a constant set for $\phi(\cdot)$, i.e., those complex z, for which we assign $\phi(z) = \phi(\lambda)$, where, for the r constructed relative to λ, the inequalities (3.4) still obtain. In fact, inspection of (3.4)(ii) yields the following immediately:

COROLLARY 3.1. *Given* $r > 0$ *and* $\phi(\lambda)$ *satisfying* (3.4), *the set of complex* z, *with* $\phi(z) = \phi(\lambda)$, *satisfying the inequalities*

$$\begin{aligned} &\text{(i)} \quad |\phi(z)| \leqslant r, \qquad \textit{and} \\ &\text{(ii)} \quad |z - \phi(z)|/|1 + \phi(z)| \leqslant r \end{aligned} \tag{3.5}$$

is the disc

$$D(\phi(\lambda), r|1 + \phi(\lambda)|) = \{z\colon |z - \phi(\lambda)| \leqslant r|1 + \phi(\lambda)|\}. \tag{3.6}$$

The Real Case

Let us concentrate on the case when $\sigma(A_1^{-1}A_2')$ is real, i.e., assume we have an estimate of real end points p and P such that $\sigma(A_1^{-1}A_2') \subset [p, P]$. As we shall see, all of $\sigma(A_1^{-1}A_2')$ can be realized as a constant set for some $\phi(\cdot)$. We can now answer the questions:

(a) What value shall we take as the upper bound r for $\sigma(m)$ = the average reduction factor for the 3-part sequence $\{x_m\}$ (cf. (3.3))?

(b) What value shall we then assign for the constant function $\phi(\cdot)$ in (3.6) so that *all* of $\sigma(A_1^{-1}A_2')$ is in the domain of $\phi(\cdot)$?

(c) What is the consequent imporvement of the bound r for $\sigma(m)$ relative to that (vis, $\rho(A_1^{-1}A_2')$) for $\sigma'(m)$ (vf. (3.2) and (3.3))? That is, what is $r/\rho(A_1^{-1}A_2')$?

The answers are contained in the next theorem.

THEOREM 3.3. *Let $\sigma(A_1^{-1}A_2')$ be a subset of the real line. Assume either*

(A) *For some r_1, $0 \leqslant r_1 < 1$,*

$$\sigma(A_1^{-1}A_2') \subset [-r_1^2, r_1^2 + 2r_1],$$

or

(B) *For some r_2, $-1 < r_2 \leqslant 0$,*

$$\sigma(A_1^{-1}A_2') \subset [r_2^2 + 2r_2\,, -r_2^2],$$

where for case (A) *and case* (B) *the spectral radius of $A_1^{-1}A''$ coincides with the appropriate interval end point, i.e., $\rho(A_1^{-1}A_2') = |\, r_i^2 + 2r_i \,|$, $i = 1, 2$.* (This defines *the r_i to be selected.) Then in each case, we may define $\phi(z) = r_i$ for all $z \in \sigma(A_1^{-1}A_2')$, resulting in a 3-part sequence $\{x_n\}$ of* (1.3) *whose average reduction factor $\sigma(m)$ of* (3.3) *is eventually bounded by $|\, r_i \,|$, $i = 1, 2$. Moreover, the ratio of improvement $|\, r_i \,|/\rho(A_1^{-1}A'')$ of the bounds of* (3.2) *and* (3.3) *is always equal to $1/(2 + r_i)$. Equivalently, if the rates of convergence for $\{x_m\}$ and $\{x_m'\}$ are $R(x_m) = -\ln |\, r_i \,|$, and $-\ln \rho(A_1^{-1}A_2')$, respectively,* (*note*: *$R(x_m)$ and $R(x_m')$ are independent of m), then*

$$R(x_m) = R(x_m') + \ln(2 + r_i).$$

Proof. To satisfy (3.5)(i), we assign the value $\phi(z) = r$, where $|\, r \,| < 1$. (We shall see presently, how r must relate to the scalar $\rho(A_1^{-1}A_2')$.) To satisfy (3.5)(ii), or (3.6), all z for which $\phi(z) = r$, must satisfy the inequality

$$|\, z - r \,| \leqslant r + r^2. \tag{3.7}$$

But if $z \geqslant r$, (3.7) implies

$$z \leqslant r^2 + 2r, \tag{3.8}$$

and if $z \leqslant r$, (3.7) implies

$$z \geqslant -r^2. \tag{3.9}$$

Thus, those real z for which (3.6) obtains must lie in the interval bounded by $r^2 + 2r$, and by $-r^2$. But $-r^2 < r^2 + 2r$ if and only if $r > 0$, so that either

(A) $\sigma(A_1^{-1}A_2') \subset [-r_1^2, r_1^2 + 2r_1]$ if

$$\phi\colon \sigma(A_1^{-1}A_2') \to r_1 > 0,$$

or

(B) $\sigma(A_1^{-1}A_2') \subset [r_2^2 + 2r_2, -r_2^2]$ if

$$\phi\colon \sigma(A_1^{-1}A_2') \to r_2 < 0.$$

Since we have assumed that for case (A), $\rho(A_1^{-1}A_2') = r_1^2 + 2r_1$, and for case (B), $\rho(A_1^{-1}A_2') = -r_2^2 - 2r_2$, it is easy to check that in both cases, the ratio of improvement $| r_i |/\rho(A_1^{-1}A_2')$ is equal to

$$\frac{| r_i |}{\rho(A_1^{-1}A_2')} = \frac{1}{2 + r_i} \qquad i = 1, 2.$$

Defining rates of convergence as the negative of the log of the "essential" upper bounds $| r_i |$ for $\sigma(m)$, and $\rho(A_1^{-1}A_2')$ for $\sigma'(m)$, i.e., $R(x_m) = -\ln | r_i |$ and $R(x_m') = -\ln \rho(A_1^{-1}A_2')$, leads us to the equation $R(x_m) = R(x_m') + \ln(2 + r_i)$. This ends the proof.

Remark. The above theorem provides us with a specific algorithm for finding that constant r for which $\phi\colon \sigma(A_1^{-1}A_2') \to r$, thus allowing construction of $A_3 = (r/1 + r)(-rA_1 + A_2')$ as per (1.4) (take $\phi(A_1^{-1}A_2') = rI$). Moreover, if we know that the largest eigenvalue in $\sigma(A_1^{-1}A_2')$ is near $r^2 + 2r$, then this allows us to solve for r and to estimate that $\sigma(m)$ will be about $1/(2 + r)$ times $\sigma'(m)$, at least for all m sufficiently large, i.e., $R(x_m) = R(x_m') + \ln(2 + r)$.

Remark. Note that case (A) includes all $A_1^{-1}A_2'$ positive semidefinite, with $\rho(A_1^{-1}A_2') = \lambda_0$, say. In this case we choose *nonnegative* $r = -1 + (1 + \lambda_0)^{1/2}$, where, in the construction of A_3 (1.4) for the 3-part splitting (1.3), we take $(A\phi_1^{-1}A_2') = rI$. This means that if for $k > 0$, $\lambda_0^{k-1} > \frac{1}{2}$, the 3-part splitting will always increase the rate of convergence by a factor of at least k. This example indicates a general property of 3-part splittings, viz, the worse the situation is (meaning, the slower the convergence) for the 2-part splitting $A = A_1 + A_2'$, the more effective is the passage to the 3-part splitting $A = A_1 + A_2 + A_3$, for increasing the rate of convergence.

4. Numerical Examples

We tabulate three examples for 6×6 matrices C_i, $i = 1, 2, 3$, with real spectrum.

TABLE I

	C_1	C_2	C_3
a_{11}	8.85680975	8.85	8.99
a_{12}	−15.9136195	−15.9	−15.78
a_{13}	41.68404875	41.65	41.55
a_{14}	−263.9155023	−264.29698	−257.11
a_{15}	69.606678	69.798792	66.7
a_{16}	−34.727239	−34.799396	−33.77
a_{21}	0	0	0
a_{22}	2.9	2.9	3.1
a_{23}	−2.8	−2.8	−3.2
a_{24}	23.4478	23.4	27.24
a_{25}	−7.0239	−7.0	−8.32
a_{26}	2.8	2.8	3.2
a_{31}	2.0	2.0	2.0
a_{32}	2.0	2.0	2.0
a_{33}	5.0	5.0	5.0
a_{34}	−4.0	−3.50302	−7.65
a_{35}	−2.8	−2.998792	−1.42
a_{36}	−1.1	−1.000604	−1.79
a_{41}	2.0	2.0	2.0
a_{42}	−2.0	−2.0	−2.0
a_{43}	8.0	8.0	8.0
a_{44}	−42.75	−42.75	−42.85
a_{45}	10.0	10.0	10.0
a_{46}	−6.0	−6.0	−6.0
a_{51}	5.0	5.0	5.0
a_{52}	−6.0	−6.0	−6.0
a_{53}	22.0	22.0	22.0
a_{54}	−125.4522	−125.5	−125.06
a_{55}	31.4761	31.5	31.18
a_{56}	−17.0	−17.0	−17.0
a_{61}	−1.0	−1.0	−1.0
a_{62}	1.0	1.0	1.0
a_{63}	−1.0	−1.0	−1.0
a_{64}	4.3456	4.74698	2.18
a_{65}	1.1522	1.001208	1.94
a_{66}	1.9	1.999396	1.21

TABLE II[a]

$$C_1 = I_6 + D_1 \qquad x_0 = (8, 4, -5, 4, 2, 0)$$
$$\sigma(D_1) = \{-0.14319025, -0.1, 0.0, 0.25, 0.4761, 0.9\}$$
$$\rho(D_1) = 0.9, \qquad r = 0.378404875, \qquad \phi(D_1) = rI_6$$

n	$\|x_n\|/\|x_0\|$	$\|x_n'\|/\|x_0\|$	$\sigma(n)$	$\sigma'(n)$
1	1.000 000	113.108 062	1.000	113.1
2	113.108 062	43.667 486	10.635	6.608
3	46.579 095	77.615 677	3.598	4.265
4	62.763 807	88.554 503	2.814	3.067
5	40.049 995	96.540 981	2.091	2.494
10	1.414 364	80.100 428	1.035	1.550
15	0.018 584	48.008 148	0.766	1.294
20	0.000 201	28.365 837	0.653	1.182
25	0.000 001	16.750 171	0.587	1.119
26	0.—	15.075 162	0.580	1.109
35	0.—	5.840 430	0.520	1.051
100	0.—	0.006 197	0.425	0.950
129	0.—	0.000 291	0.414	0.938

[a] $R(x_n)/R(x_n') = \ln(r)/\ln \rho(D_1) = 9.22$; $\sigma(129)/\sigma'(129) = 2.27$.

TABLE III[a]

$$C_2 = I_6 + D_2 \qquad x_0 = (8, 4, -5, 4, 2, 0)$$
$$\sigma(D_2) = \{-0.15, -0.1, 0.0, 0.25, 0.5, 0.999396\}$$
$$\rho(D_2) = 0.999396, \qquad r = 0.414, \qquad \phi(D_2) = rI_6$$

n	$\|x_n\|/\|x_0\|$	$\|x_n'\|/\|x_0\|$	$\sigma(n)$	$\sigma'(n)$
1	1.000 000	113.189	1.000	113.2
2	113.189 885	43.443	10.639	6.591
3	45.421 612	77.061	3.567	4.255
4	63.573 975	88.846	2.823	3.070
5	40.307 379	102.826	2.094	2.525
10	1.842 889	136.218	1.063	1.634
15	0.032 304	137.141	0.795	1.388
20	0.000 463	136.770	0.681	1.279
25	0.000 005	136.358	0.615	1.217
26	0.000 002	136.276	0.607	1.208
27	0.000 000	136.194	0.597	1.199
50	0.—	134.314	0.506	1.102
90	0.—	131.107	0.466	1.056
135	0.—	127.591	0.449	1.037

[a] $R(x_n)/R(x_n') = \ln(r)/\ln(\rho(D_2)) = 1{,}461$; $\sigma(135)/\sigma'(135) = 2.31$.

TABLE IV[a]

$C_3 = I_6 + D_3 \qquad x_0 = (8, 4, -5, 4, 2, 0)$
$\sigma(D_3) = \{-0.1, 0.0, 0.1, 0.15, 0.18, 0.21\}$
$\rho(D_3) = 0.21, \qquad r = 0.1, \qquad \phi(D_3) = rI_6$

n	$\lVert x_n \rVert / \lVert x_0 \rVert$	$\lVert x_n' \rVert / \lVert x_0 \rVert$	$\sigma(n)$	$\sigma'(n)$
1	1.000 000	111.326 583	1.000	111.326
2	111.326 683	44.873 922	10.551	6.699
3	47.834 845	81.996 116	3.630	4.344
4	72.558 852	96.555 184	2.919	3.135
5	65.427 057	70.959 400	2.308	2.345
6	30.756 910	34.021 879	1.770	1.800
7	4.461 669	12.456 920	1.238	1.434
8	0.928 876	3.888 346	0.990	1.185
9	0.111 368	1.096 571	0.783	1.010
10	0.017 483	0.288 688	0.667	0.883
14	0.000 003	0.000 949	0.409	0.608
15	0.000 000	0.000 215	0.373	0.569
18	0.—	0.000 002	0.306	0.486
19	0.—	0.000 000	0.289	0.466

[a] $R(x_n)/R(x_n) = \ln(r)/\ln(\rho(D_3)) = 1.48$; $\sigma(19)/\sigma'(19) = 1.61$.

The 2-part splittings all take $A_1 = I_6$, the identity matrix. That is

$$C_i = I_6 + D_i$$

(so that $A_1^{-1}A_2' = D_i$) defines the two part sequence $\{x_m'\}$ as per (1.2) and the 3-part splitting

$$C_i = I_6 + \left(\frac{1-r}{1+r} D_i + \frac{r^2}{1+r} I_6\right) + \left(\frac{r}{1+r} D_i - \frac{r^2}{1+r} I_6\right)$$

defines the 3-part sequence $\{x_m\}$ as per (1.3), with A_3 defined by $\phi(D_i) = r$ in (1.4). The C_i's are selected so that $\sigma(C_i)$ is real, and Theorem 3.3 will apply. In our examples r will be taken as the positive root of $r^2 + 2r - (\lambda_i - 1)$, where λ_i (resp. $\lambda_i - 1$) is the largest eigenvalue of C_i (resp. of $D_i = C_i - I_6$). Finally, we test the systems $C_i x = 0$ for convergence of the sequences $\{x_m'\}$ and $\{x_m\}$ to the solution vector 0, with $x_0 = \text{col}(8, 4, -5, 4, 2, 0)$.

The entries a_{jk} for each C_i are tabulated in Table I.

In C_2, we perturb the eigenvalues to bring $\rho(D_2)$ even closer to the unit circle. Convergence for the 2-part splitting $C_2 = I_6 + D_2$ is very much slower than that for C_1 above, but the 3-part splitting converges for C_2, about as fast as it does for C_1, reaching 6-place accuracy, for example, in 27 iterations.

We have seen two cases, C_1, C_2, where a 3-part splitting works best, i.e., when $\rho(D_i)$ is close to unity and the 2-part sequence $\{x_n'\}$ converges slowly. In the next case, $\rho(D_3)$ is reasonably small ($\rho(D_3) = 0.21$) and while the improvement by a 3-part splitting on C_3 is not as dramatically better, a faster convergence does result.

References

1. J. DE PILLIS, k-part splittings and operator parameter overrelaxation, *J. Math. Anal. Appl.* **53** (1976), 313–342.
2. R. S. VARGA, "Matrix Iterative Analysis," Prentice Hall, Englewood Cliffs, New Jersey, 1962.