Atti Accademia Nazionale dei Lincei

Classe Scienze Fisiche Matematiche Naturali

Rendiconti

David W. Showalter, Adi Ben-Israel

Representation and computation of the generalized inverse of a bounded linear operator between Hilbert spaces

Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti, Serie 8, Vol. **48** (1970), n.2, p. 184–194.

Accademia Nazionale dei Lincei

<http://www.bdim.eu/item?id=RLINA_1970_8_48_2_184_0>

Analisi matematica. — *Representation and computation of the generalized inverse of a bounded linear operator between Hilbert spaces.* Nota di DAVID W. SHOWALTER e ADI BEN–ISRAEL, presentata (*) dal Socio B. SEGRE.

RIASSUNTO. — Si dànno rappresentazioni in forma di integrale ed in forma di sviluppo in serie dell'inversa generalizzata di un operatore lineare limitato tra spazi di Hilbert (a codominio arbitrario). Si presentano inoltre dei metodi iterativi iperpotenziali di calcolo dell'inversa generalizzata mediante i proiettori associati.

§ 1. INTRODUCTION.

Series and integral representation of g.i. (generalized inverses) of bounded operators with closed range between Hilbert spaces were given in [8]. This paper is a continuation, giving the extensions to bounded operators with arbitrary ranges suggested at the end of [8], and hyperpower iterative methods for computing g.i. and projections for the same cases, extending the results of [3], [7] and [9].

To motivate the idea behind our development consider the problem of minimizing (1)

$$f(x) = (Ax - y, Ax - y). \tag{1.1}$$

Treating x as a function $x(t)$, $t \geq 0$ with $x(0) = 0$, we differentiate (1.1)

$$\begin{aligned} D_t f(x) &= 2 \operatorname{Re}(Ax - y, A\dot{x}), \qquad \dot{x} = D_t x \\ &= 2 \operatorname{Re}(A^*(Ax - y), \dot{x}) \end{aligned} \tag{1.2}$$

and setting

$$\dot{x} = -A^*(Ax - y) \tag{1.3}$$

it follows from (1.2) that

$$D_t f(x) = -2 \|A^*(Ax - y)\|^2 < 0. \tag{1.4}$$

This version of the steepest descent method results in $f(x(t))$ being a monotone decreasing function of t, asymptotically approaching its infimum as $t \to \infty$. We expect $x(t)$ to approach asymptotically $A^+ y$, so by solving (1.3)

$$x(t) = \int_0^t \exp[-A^* A(t - s)] A^* y \, ds \tag{1.5}$$

(*) Nella seduta del 14 febbraio 1970.

(1) If necessary, consult the notation in § 2.

and observing that y is arbitrary we get

(1.6) $$A^{+} = \lim_{t \to \infty} \int_0^t \exp[-A^* A (t - s)] A^* ds$$

which is the essence of § 3.

Here as elsewhere in this paper, the convergence is in the strong operator topology.

A numerical integration of (1.3) with suitably chosen step size similarly results in

(1.7) $$A^{+} = \sum_{k=0}^{\infty} (I - \alpha A^* A)^k \alpha A^*$$

where

(1.8) $$0 < \alpha < \frac{2}{\|A\|^2},$$

which is the essence of sections 4 and 5.

§ 2. NOTATION AND PRELIMINARIES.

S_1, S_2 – Hilbert spaces

A – a bounded linear operator $A : S_1 \to S_2$

A^* – the adjoint of A

I_1, I_2 – identity maps on S_1, S_2 resp.

$\bar{S}_1 \equiv cl(A^* S_2)$ – the closure of $A^* S_2$

$\bar{S}_2 \equiv cl(AS_1)$ – the closure of AS_1

$\bar{S}_1^{\perp}, \bar{S}_2^{\perp}$ – the complements of $\bar{S}_1, \bar{S}_1$ resp.

$\bar{A}, \bar{I}_1, \bar{A}^*, \bar{I}_2$ – the restrictions of these operators to $\bar{S}_1$ or $\bar{S}_2$, as appropriate

$\bar{A}^{\perp}, \bar{I}_1^{\perp}, \bar{A}^{*\perp}, \bar{I}_2^{\perp}$ – similarly for $\bar{S}_1^{\perp}, \bar{S}_2^{\perp}$.

$\bar{I}_1^{\perp}$ on all of S_1 means $\bar{I}_1^{\perp} + \bar{O}$, similarly $\bar{I}_1 = \bar{I}_1 + \bar{O}^{\perp}$, etc.

It is readily seen that

$\bar{A}^{\perp} = O$, $\bar{A}^{*\perp} = O$;

$\bar{I}_1, \bar{I}_2, \bar{I}_1^{\perp}$ and $\bar{I}_2^{\perp}$ are projections;

$I_1 = \bar{I}_1 + \bar{I}_1^{\perp}$ and $I_2 = \bar{I}_2 + \bar{I}_2^{\perp}$.

Thus for every $x \in S_1$, $x = \bar{x} + \bar{x}^{\perp}$ where $\bar{x} = \bar{I}_1 x$, $\bar{x}_1^{\perp} = \bar{I}_1^{\perp} x$; and similarly for any $y \in S_2$, $y = \bar{y} + \bar{y}^{\perp}$.

$\|\cdot\|$ – the inner product norm on S_1 or S_2, as appropriate, or the (conventional) associated operator norm.

A^{+} – the generalized inverse of A, e.g. [1], [2], [4], [5] and [6].

It is defined as the linear operator (unique on its domain of definition) satisfying:

(2.1) $$A^{+} A = \bar{I}_1$$

(2.2) $$A^{+} \bar{S}_2^{\perp} = O$$

A^{+} will not be defined on the set $\{\bar{S}_2 - AS_1\}$, whenever this set is nonempty. However

(2.3) $\bar{I}_2$ is the unique continuous extension of AA^{+} to all of S_2.

The case where $\overline{S}_2 - AS_1 = \varnothing$ was considered (for the formulae examined in this paper) previously [8]. Hence, in this paper we assume $\overline{S}_2 \neq AS_1$, i.e., A does not have closed range. In this latter case, there are three important classes for a point, y, in S_2. We abbreviate these as follows:

$$\begin{aligned} &(y \in \mathrm{I}) && - \text{ means } \bar{y} \text{ is in } AA^* S_2 \\ &(y \in \mathrm{II}) && - \text{ means } \bar{y} \in (AS_1 - AA^* S_2) \\ &(y \in \mathrm{III}) && - \text{ means } \bar{y} \in (\overline{S}_2 - AS_1) \end{aligned} \tag{2.4}$$

For simplicity, we have adopted the following convention: The y–variable (y, $y(t)$, y_N, etc.) is always in $\overline{S}_2$. The x–variable (x, $x(t)$, x_N, etc.) is always equal A^+y (if defined); and similarly w (if it exists) is given by $w = A^{*+}x$.

(2.5) Thus, when defined

$$x = A^+ y \;,\quad x(t) = A^+ y(t)\,,\quad \text{etc.}$$

$$w = A^{*+} x = A^{*+} A^+ y \;,\quad w_k = A^{*+} x_k = A^{*+} A^+ y_k\,,\quad \text{etc.}$$

Also

$$y = \bar{y} \;,\quad x = \bar{x} \;,\quad w = \bar{w}\,,$$

$$x = A^* w \;,\quad y = Ax = AA^* w \quad \text{when } x \text{ and } w \text{ are defined;}$$

$$y(t) = Ax(t) \;,\quad y_k = Ax_k\,,\quad \text{etc.}$$

If $z \in S_2$, then $A^* z = A^* \bar{z}$, and since the classification (2.4) is a property of the component in $\overline{S}_2$, $(z \in i) \Longleftrightarrow (\bar{z} \in i)$ for $i = \mathrm{I}, \mathrm{II}, \mathrm{III}$. Thus it should be noted in all succeeding theorems in this paper: The simplifying assumption that $\bar{y}^{\perp} = 0$ always, is not necessary; and if $\bar{y}^{\perp} \neq 0$, the theorems are true either when we set $\bar{y} = y$, or as stated.

§ 3. Integral representation of A^+.

For $t \geq 0$ we define

$$\begin{aligned} L_1(t) &= \int_0^t \exp\left[-A^*A(t-s)\right] ds\,, \\ L_2(t) &= \int_0^t \exp\left[-AA^*(t-s)\right] ds\,, \\ B(t) &= L_1(t) A^* = A^* L_2(t)\,. \end{aligned} \tag{3.1}$$

The integral expressions for $\overline{L}_1$ and $\overline{L}_2$ may be formally evaluated

$$\overline{L}_1(t) = (\overline{I}_1 - \exp[-A^*At])(A^*A)^+, \quad \overline{L}_2(t) = (\overline{I}_2 - \exp[-AA^*t])(AA^*)^+; \tag{3.2}$$

for example, since $A^*A(A^*A)^+ S_1 = \overline{S}_1$,

$$\overline{L}_1(t) = \int_0^t \exp\left[-A^*A(t-s)\right] A^*A(A^*A)^+ ds = (\overline{I}_1 - \exp[-A^*At])(A^*A)^+.$$

Formally, also, we can compute

(3.3) $$B(t) = A^* L_2(t) = A^+AA^* \bar{L}_2(t) = A^+ \bar{L}_2(t) AA^* = A^+ (\bar{I}_2 - \exp[-AA^* t]) = (\bar{I}_1 - \exp[-A^* At]) A^+.$$

The rest of this section deals with the strong convergence of $B(t)$ to A^+ as $t \to \infty$ and the convergence rates.

THEOREM 1.

(a) $\|(A^+ - B(t))y\|^2 \leq \dfrac{\|A^+y\|^2 \|(AA^*)^+y\|^2}{\|(AA^*)^+y\|^2 + 2\|A^+y\|^2 t}$ *if* $(y \in I)$ *and* $t \geq 0$.

(b) $\|(A^+ - B(t))y\|^2$ *is a decreasing function of* $t \geq 0$, *with limit zero as* $t \to \infty$, *if* $(y \in II)$.

(c) $\|(AA^+ - AB(t))y\|^2 \leq \dfrac{\|y\|^2 \|A^+y\|^2}{\|A^+y\|^2 + 2\|y\|^2 t}$ *if* $(y \in I)$ *or* $(y \in II)$, *and* $t \geq 0$.

(d) $\|(AA^+ - AB(t))y\|^2$ *is a decreasing function of* $t \geq 0$, *with limit zero as* $t \to \infty$, *if* $(y \in III)$.

Remarks: If $(y \in III)$, A^+y—and hence $\|(A^+ - B(t))y\|$—is undefined. In (c), (d), $AA^+ = \bar{I}_2$, e.g. (2.3).

Proof: Using the conventions of § 2, we define for $(y \in I)$ with $y = Ax = AA^* w$, and for $t \geq 0$

(3.4) $$w(t) = L_2(t) AA^* w \quad , \quad x(t) = A^* w(t) = A^* L_2(t) A (A^* w).$$

Thus

(3.5) $$x(t) = L_1(t) A^* Ax \quad , \quad y(t) = Ax(t) = AL_1(t) A^* Ax = L_2(t) AA^* Ax.$$

From which

(3.6) $$y(t) = L_2(t) AA^* y.$$

If $(y \in II)$ then w is undefined, and (3.5)—rather than (3.4)—is the defining equation for $x(t)$. Similarly, if $(y \in III)$, (3.6) is the defining equation for $y(t)$.

(3.7) From (3.1) and (2.5) it follows that the following relations hold whenever defined:

$$x(0) = y(0) = w(0) = 0 \quad , \quad w(t) = L_2(t) y,$$

$$x(t) = A^* w(t) = B(t) y \quad , \quad y(t) = Ax(t) = AA^* w(t) = AB(t) y.$$

Let

$$D_t = \frac{d}{dt}.$$

Then if $(y \in I)$:

(3.8) $$\begin{aligned} D_t \|w - w(t)\|^2 &= 2 \operatorname{Re}(w - w(t), -D_t w(t)) \\ &= 2 \operatorname{Re}(w - w(t), -D_t L_2(t) AA^* w) \\ &= 2 \operatorname{Re}(w - w(t), -AA^* (I - L_2(t) AA^*) w) \\ &= -2 \|A^* (w - w(t))\|^2. \end{aligned}$$

The identity for $D_t L_2(t)$ is proved from

$$D_t L_2(t) = D_t \int_0^t \int_0^\infty \exp[-\lambda(t-s)]\, dE_\lambda\, ds \qquad \text{where } A^* A = \int_0^\infty \lambda dE_\lambda .$$

By similar computations we obtain

$$(3.9) \qquad \begin{aligned} D_t \|x - x(t)\|^2 &= -2\| A(x - x(t)) \|^2 , \\ D_t \|y - y(t)\|^2 &= -2\| A^*(y - y(t))\|^2. \end{aligned}$$

We next derive a simple inequality, used in establishing the convergence rates given in this paper.

Let $(y \in I)$. Then

$$(3.10) \qquad \| x - x(t)\|^2 = (x - x(t), A^*(w - w(t)))$$

$$= (A(x - x(t)), w - w(t)) \leq \|A(x - x(t))\| \, \| w - w(t)\| .$$

From (3.8), $D_t \| w - w(t) \| \leq 0$ which implies

$$\| w - w(t)\| \leq \| w - w(0)\| = \| w\| , \qquad \forall t \geq 0 .$$

Therefore, if $(y \in I)$ and $t \geq 0$

$$(3.11) \qquad \|A(x - x(t))\| \geq \|x - x(t)\|^2 / \|w\| .$$

Similarly, if $t \geq 0$ and $(y \in I)$ or $(y \in II)$ then

$$(3.12) \qquad \| A^*(y - y(t))\| \geq \| y - y(t) \|^2 / \| x \| .$$

Now let $\beta(t) = \| x - x(t)\|^2$, $\forall t \geq 0$.

Then $\beta(0) = \| x\|^2$, $D_t \beta(t) = -2\| A(x - x(t))\|^2$ by (3.9), so (3.11) gives

$$(3.13) \qquad D_t \beta(t) \leq -2 \beta(t)^2 / \| w \|^2$$

Integrating (3.13) gives for $\forall t \geq 0$

$$\beta(t) \leq \frac{\| w \|^2 \beta(0)}{\| w \|^2 + 2t\beta(0)} = \frac{\| w \|^2 \| x \|^2}{\| w \|^2 + 2t \| x \|^2}$$

from which theorem 1 (*a*) follows by using (2.5), since

$$\beta(t) = \| x - x(t)\|^2 = \| A^+ y - B(t) y\|^2.$$

A similar derivation holds, using (3.9), (3.12), for theorem 1 (*c*).

Now suppose $(y \in II)$. We will establish theorem 1 (*b*). Since $A^+ y - B(t) y = x - x(t)$, (3.9) establishes that $\|(A^+ - B(t)) y\|$ is non increasing, and we need only establish the limit. Since $(y \in II)$ there is an $x = A^+ y \in \bar{S}_1 = cl(A^* S_2)$. Let $\varepsilon > 0$. Then there is a $w_\varepsilon \in \bar{S}_2$ such that $\| x - A^* w_\varepsilon\| < \frac{\varepsilon}{2}$. Define

$$(3.14) \qquad x_\varepsilon = A^* w_\varepsilon \quad , \quad y_\varepsilon = A x_\varepsilon = AA^* w_\varepsilon .$$

Then $(y_\varepsilon \in \mathrm{I})$, so there is $\tau > 0$ such that $t \geq \tau \Rightarrow \|(\mathrm{A}^+ - \mathrm{B}(t))\, y_\varepsilon\| < \frac{\varepsilon}{2}$, by theorem 1 (*a*). Also $((y - y_\varepsilon) \in \mathrm{II})$, so that $\|(\mathrm{A}^+ - \mathrm{B}(t))(y - y_\varepsilon)\|$ is nonincreasing in t.

This is shown as in (3.9):

$$\mathrm{D}_t \|(\mathrm{A}^+ - \mathrm{B}(t))(y - y_\varepsilon)\|^2 = -2\,\|\mathrm{A}(\mathrm{A}^+ - \mathrm{B}(t))(y - y_\varepsilon)\|^2 .$$

Therefore for $\forall t \geq 0$

$$\|(\mathrm{A}^+ - \mathrm{B}(t))(y - y_\varepsilon)\| \leq \|(\mathrm{A}^+ - \mathrm{B}(0))(y - y_\varepsilon)\| = \|x - x_\varepsilon\| < \frac{\varepsilon}{2},$$

by (3.14) since $\mathrm{B}(0) = 0$.

Consequently, $t > \tau \Rightarrow$

$$\|(\mathrm{A}^+ - \mathrm{B}(t))\, y\| \leq \|(\mathrm{A}^+ - \mathrm{B}(t))\, y_\varepsilon\| + \|(\mathrm{A}^+ - \mathrm{B}(t))(y - y_\varepsilon)\| < \varepsilon ,$$

establishing theorem 1 (*b*).

Theorem 1 (*d*) is shown by a similar procedure.

Remarks: The strong convergences

$$\bar{\mathrm{L}}_1(t) \to (\mathrm{A}^*\mathrm{A})^+ \quad \text{as} \quad t \to \infty \qquad \text{and} \qquad \bar{\mathrm{L}}_2(t) \to (\mathrm{A}\mathrm{A}^*)^+ \quad \text{as} \quad t \to \infty ,$$

with convergence rates analogous to those in theorem 1, can be similarly established.

Some well known properties of g. i., follow from the above results:

$$\mathrm{A}^+ = \lim_{t \to \infty} \mathrm{B}(t) = \lim_{t \to \infty} \bar{\mathrm{L}}_1(t)\,\mathrm{A}^* = (\mathrm{A}^*\mathrm{A})^+\mathrm{A}^*$$

and similarly:

$$\mathrm{A}^+ = \mathrm{A}^*(\mathrm{A}\mathrm{A}^*)^+ \quad , \quad \mathrm{A}^{*+} = \mathrm{A}^{+*} = \mathrm{A}(\mathrm{A}^*\mathrm{A})^+ = (\mathrm{A}^*\mathrm{A})^+\mathrm{A} ,$$

$$\mathrm{A}^{*+}\mathrm{A}^+ = (\mathrm{A}\mathrm{A}^*)^+ \quad , \quad \mathrm{A}^+\mathrm{A}^{*+} = (\mathrm{A}^*\mathrm{A})^+ .$$

§ 4. SERIES REPRESENTATION OF A^+.

Let c be a real number, $0 < c < 2$, considered fixed throughout this section. Let $\alpha = c/\|\mathrm{A}\|^2$, and let the sequence $\{\mathrm{B}_\mathrm{N} : \mathrm{N} = 0, 1, \cdots\}$ be defined by:

$$\mathrm{B}_\mathrm{N} = \sum_{k=0}^{\mathrm{N}} (\mathrm{I}_1 - \alpha \mathrm{A}^*\mathrm{A})^k \alpha \mathrm{A}^* . \tag{4.1}$$

Suppose $(y \in \mathrm{I}) : y = \mathrm{A}x = \mathrm{A}\mathrm{A}^* w$, $w \in \bar{\mathrm{S}}_2$. Let the sequences $\{y_\mathrm{N}\}$, $\{x_\mathrm{N}\}$, $\{w_\mathrm{N}\}$ be defined as follows:

$$(y - y_{\mathrm{N}+1}) = (\mathrm{I}_2 - \alpha \mathrm{A}\mathrm{A}^*)(y - y_\mathrm{N}) \tag{4.2}$$

$$(x - x_{\mathrm{N}+1}) = (\mathrm{I}_1 - \alpha \mathrm{A}^*\mathrm{A})(x - x_\mathrm{N}) \tag{4.3}$$

$$(w - w_{\mathrm{N}+1}) = (\mathrm{I}_2 - \alpha \mathrm{A}\mathrm{A}^*)(w - w_\mathrm{N}) \qquad (\mathrm{N} = 0, 1, \cdots) \tag{4.4}$$

with $y_0 = x_0 = w_0 = 0$.

For $(y \in \mathrm{II})$, (4.2) and (4.3) hold, but not (4.4) since w is then undefined.

Similarly, for $(y \in \mathrm{III})$ only (4.2) holds.

The sequence $\{x_N\}$ is, by (4.3) and (4.1):

$$x_{N+1} = x_N - \alpha A^* (A x_N - y) = (I_1 - \alpha A^* A) x_N + \alpha A^* y \tag{4.5}$$

$$= \left[\sum_{k=0}^{N} (I_1 - \alpha A^* A)^k \alpha A^* \right] y = B_N y \qquad (N = 0, 1, \cdots).$$

It will now be shown that $\lim_{N\to\infty} x_N = x$ i.e. $\lim_{N\to\infty} B_N = A^+$, and convergence rates analogous to theorem 1 will be established. For $N = 0, 1, \cdots$

$$\| y_{N+1} - y \|^2 = \| (I_2 - \alpha A A^*) (y_N - y) \|^2$$

$$= \| y_N - y \|^2 - 2\alpha \| A^* (y_N - y) \|^2 + \alpha^2 \| A A^* (y_N - y) \|^2 ;$$

therefore

$$\| y_{N+1} - y \|^2 \le \| y_N - y \|^2 - \frac{(2-c)\,c}{\| A \|^2} \| A^* (y_N - y) \|^2 \tag{4.6}$$

and similarly

$$\| x_{N+1} - x \|^2 \le \| x_N - x \|^2 - \frac{(2-c)\,c}{\| A \|^2} \| A (x_N - x) \|^2 \quad \text{if } (y \in \mathrm{I}, \mathrm{II}), \tag{4.7}$$

$$\| w_{N+1} - w \|^2 \le \| w_N - w \|^2 - \frac{(2-c)\,c}{\| A \|^2} \| A^* (w_N - w) \|^2 \quad \text{if } (y \in \mathrm{I}). \tag{4.8}$$

The terms on the left of (4.6), (4.7), (4.8) clearly denote nonincreasing sequences. Accordingly if $(y \in \mathrm{I})$

$$\| A (x_N - x) \|^2 \, \| w \|^2 \ge \| A (x_N - x) \|^2 \, \| w_N - w \|^2 , \qquad \forall N$$

$$\ge | (A (x_N - x) , w_N - w) |^2 = | (x_N - x , A^* (w_N - w) |^2 = \| x_N - x \|^4$$

or

$$\| A (x_N - x) \|^2 \ge \| x_N - x \|^4 / \| (A A^*)^+ y \|^2 \tag{4.9}$$

Similarly, if $(y \in \mathrm{I})$ or $(y \in \mathrm{II})$

$$\| A^* (y_N - y) \|^2 \ge \| y_N - y \|^4 / \| A^+ y \|^2 . \tag{4.10}$$

We also need an inequality:

(4.11) If $\{r_N : N = 0, 1, \cdots\}$ is a sequence of nonnegative numbers, and $h > 0$ is a constant with $r_0 h \le 1$ and

$$r_{N+1} \le r_N - h r_N^2 , \qquad \forall N \ge 0 ,$$

then:

$$r_N \le (1 + N h r_0)^{-1} r_0 , \qquad \forall N > 0 .$$

This is proved by induction:

$$r_0 \le r_0 . \quad \text{If} \quad r_N \le (1 + N h r_0)^{-1} r_0 = (r_0^{-1} + N h)^{-1}$$

then

$$1 \ge (1 - h r_N)(1 + h r_N) \ge (1 - h r_N)([r_0^{-1} + N h] r_N + h r_N)$$

$$= (r_N - h r_N^2)(r_0^{-1} + (N+1) h) \ge r_{N+1} (r_0^{-1} + (N+1) h) ,$$

(proving 4.11).

We have now the machinery to prove the following discrete analog of theorem 1.

THEOREM 2.

(a) $\|(A^+ - B_N)y\|^2 \leq \dfrac{\|A^+y\|^2 \, \|(AA^*)^+y\|^2}{\|(AA^*)^+y\|^2 + N \dfrac{(2-c)\,c}{\|A\|^2} \|A^+y\|^2}$ *if* $(y \in I)$ *and* $N = 1, 2, \cdots$.

(b) $\|(A^+ - B_N)\, y\|^2 = \|x - x_N\|^2$ *converges monotonically to zero if* $(y \in II)$.

(c) $\|(AA^+ - AB_N)y\|^2 \leq \dfrac{\|y\|^2 \, \|A^+y\|^2}{\|A^+y\|^2 + N \dfrac{(2-c)\,c}{\|A\|^2} \|y\|^2}$ *if* $(y \in I)$ *or* $(y \in II)$ *and* $N = 1, 2, \cdots$.

(d) $\|(AA^+ - AB_N)y\|^2 = \|y - y_N\|^2$ *converges monotonically to zero if* $(y \in III)$.

Proof: From (4.7) and (4.9) we have

$$\|x_{N+1} - x\|^2 \leq \|x_N - x\|^2 - \frac{(2-c)\,c}{\|A\|^2} \frac{\|x_N - x\|^4}{\|(AA^*)^+y\|^2}, \qquad \forall N > 0 .$$

Now let

$$r_0 = \|x\|^2 = \|A^+y\|^2 \quad , \quad r_N = \|x_N - x\|^2 = \|(B_N - A^+)\, y\|^2 ,$$

and $h = (2-c)\, c / \|A\|^2 \, \|(AA^*)^+ y\|^2$. Then $r_0 h \leq 1$ since $\|x_1 - x\| \geq 0$, and the inequality (4.11) is applicable to prove theorem 2(a). The proof of theorem 2(c) is similar with

$$r_0 = \|y\|^2 = \|AA^+y\|^2 \quad , \quad r_N = \|y_N - y\|^2 = \|(AB_N - AA^+)y\|^2,$$

and $h = (2-c)\, c/\|A\|^2 \, \|A^+y\|^2$ and by using (4.6), (4.10) and the inequality (4.11). The proofs of theorem 2(b) and (d) are analogous to the proofs of theorem 1(b) and (d) respectively.

Remark: These results, on the strong convergence $B_N \to A^+$, are different than the results in [7] where the convergence $B_N \to A^+$ is in the uniform operator topology, restricting A to have closed range, e.g. [7] theorem 3.

§ 5. HYPERPOWER ITERATIVE METHODS AND PROJECTION MAPPINGS.

The notation of § 4 is adhered to in this section. The sequence (4.1) is first shown to satisfy

(5.1)
$$\begin{aligned} I_2 - AB_N &= I_2 - \alpha AA^* \sum_{k=0}^{N} (I_2 - \alpha AA^*)^k \\ &= I_2 - [I_2 - (I_2 - \alpha AA^*)] \sum_{k=0}^{N} (I - \alpha AA^*)^k \\ &= I_2 - [I_2 - (I_2 - \alpha AA^*)^{N+1}] = (I_2 - \alpha AA^*)^{N+1}, \quad N = 0, 1, \cdots . \end{aligned}$$

If $p \geq 2$ is an integer, we define the sequence $\{C_{N,p} : N = 0, 1, \cdots\}$ by recursion

(5.2)
$$C_{0,p} = \alpha A^* \quad , \quad C_{N+1,p} = C_{N,p} \sum_{k=0}^{p-1} (I_2 - AC_{N,p})^k .$$

We also define

$$(5.3) \qquad D_{0,p} = \alpha A^* \quad , \quad D_{N+1,p} = D_{N,p} \sum_{k=1}^{p} \binom{p}{k} (-AD_{N,p})^{k-1}.$$

The sequence (5.2) has been studied in [7] where its uniform convergence to A^+ was established for bounded A with closed range, and for matrices in [3], [9]. The sequence (5.3) is shown in (5.7) below to be identical with (5.2), both sequences converging strongly to A^+. The series (5.2) is somewhat simpler to use if the term $(I - AC_{N,p})^k$ can be evaluated by only $k - 1$ operator multiplications, e.g. for matrices. The form (5.3) is preferable otherwise, e.g. for integral transforms. The sequence (5.2) satisfies for $N = 0, 1, \cdots$

$$(5.4) \qquad I_2 - AC_{N+1,p} = I_2 - AC_{N,p} \sum_{k=0}^{p-1} (I_2 - AC_{N,p})^k$$

$$= I_2 - [I_2 - (I_2 - AC_{N,p})] \sum_{k=0}^{p-1} (I_2 - AC_{N,p})^k = (I_2 - AC_{N,p})^p, \text{ as in (5.1)}$$

$$= (I_2 - AC_{0,p})^{p^{N+1}} = (I_2 - \alpha AA^*)^{p^{N+1}} = I_2 - AB_{(p^{N+1}-1)}, \text{ by induction.}$$

Similarly for the sequence (5.3):

$$(5.5) \qquad I_2 - AD_{N+1,p} = \sum_{k=0}^{p} \binom{p}{k} (-AD_{N,p})^k = (I_2 - AD_{N,p})^p$$

$$= (I_2 - AD_{0,p})^{p^{N+1}} = I_2 - AB_{(p^{N+1}-1)}, \text{ by induction}$$

From (5.4) and (5.5) it follows that

$$(5.6) \qquad AB_{(p^{N+1}-1)} = AC_{N+1,p} = AD_{N+1,p}.$$

Since for all integers $p, q > 0$, $A^+ AB_p = B_p$; $A^+ AC_{q,p} = C_{q,p}$; and $A^+AD_{q,p} = D_{q,p}$, we multiply each term in (5.6) by A^+ to obtain:

$$(5.7) \qquad B_{(p^{N+1}-1)} = C_{N+1,p} = D_{N+1,p}.$$

Consequently $\{C_{N,p}\}$ and $\{D_{N,p}\}$ $(N = 0, 1, \cdots)$ are hypergeometric series for $p \geq 2$, with the convergence rates established in theorem 2, e.g.

$$(5.8) \qquad \| (A^+ - C_{N,p}) y \|^2 \leq \frac{\| A^+ y \|^2 \| (AA^*)^+ y \|^2}{\| (AA^*)^+ y \|^2 + (p^N - 1) \frac{(2-c)c}{\| A \|^2} \| A^+ y \|^2}$$

if $(y \in I)$ and $N = 1, 2, \cdots$.

We return now to consider the operator $I_2 - AB_N$, (5.1). For any $y \in S_2$, $y = \bar{y} + \bar{y}^{\perp}$, it was established in § 4 that

$$(5.9) \qquad \lim_{N \to \infty} (I_2 - AB_N) \bar{y} = 0,$$

with the rate of convergence varying according as $(y \in I)$, $(y \in II)$ or $(y \in III)$.

It is clear that $B_N \bar{y}^{\perp} = 0$, $\forall N$, since $A^* \bar{y}^{\perp} = 0$. (Thus for arbitrary $y \in S_2$, $\lim_{N\to\infty} (AA^+ - AB_N) y = 0$, which was established in § 4). Consequently, $\lim_{N\to\infty} (I_2 - AB_N) y = \bar{y}^{\perp}$, or by (5.1) $\lim_{N\to\infty} (I_2 - \alpha AA^*)^N y = \bar{y}^{\perp}$, and so, $\lim_{N\to\infty} (I_2 - \alpha AA^*)^N$ is the projection onto the null space, $\bar{S}_2^{\perp}$, of A^*, with previously established convergence rates. Clearly then

$$\lim_{N\to\infty} [I_2 - (I_2 - \alpha AA^*)^N] = P_{\bar{S}_2} \tag{5.10}$$

where $P_{\bar{S}_2}$ is the projection mapping onto $\bar{S}_2$.

The following analog of (5.1)

$$I_1 - B_N A = (I_1 - \alpha A^* A)^{N+1} \tag{5.11}$$

similarly gives for any $x \in S_1$, $x = \bar{x} + \bar{x}^{\perp}$

$$(I_1 - \alpha A^* A)^{N+1} x = \bar{x}^{\perp} + (A^+ - B_N) y , \quad \forall N \tag{5.12}$$

where $y = A\bar{x}$, so that $(y \in \mathrm{I})$ or $(y \in \mathrm{II})$. By theorem 2, (5.12) converges to $\bar{x}^{\perp}$. Thus

$$\lim_{N\to\infty} [I_1 - (I_1 - \alpha A^* A)^N] = P_{\bar{S}_1} . \tag{5.13}$$

These results suggest hyperpower iterative methods for computing the projection mappings $P_{\bar{S}_1}$, $P_{\bar{S}_1^{\perp}}$, $P_{\bar{S}_2}$, $P_{\bar{S}_2^{\perp}}$ based on (5.2) or (5.3), with convergence rates given from (5.7) and theorem 2. These methods were studied in [3], [7] and [9]. It should however be noted that the uniform convergence $(I_1 - \alpha A^* A)^N \to P_{\bar{S}_1^{\perp}}$ is equivalent to the range of A being closed, see [7] corollary 3, which emphasizes the wider applicability of our results involving strong operator convergence.

Acknowledgment: Research supported by the National Science Foundation, Project GP 13546.

REFERENCES.

[1] E. ARGHIRIADE, *Sur l'inverse généralisée d'un opérateur linéaire, dans les espaces de Hilbert*, « Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. », ser. VIII, Vol. XLVI (1969).

[2] A. BEN-ISRAEL and A. CHARNES, *Contributions to the theory of generalized inverses*, « SIAM J. » *11*, 667–699 (1963).

[3] A. BEN-ISRAEL and D. COHEN, *On iterative computation of generalized inverses and associated projections*, « J. SIAM Numer. Anal. », *3*, 410–419 (1966).

[4] F. J. BEUTLER, *The operator theory of the pseudo-inverse*, I. *Bounded operators*, II. *Unbounded operators with arbitrary range*, « J. Math. Anal. Appl. » *10*, 451–470, 471–493 (1965).

[5] C. A. DESOER and B. H. WHALEN, *A note on pseudoinverses*, « SIAM J. », *11*, 442–447 (1963).
[6] M. R. HESTENES, *Relative self–adjoint operators in Hilbert space*, « Pacific J. Math. », *11*, 1315–1357 (1961).
[7] W. V. PETRYSHYN, *On generalized inverses and on the uniform convergence of* $(I - \beta K)^n$ *with application to iterative methods*, « J. Math. Anal. Appl. », *18*, 417–439 (1967).
[8] D. SHOWALTER, *Representation and computation of the pseudoinverse*, « Proc. Amer. Math. Soc. », *18*, 584–586 (1967).
[9] S. ZLOBEC, *On computing the generalized inverse of a linear operator*, « Glasnik Mat. », *2*, (22), 265–271 (1967).