APPLICATIONES MATHEMATICAE
23,4 (1996), pp. 417–432

A. ABKOWICZ and C. BREZINSKI (Lille)

ACCELERATION PROPERTIES OF THE HYBRID PROCEDURE FOR SOLVING LINEAR SYSTEMS

Abstract. The aim of this paper is to discuss the acceleration properties of the hybrid procedure for solving a system of linear equations. These properties are studied in a general case and in two particular cases which are illustrated by numerical examples.

1. The hybrid procedure. Let us consider the system of linear equations

$$Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{m\times m}$ and $x, b \in \mathbb{R}^m$. We denote by $\widetilde{x}$ the solution of (1).

Let $G = Z^T Z$ be a symmetric positive definite matrix. The G-inner product and the corresponding G-norm are respectively defined by $(x,y)_G = (x, Gy)$ and $\|x\|_G = \sqrt{(x,x)_G}$. The corresponding G-matrix norm is given by

$$\|A\|_G = \sup_{x\neq 0} \frac{\|Ax\|_G}{\|x\|_G} = \sqrt{\varrho((ZAZ^{-1})^T ZAZ^{-1})}.$$

We shall also use the notation $x \perp_G y$ if $(x,y)_G = 0$. For simplicity, the subscript G will be suppressed when unnecessary.

Let us now assume that two iterative methods for solving the system (1) are used simultaneously. Their iterates are denoted respectively by x'_n and x''_n and the corresponding residual vectors by $r'_n = b - Ax'_n$ and $r''_n = b - Ax''_n$.

The hybrid procedure defined in [1] consists of constructing a new iterate x_n and a new residual $r_n = b - Ax_n$ by

$$x_n = \alpha_n x'_n + (1-\alpha_n) x''_n, \qquad r_n = \alpha_n r'_n + (1-\alpha_n) r''_n, \tag{2}$$

1991 *Mathematics Subject Classification*: 65F10, 65B05.
Key words and phrases: acceleration, linear equations, iterative methods.

with
$$\alpha_n = -\frac{(r'_n - r''_n, r''_n)}{(r'_n - r''_n, r'_n - r''_n)}.$$
From the definition of r_n, we see that
$$\|r_n\| = \min_\alpha \|\alpha r'_n + (1-\alpha) r''_n\|.$$
We have
$$(r_n, r'_n) = (r_n, r''_n) = (r_n, r_n)$$
and, setting $p_n = r'_n - r''_n$, (2) can be written as
$$r_n = r''_n - \frac{(p_n, r''_n)}{(p_n, p_n)} p_n, \qquad r_n = r'_n - \frac{(p_n, r'_n)}{(p_n, p_n)} p_n. \tag{3}$$
It is easy to check that
$$(r_n, r_n) = \frac{(r'_n, r'_n)(r''_n, r''_n) - (r'_n, r''_n)^2}{(r'_n - r''_n, r'_n - r''_n)} \tag{4}$$
$$= (r''_n, r''_n) - \frac{(p_n, r''_n)^2}{(p_n, p_n)} \tag{5}$$
$$= (r'_n, r'_n) - \frac{(p_n, r'_n)^2}{(p_n, p_n)}. \tag{6}$$

2. Properties of the hybrid procedure. We now study the acceleration properties of the hybrid procedure.

2.1. *Asymptotic behavior of the hybrid procedure*. Let θ_n be the angle between Zr'_n and Zr''_n. Using the relation $(r'_n, r''_n) = \|r'_n\|\|r''_n\| \cos\theta_n$ we have
$$\alpha_n = -\frac{\|r'_n\|\|r''_n\| \cos\theta_n - \|r''_n\|^2}{\|r'_n\|^2 - 2\|r'_n\|\|r''_n\| \cos\theta_n + \|r''_n\|^2}$$
and
$$\|r_n\|^2 = \frac{\|r'_n\|^2 \|r''_n\|^2 (1 - \cos^2\theta_n)}{\|r'_n\|^2 - 2\|r'_n\|\|r''_n\| \cos\theta_n + \|r''_n\|^2}.$$
Setting $\varrho_n = \|r'_n\|/\|r''_n\|$ we obtain
$$\alpha_n = -\frac{\varrho_n \cos\theta_n - 1}{\varrho_n^2 - 2\varrho_n \cos\theta_n + 1},$$
$$\frac{\|r_n\|^2}{\|r'_n\|^2} = \frac{1 - \cos^2\theta_n}{\varrho_n^2 - 2\varrho_n\cos\theta_n + 1} \tag{7}$$
$$= 1 - \frac{(\varrho_n - \cos\theta_n)^2}{(\varrho_n - \cos\theta_n)^2 + \sin^2\theta_n} \tag{8}$$
$$= \frac{\sin^2\theta_n}{(\varrho_n - \cos\theta_n)^2 + \sin^2\theta_n} \tag{9}$$

$$= \frac{\sin^2 \theta_n}{\varrho_n^2 - 2\varrho_n \cos\theta_n + 1}. \tag{10}$$

From these relations, we immediately obtain

THEOREM 2.1. *Suppose that the limit* $\lim_{n\to\infty} \theta_n = \theta$ *exists.*

1. *If* $\lim_{n\to\infty} \varrho_n = 0$ *then* $\lim_{n\to\infty} \alpha_n = 1$.
2. *If* $\lim_{n\to\infty} \varrho_n = 1$ *and* $\theta \neq 0, \pi$ *then* $\lim_{n\to\infty} \alpha_n = 1/2$.
3. *If* $\lim_{n\to\infty} \varrho_n = \infty$ *then* $\lim_{n\to\infty} \alpha_n = 0$.

This theorem shows that the hybrid procedure asymptotically selects the best method among the two.

Let us now consider the convergence behavior of $\|r_n\|/\|r'_n\|$. From (10), we immediately have

THEOREM 2.2. *If the limits* $\lim_{n\to\infty} \varrho_n = \varrho$ *and* $\lim_{n\to\infty} \theta_n = \theta$ *exist and* $\varrho^2 - 2\varrho\cos\theta + 1 \neq 0$, *then*

$$\lim_{n\to\infty} \frac{\|r_n\|^2}{\|r'_n\|^2} = \frac{\sin^2\theta}{\varrho^2 - 2\varrho\cos\theta + 1} \leq 1.$$

Remark 1. Obviously if $\varrho \leq 1$, we also have $\lim_{n\to\infty} \|r_n\|^2/\|r''_n\|^2 \leq 1$. Thus $\lim_{n\to\infty} \|r_n\|/\min(\|r'_n\|, \|r''_n\|)$ exists and is not greater than 1.

Similar results can be obtained by considering the ratio $\|r_n\|^2/\|r''_n\|^2$.

It must also be noticed that $\|r_n\|^2/\|r'_n\|^2$ tends to 1 if and only if $\varrho = \cos\theta$. This result comes out directly from (8) and we also get

THEOREM 2.3. *A necessary and sufficient condition for the existence of an* N *such that*

$$0 \leq \|r_n\|^2/\|r'_n\|^2 < 1 \quad \textit{for all } n \geq N$$

is that $(r'_n - r''_n, r'_n) \neq 0$ *for all* $n \geq N$.

Proof. Suppose that $(r'_n - r''_n, r'_n) \neq 0$ for all $n \geq N$. Thus we have

$$\varrho_n - \cos\theta_n = \frac{\|r'_n\|}{\|r''_n\|} - \frac{(r'_n, r''_n)}{\|r'_n\|\|r''_n\|} = \frac{(r'_n, r'_n - r''_n)}{\|r'_n\|\|r''_n\|} \neq 0$$

and, from (8), it follows that $\|r_n\|^2/\|r'_n\|^2 < 1$. The reverse implication is proved similarly. ∎

Let us now study some cases where (r_n) converges to zero faster than (r'_n) and (r''_n). From (9), we have

THEOREM 2.4. *If there are* ϱ *and* N *such that* $0 \leq \varrho_n \leq \varrho < 1$ *for all* $n \geq N$, *then a necessary and sufficient condition for*

$$\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$$

to hold is that (θ_n) *tends to* 0 *or* π.

Proof. First let us prove the sufficiency. Suppose that (θ_n) tends to 0 or π. Thus, since $\varrho_n \le \varrho < 1$, from (9) we have $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$.

To prove the necessity, suppose that $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$. The condition $\varrho_n \le \varrho < 1$ implies that $\sin\theta_n$ tends to 0, which ends the proof. ■

Remark 2. Since $\varrho_n < 1$ we have $\|r'_n\| < \|r''_n\|$ for all $n \ge N$ and so

$$\lim_{n\to\infty} \|r_n\|/\min(\|r'_n\|, \|r''_n\|) = 0.$$

Let us now study the case where (ϱ_n) tends to 1. From (10), we first have

Theorem 2.5. *If* $\lim_{n\to\infty} \varrho_n = 1$, *then a sufficient condition for*

$$\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$$

to hold is that (θ_n) *tends to* π.

Remark 3. Since $\lim_{n\to\infty} \varrho_n = 1$, it follows that

$$\lim_{n\to\infty} \|r_n\|/\min(\|r'_n\|, \|r''_n\|) = 0.$$

Another result in the case where (ϱ_n) tends to 1 is given by

Theorem 2.6. *If* $\|r'_n\|/\|r''_n\| = 1 + a_n$ *with* $\lim_{n\to\infty} a_n = 0$, *then a sufficient condition for*

$$\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$$

to hold is that $\theta_n = o(a_n)$.

Proof. We have

$$\cos\theta_n = 1 - \theta_n^2/2 + \mathcal{O}(\theta_n^4), \qquad \sin\theta_n = \theta_n + \mathcal{O}(\theta_n^3).$$

Replacing in formula (10), we have

$$\begin{aligned}
\frac{\|r_n\|}{\|r'_n\|} &= \frac{\sin^2\theta_n}{\varrho_n^2 - 2\varrho_n\cos\theta_n + 1} \\
&= \frac{(\theta_n(1+\mathcal{O}(\theta_n^2)))^2}{(1+a_n)^2 - 2(1+a_n)(1-\theta_n^2/2+\mathcal{O}(\theta_n^4)) + 1} \\
&= \frac{\theta_n^2(1+\mathcal{O}(\theta_n^2))}{a_n^2 + \theta_n^2 + a_n\theta_n^2 + (1+a_n)\mathcal{O}(\theta_n^4)} \\
&= \frac{1+\mathcal{O}(\theta_n^2)}{(a_n/\theta_n)^2 + 1 + a_n + (1+a_n)\mathcal{O}(\theta_n^2)}
\end{aligned}$$

and the result follows. ■

Remark 4. Since $\lim_{n\to\infty} \varrho_n = 1$, the conclusion of Remark 3 still holds.

Another presentation consists in considering the angle ϑ_n between Zr'_n and Zp_n. From (6) we have $\|r_n\|^2 = \|r'_n\|^2 \sin^2 \vartheta_n$. Directly from this equation we obtain

THEOREM 2.7. *If there exists* $\vartheta \neq \pi/2$ *such that* $\lim_{n\to\infty} \vartheta_n = \vartheta$ *then* $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = |\sin\vartheta| < 1$.

Also, we have

THEOREM 2.8. $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$ *if and only if* (ϑ_n) *tends to* 0 *or* π.

These results are simpler than the preceding ones, in particular those of Theorems 2.2–2.4.

Remark 5. Similarly, if we denote by φ_n the angle between Zr''_n and Zp_n, we have $\|r_n\|^2 = \|r''_n\|^2 \sin^2 \varphi_n$. Obviously $\theta_n = \varphi_n - \vartheta_n$.

2.2. *Geometric behavior of the hybrid procedure*. A sphere in $\mathbb{R}^m$ with respect to the G-norm will be denoted by

$$\Upsilon_G(q, r) = \{y \in \mathbb{R}^m : \|y - q\|_G = r\}.$$

We have the following properties:

PROPERTY 1. $r_n \in \Upsilon_G(r'_n/2, \|r'_n\|_G/2) \cap \Upsilon_G(r''_n/2, \|r''_n\|_G/2)$.

Proof. By definition, we have $(r_n, r_n) = (r_n, r'_n) = (r_n, r''_n)$. Computing $\|r_n - r'_n/2\|^2$ we get

$$\|r_n - r'_n/2\|^2 = \|r_n\|^2 - (r_n, r'_n) + \tfrac{1}{4}\|r'_n\|^2 = \tfrac{1}{4}\|r'_n\|^2.$$

In the same way, we can prove that $\|r_n - r''_n/2\|^2 = \frac{1}{4}\|r''_n\|^2$ and the result follows. ∎

Let us denote by $e_n = \widetilde{x} - x_n$, $e'_n = \widetilde{x} - x'_n$, $e''_n = \widetilde{x} - x''_n$ the error vectors corresponding respectively to x_n, x'_n, x''_n. Using the relation $r_n = Ae_n$ and the preceding property we have

PROPERTY 2.

$$e_n \in \Upsilon_{A^TGA}(e'_n/2, \|e'_n\|_{A^TGA}/2) \cap \Upsilon_{A^TGA}(e''_n/2, \|e''_n\|_{A^TGA}/2).$$

The hybrid procedure is a projection method because there exists a matrix $\wp_n \in \mathbb{R}^{m\times m}$ such that

$$r_n = \wp_n r'_n = \wp_n r''_n \quad \text{with} \quad \wp_n = I - \frac{p_n p_n^T G}{p_n^T G p_n}.$$

It is easy to see that $\wp_n^2 = \wp_n$ and $(G\wp_n)^T = G\wp_n$. So $\wp_n$ is a G-orthogonal projection. By definition of $\wp_n$ we get

$$\begin{aligned} \wp_n v &= v \quad \text{if } v \perp_G p_n, \\ \wp_n v &= 0 \quad \text{if } v \in \operatorname{span}\{p_n\}. \end{aligned}$$

The above results can be considered as a generalization of the results given in [8].

3. Applications. It seems quite difficult to obtain more theoretical results on the convergence of the hybrid procedure in the general case. So, (r'_n) being an arbitrary sequence of residual vectors, we shall assume that we are in one of the following particular cases:

(i) $r''_n = Br_{n-1}$,
(ii) $r''_n = Br'_n$.

Such a situation arises, for example, if we consider a splitting of the matrix A,

$$A = M - N,$$

and if x''_n is obtained from y (equal to x_{n-1} or x'_n) by

$$x''_n = M^{-1}Ny + M^{-1}b.$$

In this case the associated residual has the form

$$\begin{aligned} r''_n &= b - Ax''_n = b - A(M^{-1}Ny + M^{-1}b) \\ &= b - (M-N)(M^{-1}Ny + M^{-1}b) = NM^{-1}(b - Ay). \end{aligned}$$

Thus we have $B = NM^{-1}$ with $y = x_{n-1}$ (case (i)) and $y = x'_{n-1}$ (case (ii)). It must be noticed that $B = I - AM^{-1}$. This situation also holds if $B = I - AC$ with C an arbitrary matrix. In this case, we have

$$x_n = \alpha_n x'_n + (1-\alpha_n)(y + C(b - Ay)).$$

(We are indebted to one of the referees for this remark.)

3.1. *Case* (i). Let r_n be computed by the hybrid procedure from $r''_n = Br_{n-1}$ and r'_n. We have

$$r_n = \alpha_n r'_n + (1-\alpha_n)Br_{n-1}$$

and we get

LEMMA 3.1. *Let $r_0 = r'_0$. Then, for all $n \ge 1$,*

$$H1(n) \qquad r_n = \sum_{i=0}^{n} a_i^{(n)} B^{n-i} r'_i$$

with

$$H2(n) \qquad \sum_{i=0}^{n} a_i^{(n)} = 1.$$

Proof. $a_0^{(0)} = 1$ and so $H1(0)$ and $H2(0)$ are true. Suppose that $H1(n-1)$ and $H2(n-1)$ hold. From the definition of r_n and from $H1(n-1)$, we get

$$r_n = \alpha_n r'_n + (1-\alpha_n) B r_{n-1} = \alpha_n r'_n + (1-\alpha_n) B \sum_{i=0}^{n-1} a_i^{(n-1)} B^{n-1-i} r'_i$$
$$= \alpha_n r'_n + \sum_{i=0}^{n-1} (1-\alpha_n) a_i^{(n-1)} B^{n-i} r'_i = \sum_{i=0}^{n} a_i^{(n)} B^{n-i} r'_i,$$

where the $a_i^{(n)}$'s are given by

$$a_i^{(n)} = (1-\alpha_n) a_i^{(n-1)}, \quad i = 0, \ldots, n-1,$$
$$a_n^{(n)} = \alpha_n.$$

Thus $H1(n)$ is true with $H2(n)$ obviously satisfied. ∎

Remark 6. When r'_n is computed by a polynomial method of the form $r'_n = P_n(B) r'_0$ then $r_n = Q_n(B) r'_0$ with Q_n given by $Q_n(t) = \alpha_n P_n(t) + (1-\alpha_n) t Q_{n-1}(t)$.

Let us now prove other results. We have

THEOREM 3.2. *Let γ be an eigenvector of B. If $r_n = c_n \gamma + a_n$ with*

$$\lim_{n\to\infty} (\gamma, r'_n)/\|r'_n\| = \|\gamma\| \quad \textit{and} \quad \lim_{n\to\infty} \|a_n\|/c_n = 0$$

then $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$.

Proof. Let θ_n be the angle between ZBr_{n-1} and Zr'_n. We have

$$(Br_{n-1}, r'_n)^2 = c_{n-1}^2 \lambda^2 (\gamma, r'_n)^2 + (Ba_{n-1}, r'_n)^2 + 2c_{n-1}\lambda(\gamma, r'_n)(Ba_{n-1}, r'_n),$$
$$\|Br_{n-1}\|^2 = c_{n-1}^2 \lambda^2 \|\gamma\|^2 + \|Ba_{n-1}\|^2 + 2c_{n-1}\lambda(\gamma, Ba_{n-1}),$$

where λ is the eigenvalue of B corresponding to γ. Thus

$$\lim_{n\to\infty} \cos^2\theta_n = \lim_{n\to\infty} \frac{(Br_{n-1}, r'_n)^2}{\|Br_{n-1}\|^2 \|r'_n\|^2}$$
$$= \lim_{n\to\infty} \left[\frac{\lambda^2 \dfrac{(\gamma, r'_n)^2}{\|r'_n\|^2} + \dfrac{(Ba_{n-1}, r'_n)^2}{c_{n-1}^2 \|r'_n\|^2} + 2\lambda \dfrac{(\gamma, r'_n)(Ba_{n-1}, r'_n)}{c_{n-1}\|r'_n\|^2}}{\lambda^2 \|\gamma\|^2 + \dfrac{\|Ba_{n-1}\|^2}{c_{n-1}^2} + 2\lambda \dfrac{(\gamma, a_{n-1})}{c_{n-1}}} \right]$$
$$= 1$$

and the result follows from Theorem 2.4. ∎

From the minimization property of r_n we have

$$\|r_n\| \le \|Br_{n-1}\| \le \|B\| \|r_{n-1}\|$$

and thus $\|r_n\| \le \|r_{n-1}\|$ if $\|B\| \le 1$.

In particular, consider a splitting $A = M - N$ of the matrix A. Premultiplying the system (1) from the right by M^{-1} we get a new system of the form

$$A^{(M)}x = b^{(M)}$$

with

$$B^{(M)} = M^{-1}N, \quad A^{(M)} = I - B^{(M)}, \quad b^{(M)} = M^{-1}b.$$

Applying the method described above to this new system, we get

$$\|r_n\| \leq \|B^{(M)}\|\|r_{n-1}\|.$$

Thus, a good choice of $B^{(M)}$ is equivalent to a good choice of the preconditioner M from the right-hand side.

When $B = I$, the method is called the *Minimal Residual Smoothing* (MRS) algorithm. It was introduced in [6, 7] and applied to some well known methods. For more details, see [2, 9–12].

We now apply it to an error-minimization method [4]. Set $e'_n = \widetilde{x} - x'_n$ and $e_n = \widetilde{x} - x_n$. Let φ be any norm in $\mathbb{R}^m$. For any $x \in \mathbb{R}^m$ we denote by $z(x)$ a vector such that

$$(z(x), x) = \varphi(x).$$

This is called a *decomposition* of the norm φ. Such decompositions were introduced by Gastinel [3] for the case of the l_1-norm.

Let x'_0 be a given vector. The *Transformed Norm Decomposition Method* (TNDE) [4] is defined by

$$r'_0 = b - Ax'_0, \qquad p'_0 = A^T z_0,$$

and for $n = 0, 1, \ldots,$

$$x'_{n+1} = x'_n - \beta_n p'_n, \qquad r'_{n+1} = r'_n + \beta_n A p'_n,$$
$$p'_{n+1} = A^T z_{n+1} + \sum_{i=0}^{n} \gamma_{n+1}^{(i)} p'_i,$$

where z_i is such that $(z_i, r_i) = \varphi(r_i)$. The coefficients β_n and $\gamma_{n+1}^{(i)}$ are given by

$$\beta_n = \frac{(p'_n, e'_n)}{(p'_n, p'_n)} = -\frac{\varphi(r'_n)}{(p'_n, p'_n)}$$

and

$$\gamma_{n+1}^{(i)} = -\frac{(p'_i, A^T z_{n+1})}{(p'_i, p'_i)}, \qquad i = 0, \ldots, n.$$

The sequence $(e'_n = \widetilde{x} - x'_n)$ has the following properties:

1. $e'_n \in V_n = e'_0 + \operatorname{span}\{p'_0, \ldots, p'_n\}$,
2. $V_{n-1} \subset V_n$,
3. $\|e'_n\| = \min_{e \in V_n} \|e\|$,
4. $\|e'_n\| \leq \|e'_{n-1}\|$.

If the MRS is applied to the TNDE, that is, to the sequence (r'_n) defined above with $r_0 = r'_0$, then we have

THEOREM 3.3. *If* $\alpha_n \in]0, 1[$ *then* $\|e_n\| \le \|e_{n-1}\|$.

Proof. We have

$$e_{n-1} = \alpha_{n-1} e'_{n-1} + (1 - \alpha_{n-1}) e_{n-2}.$$

It follows that $e_{n-1} = \sum_{i=1}^{n-1} a_i^{(n-1)} e'_i \in V_{n-1}$ for all n. Thus, from property 3,

$$\|e'_{n-1}\| \le \|e_{n-1}\|.$$

Using property 4, we get

$$\begin{aligned}\|e_n\| &\le \alpha_n \|e'_n\| + (1-\alpha_n)\|e_{n-1}\| \le \alpha_n \|e'_{n-1}\| + (1-\alpha_n)\|e_{n-1}\| \\ &\le \alpha_n \|e_{n-1}\| + (1-\alpha_n)\|e_{n-1}\| = \|e_{n-1}\|\end{aligned}$$

and the result follows. ■

3.2. *Case* (ii). Suppose now that r_n is given by

$$r_n = \alpha_n r'_n + (1-\alpha_n) B r'_n$$

with $B = I - AM^{-1}$. Then r_n can be written as

$$r_n = r'_n - \frac{(AM^{-1} r'_n, r'_n)}{(AM^{-1} r'_n, AM^{-1} r'_n)} AM^{-1} r'_n.$$

Remark 7. If $r'_n = r_{n-1}$ and $M = I$, then the hybrid procedure is identical to the Minimal Residual Method.

DEFINITION 1. Consider two vector sequences $(u_n), (v_n) \in \mathbb{R}^m$ such that $\lim_{n\to\infty} u_n = u$ and $\lim_{n\to\infty} v_n = v$. We say that (u_n) *converges with the same speed as* (v_n) if there exists N such that for all $n \ge N$ there are $M_n \in \mathbb{R}^{m\times m}$ and $a_n \in \mathbb{R}^m$ with $\|a_n\| \le \varepsilon$ such that

- $v_{n+1} = M_n v_n$,
- $u_{n+1} = M_n u_n + a_n$.

LEMMA 3.4. *Suppose that there exists* N *such that for all* $n \ge N$, *there is* $M_n \in \mathbb{R}^{m\times m}$ *such that* $r'_{n+1} = M_n r'_n$ *and* $AM^{-1} M_n = M_n AM^{-1}$. *If* $\lim_{n\to\infty} \alpha_n = \alpha$ *exists and if there is* K *such that* $\|M_n\| < K$ *for all* n, *then* (r_n) *converges with the same speed as* (r'_n).

Proof. If (α_n) converges, then there is a sequence (ε_n) with $\lim_{n\to\infty} \varepsilon_n = 0$ such that $\alpha_{n+1} = \alpha_n - \varepsilon_n$ for all n. Setting $a_n = \varepsilon_n AM^{-1} M_n r'_n$, we get from the definition

$$\begin{aligned}r_{n+1} &= r'_{n+1} - (1-\alpha_{n+1}) AM^{-1} r'_{n+1} = M_n r'_n - (1 - \alpha_n + \varepsilon_n) AM^{-1} M_n r'_n \\ &= M_n (r'_n - (1-\alpha_n) AM^{-1} r'_n) + \varepsilon_n AM^{-1} M_n r'_n = M_n r_n + a_n.\end{aligned}$$

Obviously $\lim_{n\to\infty} a_n = 0$ and the result follows. ■

We now assume that $r'_n = c_n\gamma + a_n$, where $c_n \in \mathbb{R}$, $a_n \in \mathbb{R}^m$, and that γ is an eigenvector of B. In this case we get

LEMMA 3.5. *Let γ be an eigenvector of B. If $r'_n = c_n\gamma + a_n$, then there are $K \in \mathbb{R}$ and $M_n \in \mathbb{R}^m$ such that for all n, $\|M_n\| \le K$ and $r_n = M_n a_n$.*

Proof. We know that

$$r_n = \wp_n r'_n = c_n \wp_n \gamma + \wp_n a_n$$

with

$$\wp_n = I - p_n p_n^T G/(p_n^T G p_n),$$

where $p_n = AM^{-1} r'_n$. Premultiplying p_n by $\wp_n$ we get

$$0 = \wp_n p_n = (1-\lambda) c_n \wp_n \gamma + \wp_n AM^{-1} a_n,$$

where λ is the eigenvalue of B corresponding to γ. Thus, since A is assumed to be regular,

$$c_n \wp_n \gamma = -\frac{1}{1-\lambda} \wp_n AM^{-1} a_n.$$

Setting

$$M_n = \wp_n \left(I - \frac{1}{1-\lambda} AM^{-1} \right),$$

we get $r_n = M_n a_n$. The matrix $\wp_n$ is a G-orthogonal projection and thus $\|\wp_n\|_G = 1$. It follows that

$$\|M_n\| \le 1 + \frac{1}{|1-\lambda|} \|AM^{-1}\|$$

which ends the proof. ∎

Remark 8. As a consequence of Lemma 3.5 we have $\|r_n\| = \mathcal{O}(\|a_n\|)$.

From Theorem 2.8, we easily get

THEOREM 3.6. *Let γ be an eigenvector of B with the corresponding eigenvalue λ. If $r'_n = c_n\gamma + a_n$ with $\lim_{n\to\infty} \|a_n\|/c_n = 0$ then*

$$\lim_{n\to\infty} \alpha_n = -\frac{\lambda}{1-\lambda} \quad \text{and} \quad \lim_{n\to\infty} \frac{\|r_n\|}{\|r'_n\|} = 0.$$

Proof. We have

$$r'_n = c_n\gamma + a_n, \quad Br'_n = \lambda c_n \gamma + Ba_n, \quad AM^{-1} r'_n = (1-\lambda) c_n \gamma + AM^{-1} a_n,$$

and

$$\begin{aligned}
\lim_{n\to\infty} \alpha_n &= -\frac{(Br'_n, AM^{-1}r'_n)}{(AM^{-1}r'_n, AM^{-1}r'_n)} \\
&= \lim_{n\to\infty} \left[-\frac{\lambda(1-\lambda)c_n^2(\gamma,\gamma) + c_n\lambda(\gamma, AM^{-1}a_n)}{(1-\lambda)^2c_n^2(\gamma,\gamma) + 2(1-\lambda)c_n(\gamma, AM^{-1}a_n) + (AM^{-1}a_n, AM^{-1}a_n)} \right. \\
&\qquad \left. + \frac{(1-\lambda)c_n(Ba_n,\gamma) + (Ba_n, AM^{-1}a_n)}{(1-\lambda)^2c_n^2(\gamma,\gamma) + 2(1-\lambda)c_n(\gamma, AM^{-1}a_n) + (AM^{-1}a_n, AM^{-1}a_n)} \right] \\
&= \lim_{n\to\infty} \left[-\frac{\lambda(1-\lambda)(\gamma,\gamma) + \lambda\dfrac{(\gamma, AM^{-1}a_n)}{c_n}}{(1-\lambda)^2(\gamma,\gamma) + 2(1-\lambda)\dfrac{(\gamma, AM^{-1}a_n)}{c_n} + \dfrac{(AM^{-1}a_n, AM^{-1}a_n)}{c_n^2}} \right. \\
&\qquad \left. + \frac{(1-\lambda)\dfrac{(Ba_n,\gamma)}{c_n} + \dfrac{(Ba_n, AM^{-1}a_n)}{c_n^2}}{(1-\lambda)^2(\gamma,\gamma) + 2(1-\lambda)\dfrac{(\gamma, AM^{-1}a_n)}{c_n} + \dfrac{(AM^{-1}a_n, AM^{-1}a_n)}{c_n^2}} \right] \\
&= -\frac{\lambda}{1-\lambda}.
\end{aligned}$$

Let θ_n be the angle between Zr'_n and $ZAM^{-1}r'_n$. Replacing r'_n and $AM^{-1}r'_n$ by their expressions above, we also get

$$\begin{aligned}
\lim_{n\to\infty} \cos^2\theta_n &= \lim_{n\to\infty} \frac{(r'_n, AM^{-1}r'_n)^2}{\|r'_n\|^2\|AM^{-1}r'_n\|^2} \\
&= \lim_{n\to\infty} \left[\frac{1}{c_n^2(\gamma,\gamma) + 2c_n(\gamma, a_n) + (a_n, a_n)} \right. \\
&\qquad \left. \times \frac{[(1-\lambda)c_n^2(\gamma,\gamma) + c_n(\gamma, AM^{-1}a_n) + (1-\lambda)c_n(a_n,\gamma) + (a_n, AM^{-1}a_n)]^2}{(1-\lambda)^2c_n^2(\gamma,\gamma) + 2(1-\lambda)c_n(\gamma, AM^{-1}a_n) + (AM^{-1}a_n, AM^{-1}a_n)} \right] \\
&= \lim_{n\to\infty} \left[\frac{1}{(\gamma,\gamma) + 2\dfrac{(\gamma,a_n)}{c_n} + \dfrac{(a_n,a_n)}{c_n^2}} \right. \\
&\qquad \left. \times \frac{\left[(1-\lambda)(\gamma,\gamma) + \dfrac{(\gamma, AM^{-1}a_n)}{c_n} + (1-\lambda)\dfrac{(a_n,\gamma)}{c_n} + \dfrac{(a_n, AM^{-1}a_n)}{c_n^2}\right]^2}{(1-\lambda)^2(\gamma,\gamma) + 2(1-\lambda)\dfrac{(\gamma, AM^{-1}a_n)}{c_n} + \dfrac{(AM^{-1}a_n, AM^{-1}a_n)}{c_n^2}} \right] \\
&= 1
\end{aligned}$$

and the result follows by Theorem 2.8. ■

The conditions of Lemma 3.5 and Theorem 3.6 seem difficult to check in practice. We now give an example where these results can be applied.

EXAMPLE. Let $\{\lambda_i\}_{i=1}^m$ be the eigenvalues of $B = I - A$ with the corresponding eigenvectors $\{\gamma_i\}_{i=1}^m$. Suppose that $|\lambda_1| \geq \ldots \geq |\lambda_m|$ and that the eigenvectors form a basis of $\mathbb{R}^m$. Let r'_n be such that $r'_n = Br'_{n-1}$ and let r_n be obtained by the hybrid procedure from r'_n and r'_{n+1}. Let $r'_0 = \sum_{i=1}^m d_i\gamma_i$. Thus

$$r'_n = \sum_{i=1}^m d_i\lambda_i^n\gamma_i = d_1\lambda_1^n\gamma_1 + \sum_{i=2}^m d_i\lambda_i^n\gamma_i.$$

Setting

$$c_n = d_1\lambda_1^n, \qquad a_n = \sum_{i=2}^m d_i\lambda_i^n\gamma_i,$$

we get from Remark 8 and Theorem 3.6

THEOREM 3.7. *If* $r'_n = Br'_{n-1}$, $r_0 = r'_0$ *and if* r_n *is obtained by the hybrid procedure from* r'_n *and* r'_{n+1}, *then* $\|r_n\| = \mathcal{O}(|\lambda_2|^n)$. *Moreover*, *if* $|\lambda_2| < |\lambda_1|$, *then* $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$.

Remark 9. This theorem holds even if $|\lambda_2| < 1 < |\lambda_1|$.

Remark 10. Since (α_n) converges, Lemma 3.4 shows that (r_n) converges with the same speed as (r'_n). In this case, the iterations will be stopped when $|\alpha_n + \lambda_1/(1-\lambda_1)| \leq \varepsilon$, where ε is an arbitrary threshold. Of course the value of λ_1 is usually unknown and this test cannot be used in practice. Thus the iterations will be stopped when $|\alpha_n - \alpha_{n-1}| \leq \varepsilon$. However, it must be noticed that, due to a possible stagnation of the method, this test does not guarantee that the recurrence is close to the limit.

4. Numerical examples. In all the examples we take $G = I$, $M = I$, $B = I - A$ and $x_0 = 0$. The right-hand side is computed in order that the solution be $\widetilde{x} = [1,\ldots,1]^T$. Each figure shows $\log\|r'_n\|$ and $\log\|r_n\|$ as a function of the number n of iterations and the lowest curve always corresponds to the hybrid procedure.

Let $\{\lambda_i\}_{i=1}^m$ be the set of eigenvalues of B. The elements of the matrix $A \in \mathbb{R}^{100\times100}$ were randomly chosen in $[0,1]$. The values of $\|B\|$, $|\lambda_i|$ ($i = 1,\ldots,100$) were computed with Matlab with a precision of 10^{-20}.

4.1. *Case* (i). Let r'_n be obtained by the norm decomposition method of Gastinel [3] with $\varphi_1(r) = \sum_{i=1}^m |r_i|$. This method is as follows: for $n = 0,1,\ldots,$

$$x'_{n+1} = x'_n - \alpha'_n A^T z_n, \qquad r'_{n+1} = r'_n + \alpha'_n AA^T z_n,$$

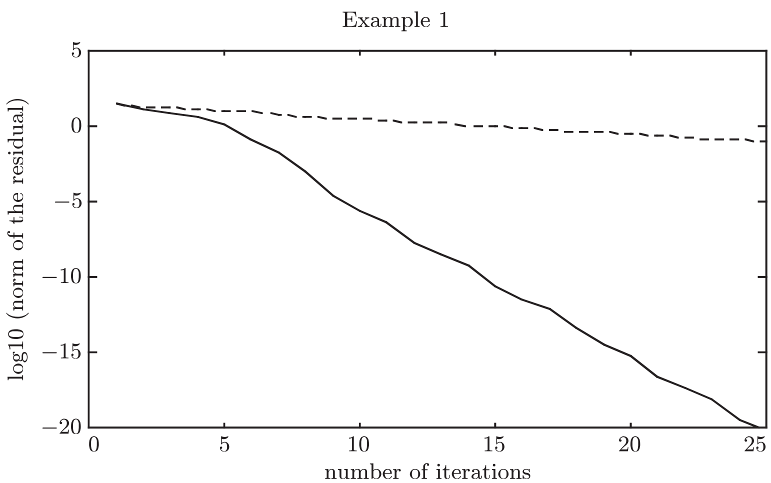
where $z_n = \mathrm{sgn}(r'_n)$. Thus, $(z_n, r'_n) = \varphi_1(r'_n)$ and

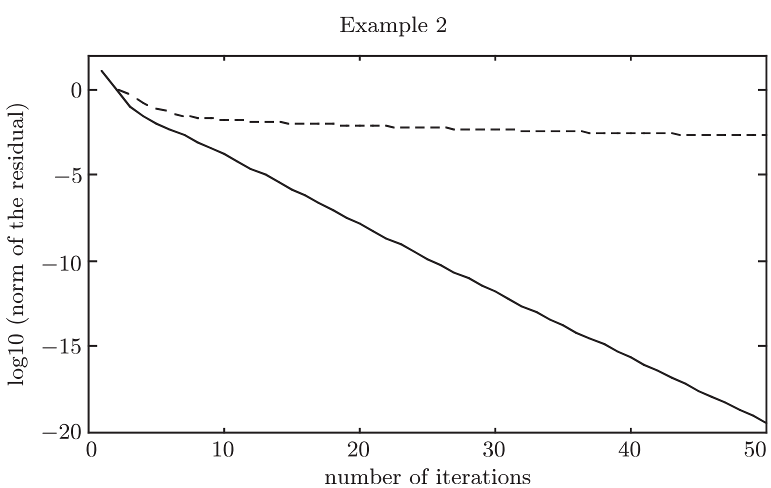
$$\alpha'_n = -\frac{\varphi_1(r'_n)}{(A^T z_n, A^T z_n)}.$$

Let r_n be computed by the hybrid procedure from r'_n and Br_{n-1}.

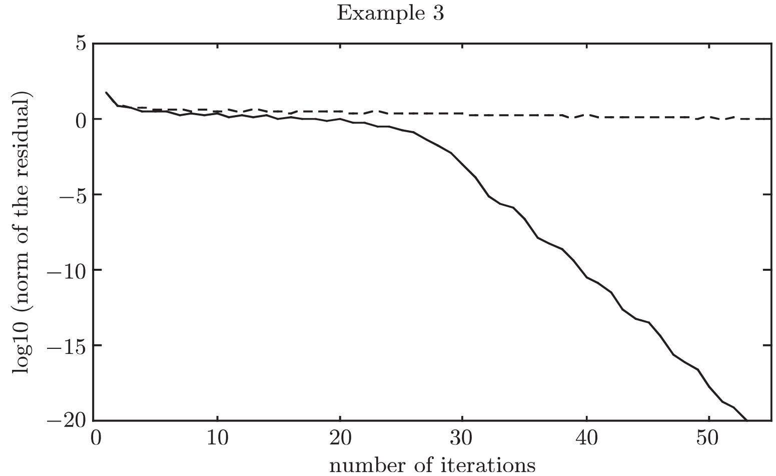
	Example 1	Example 2	Example 3
$\lVert B \rVert$	0.998605	0.663839	1.485374
$\lvert\lambda_1\rvert$	0.030418	0.661562	0.078104
$\lvert\lambda_2\rvert$	0.030372	0.040093	0.046249

For each example $|\lambda_2| < |\lambda_1|$ and thus condition 2 of Theorem 3.2 is satisfied. We did not check condition 1 but the numerical results show that, in this case, the convergence of Gastinel's method has been accelerated.





Example 3



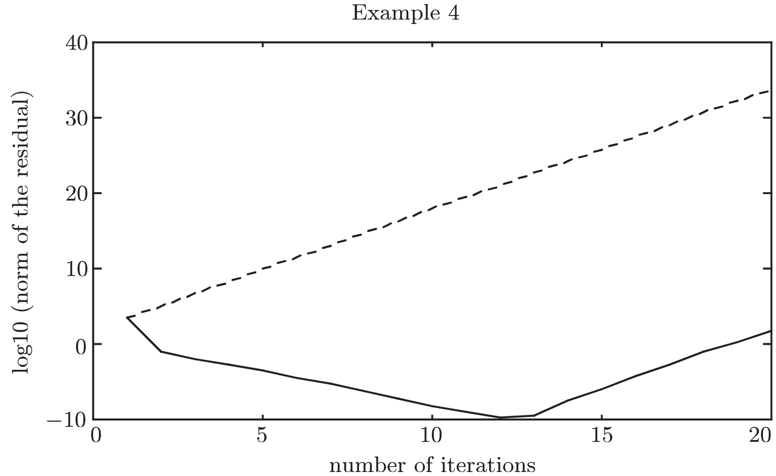
4.2. *Case* (ii). Let r'_n be such that $r'_n = Br'_{n-1}$ and let r_n be computed by the hybrid procedure from r'_n and r'_{n+1}.

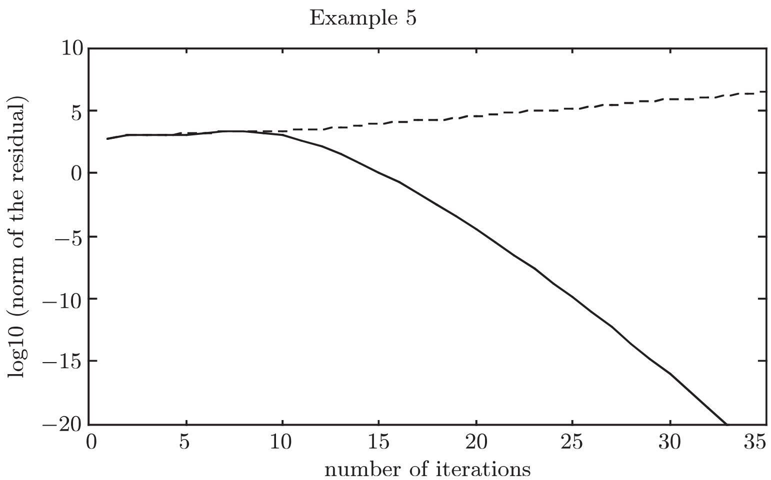
	Example 4	Example 5	Example 6
$\|B\|$	6.296298	3.273282	6.457731
$\|\lambda_1\|$	6.274695	1.158723	0.822448
$\|\lambda_2\|$	0.380272	0.099341	0.195185

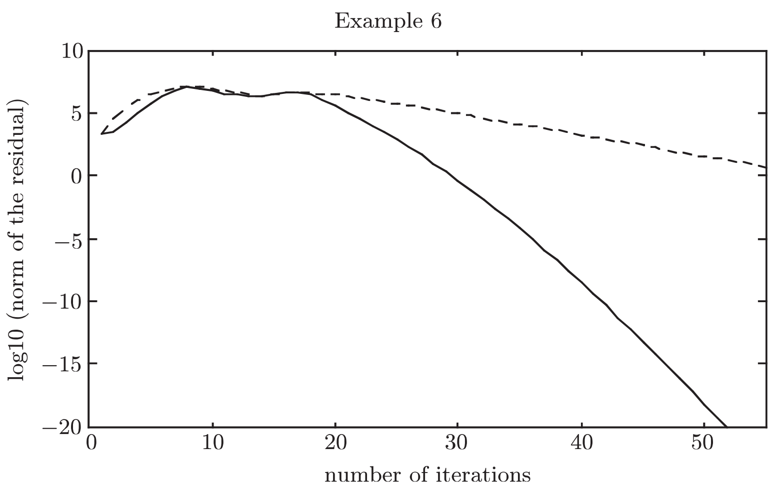
Let N be the index such that $|\alpha_N + \lambda_1/(1-\lambda_1)| \le 10^{-20}$. We get

	Example 4	Example 5	Example 6
N	12	> 35	> 55
$\log \|r'_N\|$	20.992904		
$\log \|r_N\|$	-9.771103		

Example 4







For each example we have $|\lambda_2| < |\lambda_1|$. Thus the conditions of Theorem 3.7 are satisfied and we get $\lim_{n\to\infty} \|r_n\|/\|r'_n\| = 0$ even if $\lim_{n\to\infty} \|r'_n\| = \infty$ (see Examples 4 and 5). For Example 1 we get, at iteration 12, $|\alpha_{12} + \lambda_1/(1-\lambda_1)| \leq 10^{-20}$. Moreover, we also have $|\alpha_n + \lambda_1/(1-\lambda_1)| \leq 10^{-20}$ for $n \in [12, 20]$, and thus we see that (r_n) converges with the same speed as (r'_n). We can also remark that, since the sequence (r'_n) diverges, so does (r_n) (from iteration 12) and thus it is better to stop the iterations at $n = 12$.

References

[1] C. Brezinski and M. Redivo Zaglia, *Hybrid procedures for solving linear systems*, Numer. Math. 67 (1994), 1–19.

[2] R. W. Freund, *A transpose-free quasi-minimal residual algorithm for non-Hermitian linear systems*, SIAM J. Sci. Statist. Comput. 14 (1993), 470–482.
[3] N. Gastinel, *Procédé itératif pour la résolution numérique d'un système d'équations linéaires*, C. R. Acad. Sci. Paris 246 (1958), 2571–2574.
[4] K. Jbilou, *Projection-minimization methods for nonsymmetric linear systems*, Linear Algebra Appl. 229 (1995), 101–125.
[5] —, *G-orthogonal projection methods for solving linear systems*, to appear.
[6] W. Schönauer, *Scientific Computing on Vector Computers*, North-Holland, Amsterdam, 1987.
[7] W. Schönauer, H. Müller and E. Schnepf, *Numerical tests with biconjugate gradient type methods*, Z. Angew. Math. Mech. 65 (1985), T400–T402.
[8] R. Weiss, *Convergence behavior of generalized conjugate gradient methods*, Ph.D. Thesis, University of Karlsruhe, 1990.
[9] —, *Error-minimizing Krylov subspace methods*, SIAM J. Sci. Statist. Comput. 15 (1994), 511–527.
[10] —, *Properties of generalized conjugate gradient methods*, Numer. Linear Algebra Appl. 1 (1994), 45–63.
[11] R. Weiss and W. Schönauer, *Accelerating generalized conjugate gradient methods by smoothing*, in: *Iterative Methods in Linear Algebra*, R. Beauwens and P. de Groen (eds.), North-Holland, Amsterdam, 1992, 283–292.
[12] L. Zhou and H. F. Walker, *Residual smoothing techniques for iterative methods*, SIAM J. Sci. Statist. Comput. 15 (1994), 297–312.

Anna Abkowicz and Claude Brezinski
Laboratoire d'Analyse Numérique et d'Optimisation
UFR IEEA–M3
Université des Sciences et Technologies de Lille
F-59655 Villeneuve d'Ascq Cedex, France
E-mail: brezinsk@omega.univ-lille1.fr

Received on 24.1.1995