Hindawi Publishing Corporation
Fixed Point Theory and Applications
Volume 2010, Article ID 821928, 13 pages
doi:10.1155/2010/821928

Research Article

Nonexpansive Matrices with Applications to Solutions of Linear Systems by Fixed Point Iterations

Teck-Cheong Lim

Department of Mathematical Sciences, George Mason University, 4400, University Drive, Fairfax, VA 22030, USA

Correspondence should be addressed to Teck-Cheong Lim, tlim@gmu.edu

Received 28 August 2009; Accepted 19 October 2009

Academic Editor: Mohamed A. Khamsi

We characterize (i) matrices which are nonexpansive with respect to some matrix norms, and (ii) matrices whose average iterates approach zero or are bounded. Then we apply these results to iterative solutions of a system of linear equations.

Throughout this paper, $\mathbb{R}$ will denote the set of real numbers, $\mathbb{C}$ the set of complex numbers, and M_n the complex vector space of complex $n \times n$ matrices. A function $\|\cdot\| : M_n \to \mathbb{R}$ is a *matrix norm* if for all $A, B \in M_n$, it satisfies the following five axioms:

(1) $\|A\| \geq 0$;

(2) $\|A\| = 0$ if and only if $A = 0$;

(3) $\|cA\| = |c|\|A\|$ for all complex scalars c;

(4) $\|A + B\| \leq \|A\| + \|B\|$;

(5) $\|AB\| \leq \|A\|\ \|B\|$.

Let $|\cdot|$ be a norm on $\mathbb{C}^n$. Define $\|\cdot\|$ on M_n by

$$\|A\| = \max_{|x|=1} |Ax|. \tag{1}$$

This norm on M_n is a matrix norm, called the *matrix norm induced by* $|\cdot|$. A matrix norm on M_n is called an *induced matrix norm* if it is induced by some norm on $\mathbb{C}^n$. If $\|\cdot\|_1$ is a matrix norm on M_n, there exists an induced matrix norm $\|\cdot\|_2$ on M_n such that $\|A\|_2 \leq \|A\|_1$ for all

$A \in M_n$ (cf. [1, page 297]). Indeed one can take $\|\cdot\|_2$ to be the matrix norm induced by the norm $|\cdot|$ on $\mathbb{C}^n$ defined by

$$|x| = \|C(x)\|_1, \tag{2}$$

where $C(x)$ is the matrix in M_n whose columns are all equal to x. For $A \in M_n$, $\rho(A)$ denotes the spectral radius of A.

Let $|\cdot|$ be a norm in $\mathbb{C}^n$. A matrix $A \in M_n$ is a *contraction* relative to $|\cdot|$ if it is a contraction as a transformation from $\mathbb{C}^n$ into $\mathbb{C}^n$; that is, there exists $0 \leq \lambda < 1$ such that

$$\left|Ax - Ay\right| \leq \lambda\left|x - y\right|, \quad x, y \in \mathbb{C}^n. \tag{3}$$

Evidently this means that for the matrix norm $\|\cdot\|$ induced by $|\cdot|$, $\|A\| < 1$. The following theorem is well known (cf. [1, Sections 5.6.9–5.6.12]).

Theorem 1. *For a matrix $A \in M_n$, the following are equivalent:*

(a) *A is a contraction relative to a norm in $\mathbb{C}^n$;*

(b) $\|A\| < 1$ *for some induced matrix norm* $\|\cdot\|$;

(c) $\|A\| < 1$ *for some matrix norm* $\|\cdot\|$;

(d) $\lim_{k\to\infty} A^k = 0$;

(e) $\rho(A) < 1$.

That (b) follows from (c) is a consequence of the previous remark about an induced matrix norm being less than a matrix norm. Since all norms on M_n are equivalent, the limit in (d) can be relative to any norm on M_n, so that (d) is equivalent to all the entries of A^k converge to zero as $k \to \infty$, which in turn is equivalent to $\lim_{k\to\infty} A^k x = 0$ for all $x \in \mathbb{C}^n$.

In this paper, we first characterize matrices in M_n that are *nonexpansive* relative to some norm $|\cdot|$ on $\mathbb{C}^n$, that is,

$$\left|Ax - Ay\right| \leq \left|x - y\right|, \quad x, y \in \mathbb{C}^n. \tag{4}$$

Then we characterize those $A \in M_n$ such that

$$A_k = \frac{1}{k}\left(I + A + A^2 + \cdots + A^{k-1}\right) \tag{5}$$

converges to zero as $k \to \infty$, and those that $\{A_k : k = 0, 1, 2, \ldots\}$ is bounded.

Finally we apply our theory to approximation of solution of $Ax = b$ using iterative methods (fixed point iteration methods).

Theorem 2. *For a matrix $A \in M_n$, the following are equivalent:*

(a) *A is nonexpansive relative to some norm on $\mathbb{C}^n$;*

(b) *$\|A\| \leq 1$ for some induced matrix norm $\|\cdot\|$;*

(c) *$\|A\| \leq 1$ for some matrix norm $\|\cdot\|$;*

(d) *$\{A^k : k = 0, 1, 2, \ldots\}$ is bounded;*

(e) *$\rho(A) \leq 1$, and for any eigenvalue λ of A with $|\lambda| = 1$, the geometric multiplicity is equal to the algebraic multiplicity.*

Proof. As in the previous theorem, (a), (b), and (c) are equivalent. Assume that (b) holds. Let the norm $\|\cdot\|$ be induced by a vector norm $|\cdot|$ of $\mathbb{C}^n$. Then

$$\left|A^k(x)\right| \leq \left\|A^k\right\| |x| \leq \|A\|^k |x| \leq |x|, \quad k = 0, 1, 2, \ldots, \tag{6}$$

proving that $A^k(x)$ is bounded in norm $|\cdot|$ for every $x \in \mathbb{C}^n$. Taking $x = e_i$, we see that the set of all columns of $A^k, k = 0, 1, 2, \ldots$, is bounded. This proves that $A^k, k = 0, 1, 2, \ldots$, is bounded in maximum column sum matrix norm ([1, page 294]), and hence in any norm in M_n. Note that the last part of the proof also follows from the Uniform Boundedness Principle (see, e.g., [2, Corollary 21, page 66])

Now we prove that (d) implies (e). Suppose that A has an eigenvalue λ with $\lambda > 1$. Let x be an eigenvector corresponding to λ. Then

$$\left\|A^k x\right\| = |\lambda|^k \|x\| \longrightarrow \infty \tag{7}$$

as $k \to \infty$, where $\|\cdot\|$ is any vector norm of $\mathbb{C}^n$. This contradicts (d). Hence $|\lambda| \leq 1$. Now suppose that λ is an eigenvalue with $|\lambda| = 1$ and the Jordan block corresponding to λ is not diagonal. Then there exist nonzero vectors v_1, v_2 such that $Av_1 = \lambda v_1$, $A(v_2) = v_1 + \lambda v_2$. Let $u = v_1 + v_2$. Then

$$A^k u = \lambda^{k-1}(\lambda + k)v_1 + \lambda^k v_2, \tag{8}$$

and $\|A^k(u)\| \geq k\|v_1\| - \|v_1\| - \|v_2\|$. It follows that $A^k u, k = 0, 1, 2, \ldots$, is unbounded, contradicting (d). Hence (d) implies (e).

Lastly we prove that (e) implies (c). Assume that (e) holds. A is similar to its Jordan canonical form J whose nonzero off-diagonal entries can be made arbitrarily small by similarity ([1, page 128]). Since the Jordan block for each eigenvalue with modulus 1 is diagonal, we see that there is an invertible matrix S such that the l_1-sum of each row of SAS^{-1} is less than or equal to 1, that is, $\|SAS^{-1}\|_\infty \leq 1$, where $\|\cdot\|_\infty$ is the maximum row sum matrix norm ([1, page 295]). Define a matrix norm $\|\cdot\|$ by $\|M\| = \|SMS^{-1}\|_\infty$. Then we have $\|A\| \leq 1$. □

Let λ be an eigenvalue of a matrix $A \in M_n$. The index of λ, denoted by index(λ) is the smallest value of k for which $\operatorname{rank}(A - \lambda I)^k = \operatorname{rank}(A - \lambda I)^{k+1}$ ([1, pages 148 and 131]). Thus condition (e) above can be restated as $\rho(A) \leq 1$, and for any eigenvalue λ of A with $|\lambda| = 1$, index$(\lambda) = 1$.

Let $A \in M_n$. Consider

$$A_k = \frac{1}{k}\left(I + A + \cdots + A^{k-1}\right). \tag{9}$$

We call A_k the k-average of A. As with A^k, we have $A_k x \to 0$ for every x if and only if $A_k \to 0$ in M_n, and that $A_k x$ is bounded for every x if and only if A_k is bounded in M_n. We have the following theorem.

Theorem 3. *Let $A \in M_n$. Then*

(a) $A_k, k = 1, 2, \ldots,$ *converges to* 0 *if and only if* $\|A\| \leq 1$ *for some matrix norm* $\|\cdot\|$ *and that* 1 *is not an eigenvalue of* A,

(b) $A_k, k = 1, 2, \ldots,$ *is bounded if and only if* $\rho(A) \leq 1, \mathit{index}(\lambda) \leq 2$ *for every eigenvalue* λ *with* $|\lambda| = 1$ *and that* $\mathit{index}(1) = 1$ *if* 1 *is an eigenvalue of* A.

Proof. First we prove the sufficiency part of (a). Let x be a vector in $\mathbb{C}^n$. Let

$$y_k = \frac{1}{k}\left(I + A + \cdots + A^{k-1}\right)(x). \tag{10}$$

By Theorem 2 for any eigenvalues λ of A either $|\lambda| < 1$ or $|\lambda| = 1$ and $\text{index}(\lambda) = 1$.

If A is written in its Jordan canonical form $A = SJS^{-1}$, then the k-average of A is $SJ'S^{-1}$, where J' is the k-average of J. J' is in turn composed of the k-average of each of its Jordan blocks. For a Jordan block of J of the form

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda \end{pmatrix}, \tag{11}$$

$|\lambda|$ must be less than 1. Its k-average has constant diagonal and upper diagonals. Let D_j be the constat value of its jth upper diagonal (D_0 being the diagonal) and let $S_j = kD_j$. Then ($C(m,n) = 0$ for $n > m$)

$$\begin{gathered} S_0 = \frac{1-\lambda^k}{1-\lambda}, \\ S_j = C(j,j) + C(j+1,j)\lambda + \cdots + C(k-1,j)\lambda^{k-1-j}, \quad j = 1, 2, \ldots, n-1. \end{gathered} \tag{12}$$

Using the relation $C(m+1,j) - C(m,j) = C(m,j-1)$, we obtain

$$S_j - \lambda S_j = S_{j-1} - \lambda^{k-j} C(k,j). \tag{13}$$

Thus, we have $S_0 \to 1/(1-\lambda)$ as $k \to \infty$. By induction, using (13) above and the fact that $\lambda^{k-j}C(k,j) \to 0$ as $k \to \infty$, we obtain $S_j \to 1/(1-\lambda)^{j+1}$ as $k \to \infty$. Therefore $D_j = S_j/k = O(1/k)$ as $k \to \infty$.

If the Jordan block is diagonal of constant value λ, then $\lambda \neq 1, |\lambda| \leq 1$ and the k-average of the block is diagonal of constant value $(1-\lambda^k)/k(1-\lambda) = O(1/k)$.

We conclude that $\|A_k\| = O(1/k)$ and hence $\|y_k\| \leq \|A_k\|\|x\| = O(1/k)$ as $k \to \infty$.

Now we prove the necessity part of (a). If 1 is an eigenvalue of A and x is a corresponding eigenvector, then $A_k x = x \neq 0$ for every k and of course $B_k x$ fails to converge to 0. If λ is an eigenvalue of A with $|\lambda| > 1$ and x is a corresponding eigenvector, then

$$\|A_k x\| = \left|\frac{\lambda^k - 1}{k(\lambda - 1)}\right| \|x\| \geq \frac{|\lambda|^k - 1}{k|\lambda - 1|} \|x\|. \tag{14}$$

which approaches to ∞ as $k \to \infty$. If λ is an eigenvalue of A with $|\lambda| = 1$, $\lambda \neq 1$, and $\operatorname{index}(\lambda) \geq 2$, then there exist nonzero vectors v_1, v_2 such that $A(v_1) = \lambda v_1, A(v_2) = v_1 + \lambda v_2$. Then by using the identity

$$1 + 2\lambda + 3\lambda^2 + \cdots + (k-1)\lambda^{k-2} = \frac{1-\lambda^{k-1}}{(1-\lambda)^2} - (k-1)\frac{\lambda^{k-1}}{1-\lambda} \tag{15}$$

we get

$$A_k(v_2) = \left(\frac{1-\lambda^{k-1}}{k(1-\lambda)^2} - \left(1 - \frac{1}{k}\right)\frac{\lambda^{k-1}}{1-\lambda}\right)v_1 + \frac{1-\lambda^k}{k(1-\lambda)}v_2. \tag{16}$$

It follows that $\lim_{k\to\infty} A_k(v_2)$ does not exist. This completes the proof of part (a).

Suppose that A satisfies the conditions in (b) and that $A = SJS^{-1}$ is the Jordan canonical form of A. Let λ be an eigenvalue of A and let v be a column vector of S corresponding to λ. If $|\lambda| < 1$, then the restriction B of A to the subspace spanned by $v, Av, A^2v, \ldots$ is a contraction, and we have $\|A_k v\| = \|B_k v\| \leq \|v\|$. If $|\lambda| = 1$, and $\lambda \neq 1$, then by conditions in (b) either $Av = \lambda v$, or there exist v_1, v_2 with $v = v_2$ such that $A(v_1) = \lambda v_1, A(v_2) = v_1 + \lambda v_2$. In the former case, we have $\|A_k\| \leq \|v\|$ and in the latter case, we see from (16) that $A_k(v) = A_k(v_2)$ is bounded. Finally if $\lambda = 1$ then since $\operatorname{index}(1) = 1$, we have $Av = v$ and hence $A_k v = v$. In all cases, we proved that $A_k v, k = 0, 1, 2, \ldots,$ is bounded. Since column vectors of S form a basis for $\mathbb{C}^n$, the sufficiency part of (b) follows.

Now we prove the necessity part of (b). If A has an eigenvalue λ with $|\lambda| > 1$ and eigenvector v, then as shown above $A_k(v) \to \infty$ as $k \to \infty$. If A has 1 as an eigenvalue and $\operatorname{index}(1) \geq 2$, then there exist nonzero vectors v_1, v_2 such that $Av_1 = v_1$ and $Av_2 = v_1 + v_2$. Then $A_k(v_2) = ((k-1)/2) + v_2$ which is unbounded. If λ is an eigenvalue of A with $|\lambda| = 1, \lambda \neq 1$ and $\operatorname{index}(\lambda) \geq 3$, then there exist nonzero vectors v_1, v_2 and v_3 such that $Av_1 = \lambda v_1, A(v_2) = v_1 + \lambda v_2$ and $A(v_3) = v_2 + \lambda v_3$. By expanding $A^j(v_3), j = 0, 1, 2, \ldots, k-1$ and using the identity

$$\sum_{j=2}^{k-1} C(j,2)\lambda^{j-2} = \frac{1}{(1-\lambda)^2}\left(\frac{1-\lambda^{k-2}}{1-\lambda} + \frac{1}{2}(k-2)\lambda^{k-2}((k-1)\lambda - (k+1))\right), \tag{17}$$

we obtain

$$A_k(v_3) = \frac{1}{(1-\lambda)^2}\left(\frac{1-\lambda^{k-2}}{k(1-\lambda)} + \frac{1}{2}(k-2)\lambda^{k-2}\left(\frac{k-1}{k}\lambda - \frac{k+1}{k}\right)\right)v_1 + \left(\frac{1-\lambda^{k-1}}{k(1-\lambda)^2} - \left(1-\frac{1}{k}\right)\frac{\lambda^{k-1}}{1-\lambda}\right)v_2 + \frac{1-\lambda^k}{k(1-\lambda)}v_3 \tag{18}$$

which approaches to ∞ as $k \to \infty$. This completes the proof. □

We now consider applications of preceding theorems to approximation of solution of a linear system $Ax = b$, where $A \in M_n$ and b a given vector in $\mathbb{C}^n$. Let Q be a given invertible matrix in M_n. x is a solution of $Ax = b$ if and only if x is a fixed point of the mapping T defined by

$$Tx = \left(I - Q^{-1}A\right)x + Q^{-1}b. \tag{19}$$

T is a contraction if and only if $I - Q^{-1}A$ is. In this case, by the well known Contraction Mapping Theorem, given any initial vector x_0, the sequence of iterates $x_k = T^k x_0, k = 0, 1, 2, \ldots$, converges to the unique solution of $Ax = b$. In practice, given x_0, each successive x_k is obtained from x_{k-1} by solving the equation

$$Q(x_k) = (Q - A)x_{k-1} + b. \tag{20}$$

The classical methods of Richardson, Jacobi, and Gauss-Seidel (see, e.g., [3]) have $Q = I, D$, and L respectively, where I is the identity matrix, D the diagonal matrix containing the diagonal of A, and L the lower triangular matrix containing the lower triangular portion of A. Thus by Theorem 1 we have the following known theorem.

Theorem 4. *Let $A, Q \in M_n$, with Q invertible. Let $b, x_0 \in \mathbb{C}^n$. If $\rho(I - Q^{-1}A) < 1$, then A is invertible and the sequence $x_k, k = 1, 2, \ldots$, defined recursively by*

$$Q(x_k) = (Q - A)x_{k-1} + b \tag{21}$$

converges to the unique solution of $Ax = b$.

Theorem 4 fails if $\rho(I - Q^{-1}A) = 1$, For a simple 2×2 example, let $Q = I, b = 0, A = 2I$ and x_0 any nonzero vector.

We need the following lemma in the proof of the next two theorems. For a matrix $A \in M_n$, we will denote $R(A)$ and $N(A)$ the range and the null space of A respectively.

Lemma 5. *Let A be a singular matrix in M_n such that the geometric multiplicity and the algebraic multiplicity of the eigenvalue 0 are equal, that is,* index(0) = 1. *Then there is a unique projection P_A whose range is the range of A and whose null space is the null space of A, or equivalently, $\mathbb{C}^n = R(A) \oplus N(A)$. Moreover, A restricted to $R(A)$ is an invertible transformation from $R(A)$ onto $R(A)$.*

Proof. If $A = SJS^{-1}$ is a Jordan canonical form of A where the eigenvalues 0 appear at the end portion of the diagonal of J, then the matrix

$$P_A = S \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \cdot & \\ & & & & & & \cdot \\ & & & & & & & 0 \end{bmatrix} S^{-1} \tag{22}$$

is the required projection. Obviously A maps $R(A)$ into $R(A)$. If $z \in R(A)$ and $Az = 0$, then $z \in N(A) \cap R(A) = \{0\}$ and so $z = 0$. This proves that A is invertible on $R(A)$. □

Remark 6. Under the assumptions of Lemma 5, we will call the component of a vector c in $N(A)$ the projection of c on $N(A)$ along $R(A)$. Note that by definition of index, the condition in the lemma is equivalent to $N(A^2) = N(A)$.

Theorem 7. *Let A be a matrix in M_n and b a vector in $\mathbb{C}^n$. Let Q be an invertible matrix in M_n and let $B = I - Q^{-1}A$. Assume that $\rho(B) \leq 1$ and that $\operatorname{index}(\lambda) = 1$ for every eigenvalue λ of B with modulus 1, that is, B is nonexpansive relative to a matrix norm. Starting with an initial vector x_0 in $\mathbb{C}^n$ define x_k recursively by*

$$Q(x_k) = (Q - A)x_{k-1} + b \tag{23}$$

for $k = 1, 2, \ldots$. Let

$$y_k = \frac{x_0 + x_1 + \cdots + x_{k-1}}{k}. \tag{24}$$

If $Ax = b$ is consistent, that is, has a solution, then $y_k, k = 1, 2, \ldots,$ converge to a solution vector z with rate of convergence $\|y_k - z\| = O(1/k)$. If $Ax = b$ is inconsistent, then $\lim_k \|x_k\| = \lim_k \|y_k\| = \infty$. More precisely, $\lim_k x_k/k = c'$ and $\lim_k y_k/k = c'/2$, where $c = Q^{-1}b$ and c' is the projection of c on $N(A) = N(Q^{-1}A)$ along $R(Q^{-1}A)$.

Proof. First we assume that A is invertible so that $I - B = Q^{-1}A$ is also invertible. Let T be the mapping defined by $Tx = Bx + c$. Then $T^k x = B^k x + c + Bc + \cdots + B^{k-1}c$. Let $s = c + Bc + \cdots + B^{k-1}c$. Then $s - Bs = c - B^k c$ and hence $s = (I - B)^{-1}c - (I - B)^{-1}B^k c = (I - B)^{-1}c - B^k(I - B)^{-1}c$. Let $z = (I - B)^{-1}c = A^{-1}b$. z is the unique solution of $Ax = b$ and

$$T^k x = B^k x + z - B^k z = B^k(x - z) + z. \tag{25}$$

Since the sequence x_k in the theorem is $T^k x_0$, we have

$$y_k = \frac{1}{k}\left(I + B + \cdots + B^{k-1}\right)(x_0 - z) + z = B_k(x_0 - z) + z. \tag{26}$$

Since $I - B$ is invertible, 1 is not an eigenvalue of B, and by Theorem 3 part (a) $\|y_k - z\| = \|B_k(x_0 - z\| \to 0$ as $k \to \infty$. Moreover, from the proof of the same theorem, $\|y_k - z\| = O(1/k)$.

Next we consider the case when A is not invertible. Since Q is invertible, we have $R(Q^{-1}A) = Q^{-1}(R(A))$ and $N(Q^{-1}A) = N(A)$. The index of the eigenvalue 0 of $Q^{-1}A$ is the index of eigenvalue 1 of $B = I - Q^{-1}A$. Thus by Lemma 5, $\mathbb{C}^n = Q^{-1}(R(A)) \oplus N(A)$. For every vector $v \in \mathbb{C}^n$, let $v^{(r)}$ and $v^{(n)}$ denote the component of v in the subspace $Q^{-1}(R(A))$ and $N(A)$, respectively.

Assume that $Ax = b$ is consistent, that is, $b \in R(A)$. Then $c \in R(Q^{-1}A)$. By Lemma 5, the restriction of $Q^{-1}A$ on its range is invertible, so there exists a unique z' in $R(Q^{-1}A)$ such that $Q^{-1}Az' = c$, or equivalently, $(I - B)z' = c$. For any vector x, we have

$$\begin{aligned} T^k x &= B^k x + c + Bc + \cdots + B^{k-1}c \\ &= B^K\left(x^{(r)} + x^{(n)}\right) + \left(I + B + \cdots + B^{k-1}\right)(I - B)z' \\ &= B^k\left(x^{(r)}\right) + x^{(n)} + z' - B^k\left(z'\right) \\ &= B^k\left(x^{(r)} - z'\right) + x^{(n)} + z'. \end{aligned} \tag{27}$$

Since B maps $R(Q^{-1}A)$ into $R(Q^{-1}A)$ and $I - B = Q^{-1}A$ restricted to $R(Q^{-1}A)$ is invertible, we can apply the preceding proof and conclude that the sequence y_k as defined before converges to $z = x_0^{(n)} + z'$ and $\|y_k - z\| = O(1/k)$. Now $Az = A(x_0^{(n)}) + A(z') = A(z') = Qc = b$, showing that z is a solution of $Ax = b$.

Assume now that $b \notin R(A)$, that is, $Ax = b$ is inconsistent. Then $c \notin R(Q^{-1}A)$ and $c = c^{(r)} + c^{(n)}$ with $c^{(n)} \neq 0$. As in the preceding case there exists a unique $z' \in R(Q^{-1}A)$ such that $(I - B)z' = c^{(r)}$. Note that for all $y \in N(A)$, $B(y) = (I - Q^{-1}A)(y) = y$. Thus for any vector x and any positive integer j

$$\begin{aligned} x_j &= T^j x \\ &= B^j x + c + Bc + \cdots + B^{j-1}c \\ &= B^j\left(x^{(r)} + x^{(n)}\right) + \left(I + B + \cdots + B^{j-1}\right)(I - B)z' + jc^{(n)} \\ &= B^j\left(x^{(r)}\right) + x^{(n)} + z' - B^j\left(z'\right) + jc^{(n)} \\ &= B^j\left(\mathrm{x}^{(r)} - z'\right) + x^{(n)} + z' + jc^{(n)}, \\ y_k &= \frac{1}{k}\left(x + Tx + \cdots + T^{k-1}x\right) \\ &= B_k\left(x^{(r)} - z'\right) + x^{(n)} + z' + \frac{k-1}{2}c^{(n)}, \end{aligned} \tag{28}$$

where $B_k = (I + B + \cdots + B^{k-1})$. As in the preceding case, $B^k(x^{(r)} - z'), k = 0, 1, 2, \ldots$ is bounded and $B_k(x^{(r)} - z'), k = 1, 2, \ldots,$ converges to 0. Thus $\lim_{k\to\infty}(x_k/k) = c(n)$ and $\lim_{k\to\infty}(y_k/k) = c^{(n)}/2$, and hence $\lim_{k\to\infty}\|x_k\| = \lim_{k\to\infty}\|y_k\| = \infty$. This completes the proof. □

Next we consider another kind of iteration in which the nonlinear case was considered in Ishikawa [4]. Note that the type of mappings in this case is slightly weaker than nonexpansivity (see condition (c) in the next lemma).

Lemma 8. *Let B be an $n \times n$ matrix. The following are equivalent:*

(a) *for every $0 < \mu < 1$, there exists a matrix norm $\|\cdot\|_\mu$ such that $\|\mu I + (1-\mu)B\|_\mu \leq 1$,*

(b) *for every $0 < \mu < 1$, there exists an induced matrix norm $\|\cdot\|_\mu$ such that $\|\mu I + (1-\mu)B\|_\mu \leq 1$,*

(c) *$\rho(B) \leq 1$ and* index(1) = 1 *if 1 is an eigenvalue of B.*

Proof. As in the proof of Theorem 2, (a) and (b) are equivalent. For $0 < \mu < 1$, denote $\mu I + (1-\mu)B$ by $B(\mu)$. Suppose now that (a) holds. Let λ be an eigenvalue of B. Then $\mu + (1-\mu)\lambda$ is an eigenvalue of $B(\mu)$. By Theorem 2 $|\mu + (1-\mu)\lambda| \leq 1$ for every $0 < \mu < 1$ and hence $|\lambda| \leq 1$. If 1 is an eigenvalue of B, then it is also an eigenvalue of $B(\mu)$. By Theorem 2, the index of 1, as an eigenvalue of $B(\mu)$, is 1. Since obviously B and $B(\mu)$ have the same eigenvectors corresponding to the eigenvalue 1, the index of 1, as an eigenvalue of B, is also 1. This proves (c).

Now assume (c) holds. Since $|\mu + (1-\mu)\lambda| < 1$ for $|\lambda| = 1, \lambda \neq 1$, every eigenvalue of $B(\mu)$, except possibly for 1, has modulus less than 1. Reasoning as above, if 1 is an eigenvalue of $B(\mu)$, then its index is 1. Therefore by Theorem 2, (a) holds. This completes the proof. □

Theorem 9. *Let $A \in M_n$ and $b \in \mathbb{C}^n$. Let Q be an invertible matrix in M_n, and $B = I - Q^{-1}A$. Suppose $\rho(B) \leq 1$ and that* index(1) = 1 *if 1 is an eigenvalue of B. Let $0 < \mu < 1$ be fixed. Starting with an initial vector x_0, define $x_k, y_k, k = 0, 1, 2, \ldots,$ recursively by*

$$\begin{aligned} y_0 &= x_0, \\ Q(x_k) &= (Q - A)(y_{k-1}) + b, \\ y_k &= \mu y_{k-1} + (1-\mu)x_k. \end{aligned} \tag{29}$$

If $Ax = b$ is consistent, then $y_k, k = 0, 1, 2, \ldots,$ converges to a solution vector z of $Ax = b$ with rate of convergence given by

$$\|y_k - z\| = o\left(\zeta^k\right), \tag{30}$$

where ζ is any number satisfying

$$\max\left\{\left|\mu + (1-\mu)\lambda\right| : \lambda \text{ an eigenvalue of } B,\ \lambda \neq 1\right\} < \zeta < 1. \tag{31}$$

If $Ax = b$ *is inconsistent, then* $\lim_{k\to\infty}\|y_k\| = \infty$*; more precisely,*

$$\lim_{k\to\infty}\frac{y_k}{k} = (1-\mu)c^{(n)}, \tag{32}$$

where $c^{(n)}$ *is the projection of* c *on* $N(A)$ *along* $R(Q^{-1}A)$*.*

Proof. Let $c = Q^{-1}b, B_1 = \mu I + (1-\mu)B = I - (1-\mu)Q^{-1}A$, and $Tx = B_1x + (1-\mu)c$. Then $y_k = T^k(x_0)$.

First we assume that A is invertible. Then $I - B_1 = (1-\mu)Q^{-1}A$ is invertible and 1 is not an eigenvalue of B_1; thus $\rho(B_1) < 1$. Let $z = (1-\mu)(I-B_1)^{-1}c = A^{-1}b$. We have

$$\begin{aligned} y_k &= T^k x_0 \\ &= B_1^k x_0 + (1-\mu)\left(c + B_1 c + \cdots + B_1^{k-1}c\right) \\ &= B_1^k x_0 + (1-\mu)\frac{1}{1-\mu}\left(I + B_1 + \cdots + B_1^{k-1}\right)(I-B_1)z \\ &= B_1^k(x_0 - z) + z. \end{aligned} \tag{33}$$

By a well known theorem (see, e.g. [1]), $\|y_k - z\| = o(\zeta^k)$ for every $\zeta > \rho(B_1)$.

Assume now that A is not invertible and $b \in R(A)$. Then c is in the range of $Q^{-1}A$. Since $B = I - Q^{-1}A$ satisfies the condition in Lemma 8, $Q^{-1}A$ satisfies the condition in Lemma 5. Thus the restriction of $Q^{-1}A$ on its range is invertible and there exists z' in $R(Q^{-1}A)$ such that $Q^{-1}Az' = c$, or equivalently, $(I - B_1)z' = (1-\mu)c$. For any vector $x = x_0$, we have

$$\begin{aligned} y_k &= T^k(x) \\ &= B_1^k x + (1-\mu)\left(c + B_1 c + \cdots + B_1^{k-1}c\right) \\ &= B_1^k\left(x^{(r)} + x^{(n)}\right) + \left(I + B_1 + \cdots + B_1^{k-1}\right)(I-B_1)z' \\ &= B_1^k\left(x^{(r)}\right) + x^{(n)} + z' - B_1^k(z') \\ &= B_1^k\left(x^{(r)} - z'\right) + x^{(n)} + z'. \end{aligned} \tag{34}$$

Since B_1 maps $R(Q^{-1}A)$ into $R(Q^{-1}A)$ and $I - B = Q^{-1}A$ restricted to $R(Q^{-1}A)$ is invertible, we can apply the preceding proof and conclude that the sequence $y_k, k = 0, 1, 2, \ldots$ converges to $z = x^{(n)} + z'$ and $\|y_k - z\| = o(\zeta^k)$. z solves $Ax = b$ since $Az = A(x^{(n)}) + A(z') = A(z') = Qc = b$.

Assume lastly that $b \notin R(A)$, that is, $Ax = b$ is inconsistent. Then $c \notin R(Q^{-1}A)$ and $c = c^{(r)} + c^{(n)}$ with $c^{(n)} \neq 0$. As before there exists $z' \in R(Q^{-1}A)$ such that $(I - b_1)z' = (1-\mu)c^{(r)}$. Note that $B_1(p) = p$ for $p \in N(A)$. Then

$$\begin{aligned} y_k &= T^k(x) \\ &= B_1^k x + (1-\mu)\left(c + B_1 c + \cdots + B_1^{k-1} c\right) \\ &= B_1^k\left(x^{(r)} + x^{(n)}\right) + \left(I + B_1 + \cdots + B_1^{k-1}\right)(I - B_1)z' + k(1-\mu)c^{(n)} \\ &= B_1^k\left(x^{(r)} - z'\right) + x^{(n)} + z' + k(1-\mu)c^{(n)}. \end{aligned} \tag{35}$$

Since $B_1^k(x^{(r)} - z'), k = 0, 1, 2, \ldots,$ converges to 0, we have

$$\lim_{k\to\infty} \frac{y_k}{k} = (1-\mu)c^{(n)}, \tag{36}$$

and hence $\lim_{k\to\infty}\|y_k\| = \infty$. This completes the proof. □

By taking $Q = I$ and considering only nonexpansive matrices in Theorems 7 and 9, we obtain the following corollary.

Corollary 10. *Let A be an $n \times n$ matrix such that $\|I - A\| \le 1$ for some matrix norm $\|\cdot\|$. Let b be a vector in $\mathbb{C}^n$. Then:*

(a) *starting with an initial vector x_0 in $\mathbb{C}^n$ define x_k recursively as follows:*

$$x_k = (I - A)(x_{k-1}) + b \tag{37}$$

for $k = 1, 2, \ldots$. Let

$$y_k = \frac{x_0 + x_1 + \cdots + x_{k-1}}{k} \tag{38}$$

for $k = 1, 2, \ldots$. If $Ax = b$ is consistent, then $y_k, k = 1, 2, \ldots,$ converges to a solution vector z with rate of convergence given by

$$\|y_k - z\| = O\left(\frac{1}{k}\right). \tag{39}$$

If $Ax = b$ is inconsistent, then $\lim_{k\to\infty}\|x_k\| = \lim_{k\to\infty}\|y_k\| = \infty$.

(b) *let $0 < \mu < 1$ be a fixed number. Starting with an initial vector x_0, let*

$$\begin{aligned} y_0 &= x_0, \\ x_k &= (I - A)(y_{k-1}) + b, \\ y_k &= \mu y_{k-1} + (1-\mu)x_k. \end{aligned} \tag{40}$$

If $Ax = b$ is consistent, then $y_k, k = 0, 1, 2, \ldots$, converges to a solution vector z of $Ax = b$ with rate of convergence given by

$$\|y_k - z\| = o\left(\zeta^k\right) \tag{41}$$

where ζ is any number satisfying

$$\max\left\{\left|\mu + (1-\mu)\lambda\right| : \lambda \text{ an eigenvalue of } B,\ \lambda \neq 1\right\} < \zeta < 1. \tag{42}$$

If $Ax = b$ is inconsistent, then $\lim_{k\to\infty}\|y_k\| = \infty$.

Remark 11. If in the previous corollary, $\|I - A\| < 1$, and $\mu = 0$ in part (b), the sequence $y_k = x_k$ converges to a solution. This is the Richardson method, see for example, [3]. Even in this case, our method in part (b) may yield a better approximation. For example if

$$A = \begin{pmatrix} 1 & 0.9 \\ -0.9 & 1 \end{pmatrix}, \tag{43}$$

$b = \mathbf{0}$, and $x_0 = \mathbf{e}_1$, then the nth iterate in the Richardson method is 0.9^n away from the solution 0, while the nth iterate using the method in the corollary part (b) with $\mu = 1/2$ is less than $(0.5)^{n/2}$.

An $n \times n$ matrix $A = (a_{ij})$ is called *diagonally dominant* if

$$|a_{ii}| \geq \sum_{j=1, j\neq i}^{n} \left|a_{ij}\right| \tag{44}$$

for all $i = 1, \ldots, n$. If A is diagonally dominant with $a_{ii} \neq 0$ for every i and if $Q = D$ or L, where D is the diagonal matrix containing the diagonal of A, and L the lower triangular matrix containing the lower triangular entries of A, then it is easy to prove that $\|I - Q^{-1}A\|_\infty \leq 1$ where $\|\cdot\|_\infty$ denotes the maximum row sum matrix norm; see, for example, [1, 3]. The following follows from Theorems 7 and 9.

Corollary 12. *Let A be a diagonally dominant $n\times n$ matrix with $a_{ii} \neq 0$ for all $i = 1, \ldots, n$. Let $Q = D$ or L, where D is the diagonal matrix containing the diagonal of A, and L the lower triangular matrix containing the lower triangular entries of A. Let b be a vector in $\mathbb{C}^n$. Then:*

(a) *starting with an initial vector x_0 in $\mathbb{C}^n$ define x_k recursively as follows:*

$$Q(x_k) = (Q - A)(x_{k-1}) + b \tag{45}$$

for $k = 1, 2, \ldots$. Let

$$y_k = \frac{x_0 + x_1 + \cdots + x_{k-1}}{k} \tag{46}$$

for $k = 1, 2, \ldots$. *If* $Ax = b$ *is consistent, then* $y_k, k = 1, 2, \ldots$ *converges to a solution vector* z *with rate of convergence given by*

$$\|y_k - z\| = O\left(\frac{1}{k}\right). \tag{47}$$

If $Ax = b$ *is inconsistent, then* $\lim_{k\to\infty}\|x_k\| = \lim_{k\to\infty}\|y_k\| = \infty$.

(b) *Let* $0 < \mu < 1$ *be a fixed number. Starting with an initial vector* x_0, *let*

$$\begin{aligned} y_0 &= x_0, \\ Q(x_k) &= (Q - A)(y_{k-1}) + b, \\ y_k &= \mu y_{k-1} + (1-\mu)x_k. \end{aligned} \tag{48}$$

If $Ax = b$ *is consistent, then* $y_k, k = 0, 1, 2, \ldots$, *converges to a solution vector* z *of* $Ax = b$ *with rate of convergence given by*

$$\|y_k - z\| = o\left(\zeta^k\right), \tag{49}$$

where ζ *is any number satisfying*

$$\max\left\{\left|\mu + (1-\mu)\lambda\right| : \lambda \text{ an eigenvalue of } B,\ \lambda \neq 1\right\} < \zeta < 1. \tag{50}$$

If $Ax = b$ *is inconsistent, then* $\lim_{k\to\infty}\|y_k\| = \infty$.

References

[1] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

[2] N. Dunford and J. T. Schwartz, *Linear Operators, Part I*, Interscience Publishers, New York, NY, USA, 1957.

[3] D. Kincaid and W. Cheney, *Numerical Analysis*, Brooks/Cole, Pacific Grove, Calif, USA, 1991.

[4] S. Ishikawa, "Fixed points and iteration of a nonexpansive mapping in a Banach space," *Proceedings of the American Mathematical Society*, vol. 59, no. 1, pp. 65–71, 1976.