NEW ZEALAND JOURNAL OF MATHEMATICS
Volume 38 (2008), 171–178

THE nTH POWER OF A MATRIX AND APPROXIMATIONS FOR LARGE n

Christopher S. Withers and Saralees Nadarajah

(Received July 2008)

Abstract. When a square matrix A, is diagonalizable, (for example, when A is Hermitian or has distinct eigenvalues), then A^n can be written as a sum of the nth powers of its eigenvalues with matrix weights. However, if a 1 occurs in its Jordan form, then the form is more complicated: A^n can be written as a sum of polynomials of degree n in its eigenvalues with coefficients depending on n. In this case to a first approximation for large n, A^n is proportional to $n^{m-1}\lambda^n$ with a constant matrix multiplier, where λ is the eigenvalue of maximum modulus and m is the maximum multiplicity of λ.

1. Introduction and Summary

Let A be an $s \times s$ complex matrix with eigenvalues $\lambda_1, \cdots, \lambda_s$ ordered so that

$$r_1 \stackrel{\text{def}}{=} |\lambda_1| = \cdots = |\lambda_N| > r_0 \stackrel{\text{def}}{=} |\lambda_{N+1}| \geq \cdots \geq |\lambda_s|. \tag{1.1}$$

If all eigenvalues have the same modulus then (1.1) should be interpreted as

$$r_1 \stackrel{\text{def}}{=} |\lambda_1| = \cdots = |\lambda_s|$$

with $N = s$. If A is Hermitian, or more generally if it has diagonal Jordan form, then for $n \geq 1$, $A^n = \sum_{j=1}^{s} \lambda_j^n \, W_j$, for certain matrices $\{W_j\}$. Let us write r_j and θ_j for the amplitude and phase of λ_j, that is,

$$\lambda_j = r_j \ e^{i\theta_j} \tag{1.2}$$

for $1 \leq j \leq s$. Then this implies the approximations

$$A^n = r_1^n C_n + r_0^n E_n, \quad A^n = \lambda_1^n W_1 + r_0^n F_n \tag{1.3}$$

(with the latter holding if $N = 1$), where

$$C_n = \sum_{j=1}^{N} e^{in\theta_j} W_j$$

with $||E_n||$, $||F_n||$ and $||C_n||$ bounded as n increases. These cases are covered in Sections 2 and 3. Otherwise, A^n can be written as a sum of polynomials in each eigenvalue:

$$A^n = \sum_{j=1}^{r} f_j(\lambda_j),$$

2000 *Mathematics Subject Classification* Primary 15A21 Canonical forms, reductions, classification Secondary 15A24 Matrix equations and identities.

Key words and phrases: powers of matrices, approximations, Jordan form, Singular value decomposition.

where

$$f_j(\lambda) = \sum_{k=0}^{m_j-1} \binom{n}{k} \lambda^{n-k} W_{jk}$$

for $n \geq 1$ and certain matrices $\{W_{jk}\}$. Here m_j is the dimension of the jth Jordan block of A and r is the number of blocks. So, λ_j has multiplicity at least m_j. Suppose the blocks for λ_1 are ordered so that

$$m = m_1 = \cdots = m_M > m_j, \ M < j \leq N \tag{1.4}$$

with (1.4) interpreted as

$$m = m_1 = \cdots = m_N$$

when all Jordan blocks have dimension m. Then for $n \geq 1$,

$$A^n = \binom{n}{m-1} r_1^n \ [C_n + n^{-1} E_n^*],$$

where

$$C_n = \sum_{j=1}^{M} e^{i(n-m+1)\theta_j} \ W_{j,m-1}$$

with $||E_n^*||$ and $||C_n||$ bounded as n increases. For example, if there is only one Jordan block with eigenvalue λ_1 then

$$A^n = \binom{n}{m-1} \lambda_1^n \ [W_{1,m-1} + n^{-1} F_n^*],$$

where $||F_n^*||$ is bounded as n increases. Details are given in Section 4. Section 5 gives results similar to (1.3) for $(AA^T)^n$ and related forms for large n based on the singular-value decomposition (SVD). In this section A need not be square. Throughout this note, $a_n = O(b_n)$ means that an integer m and a positive constant K exist such that for all $n \geq m$, $|a_n| \leq K|b_n|$.

2. A^n for A Hermitian

Consider the case where A is Hermitian, that is $A^T = A$, where A^T is the transpose of the complex conjugate of A. Then the eigenvalues are real,

$$A = H\Lambda H^T = \sum_{j=1}^{s} \lambda_j p_j p_j^T, \tag{2.1}$$

where $\Lambda = \lceil \lambda_1, \cdots, \lambda_s \rfloor$, $H = (p_1, \cdots, p_s)$, and p_j is the eigenvector of λ_j, that is, a solution of $Ap_j = \lambda_j p_j$. The eigenvectors can be taken as orthogonal and scaled to have unit norm, so that $p_j^T p_k = \delta_{jk}$ and $\sum_{j=1}^{s} p_j p_j^T = I_s$, where $\delta_{jj} = 1$ and $\delta_{jk} = 0$ for $j \neq k$. If A is real symmetric then H can be taken as real. So, by (1.1) we have the well known expression

$$A^n = H\Lambda^n H^T = \sum_{j=1}^{s} \lambda_j^n \ p_j p_j^T. \tag{2.2}$$

It follows that $A^n = r_1^n C_n + r_0^n E_n$, $n = 0, 1, \cdots$, where $C_n = \sum_{j=1}^{N} s_j^n p_j p_j^T$, $s_j = \text{sign}(\lambda_j)$ and $||E_n||$ is bounded as n increases. So, for $n \geq 1$, if $N = 1$ then

$A^n = \lambda_1^n D + r_0^n F_n$, where $D = p_1 p_1^T$ and $||F_n||$ is bounded as n increases. On the other hand, if $N \geq 2$ then the multiplier C_n alternates between C_+, its form for n even, and C_-, its form for n odd, where $C_+ = \sum_{j=1}^s p_j p_j^T$, $C_- = \sum_{s_j=1} p_j p_j^T - \sum_{s_j=-1} p_j p_j^T$. Also if $det(A) \neq 0$ then (2.2) extends to $n = -1, -2, \cdots$, and in fact to complex n. For example, $A^{-1} = H\Lambda^{-1}H^T = \sum_{j=1}^s \lambda_j^{-1} p_j p_j^T$. Section 5 gives results similar to those of Section 3 but for $(AA^T)^n$, $(AA^T)^n A$. These are based on the SVD of A which need not be square.

3. Diagonal Jordan Form

Now suppose that A has diagonal Jordan form, that is $A = P\Lambda P^{-1}$ for some matrix P and Λ of (2.1). Such a matrix is said to be *diagonalizable*. If A and its eigenvalues are real, then P can be taken as real. Writing $P = (p_1, \cdots, p_s)$, $Q^T = P^{-1} = (q_1, \cdots, q_s)^T$, we have $q_j^T p_k = \delta_{jk}$, $\sum_{j=1}^s p_j q_j^T = I_s$. By (1.1) and (1.2), for $n \geq 1$,

$$A^n = P\Lambda^n P^{-1} = \sum_{j=1}^{s} \lambda_j^n \; p_j q_j^T = r_1^n \; C_n + r_0^n E_n, \tag{3.1}$$

where $C_n = \sum_{j=1}^N e^{in\theta_j} \; p_j q_j^T$ and $||E_n||$ is bounded as n increases. For example, if $N = 1$ then $A^n = \lambda_1^n D + r_0^n F_n$, where $D = p_1 q_1^T$ and $||F_n||$ is bounded as n increases. On the other hand, if $N = 2$ then $A^n = \lambda_1^n D_n + r_0^n F_n$, where $D_n = p_1 q_1^T + e^{in(\theta_2 - \theta_1)} p_2 q_2^T$ and $||F_n||$ is bounded as n increases. In particular if $\theta_2 - \theta_1 = 2\pi/K$ for some $K = 1, 2, \cdots$, then D_n takes K distinct forms: $D_n = D_m$ when n modulo$(K) = m$, $m = 0, 1, \cdots, K-1$. If $det(A) \neq 0$ then (3.1) extends to $n = -1, -2, \cdots$, and to complex n. For example, $A^{-1} = P\Lambda^{-1}P^{-1} = \sum_{j=1}^s \lambda_j^{-1} p_j q_j^T$.

4. General Jordan Form

Now suppose that A has non-diagonal Jordan form, that is, it has one or more multiple eigenvalues, and $A = PJP^{-1}$, $J = \lceil J_1, \cdots, J_r \rfloor$, $J_j = J_{m_j}(\lambda_j)$, where

$$J_m(\lambda) = \lambda I_m + U_m = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

for some matrix P, and U_m is the $m \times m$ matrix with 1s on the superdiagonal and 0s elsewhere: $(U_m)_{jk} = \delta_{j,k-1}$. Again if A and its eigenvalues are real, then P can be taken as real.

Methods for obtaining P are known, see, for example, Dunford and Schwartz (1958), Finkbeiner II (1978), Horn and Johnson (1985) and Golub and van Loan

(1996). The $m \times m$ matrix U_m^n is defined as $J_m(0)^n$, so that $U_m^0 = I$,

$$
\begin{aligned}
U_m^1 &= \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \\ & & & & & 0 \end{pmatrix}, \\
U_m^2 &= \begin{pmatrix} 0 & 0 & 1 & & & \\ & 0 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 0 & 0 & 1 \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}, \\
U_m^3 &= \begin{pmatrix} 0 & 0 & 0 & 1 & & \\ & 0 & 0 & 0 & 1 & \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}, \\
U_m^{m-1} &= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ & 0 & 0 & \ddots & 0 & 0 \\ & & \ddots & \ddots & \ddots & \ddots \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}
\end{aligned}
$$

with $U_m^k = 0$ for $k \geq m$. Also by the binomial expansion, for $\lambda \neq 0$ and $n \geq 1$,

$$
\begin{aligned}
J_m(\lambda)^n &= \sum_{j=0}^{\min(n,m-1)} \binom{n}{j} \lambda^{n-j} U_m^j \qquad (4.1) \\
&= \lambda^n e_{n,m-1}(\lambda)\ [U_m^{m-1} + n^{-1}\Omega_n] \\
&= d_{n,m-1}(\lambda) + n^{m-2}\lambda^n \Omega_n^* \\
&= n^{m-1}\lambda^n \Omega_n^{**},
\end{aligned}
$$

where $e_{nk}(\lambda) = \binom{n}{k}\lambda^{-k}$, $d_{nk}(\lambda) = \lambda^n e_{nk}(\lambda) U_m^{m-1}$ and U_m^{m-1} is the $m \times m$ matrix of 0s except for a 1 in the upper right corner. Further, $||\Omega_n||$, $||\Omega_n^*||$ and $||\Omega_n^{**}||$ are bounded as n increases. Another way to write (4.1) is

$$
J_m(\lambda)^n = \lambda^n \begin{pmatrix} 1 & e_{n1} & e_{n2} & \cdots & e_{n,m-1} \\ 0 & 1 & e_{n1} & \cdots & e_{n,m-2} \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}
$$

at $e_{nk} = e_{nk}(\lambda)$. So, we obtain the exact formula

$$A^n = PJ^nP^{-1} = P\lceil J_1^n, \cdots, J_r^n \rfloor P^{-1},$$

where J_j^n is given by (4.1). Suppose now that N of the Jordan blocks $\{J_j\}$ have eigenvalues with the maximum amplitude r_1, and that these are ordered so that

$$r_1 = |\lambda_1| = \cdots = |\lambda_N| > r_0 = \max_{N<j\leq s} |\lambda_j|.$$

Because of multiplicities, this N is not same as N of (1.1). Suppose also that the first $M \leq N$ blocks have the maximum multiplicity m, that is, (1.4) holds, and that m_0 is the maximum multiplicity of eigenvalues of amplitude r_0. Then for $n \geq 1$,

$$\begin{aligned} A^n &= P\lceil J_1^n, \cdots, J_N^n, 0, \cdots, 0\rfloor P^{-1} + n^{m_0-1} r_0^n \Psi_n \\ &= r_1^n e_{n,m-1}(r_1) P \lceil d_{n,m-1}(e^{i\theta_1}), \cdots, d_{n,m-1}(e^{i\theta_M}), \cdots, 0, \cdots, 0\rfloor P^{-1} + n^{m-2} r_1^n \Psi_n^* \end{aligned}$$

for $\{\theta_j\}$ of (1.2), where $||\Psi_n||$ and $||\Psi_n^*||$ are bounded as n increases. Now partition P and its inverse into $r \times r$ blocks to match J, say $P = (P_{jk}),\ P^{-1} = (P^{jk})$. So, A^n is the $r \times r$ block matrix with (j, k) element

$$\begin{aligned} (A^n)_{jk} &= \sum_{c=1}^{r} P_{jc}\ J_c^n\ P^{ck} \\ &= \sum_{c=1}^{N} P_{jc}\ J_c^n\ P^{ck} + O(n^{m_0-1} r_0^n) \\ &= r_1^n\ e_{n,m-1}(r_1) \sum_{c=1}^{M} P_{jc}\ d_{n,m-1}(e^{i\theta_c})\ P^{ck} + O(n^{m-1} r_1^n), \end{aligned} \tag{4.2}$$

giving

$$A^n = \binom{n}{m-1} r_1^{n-m+1} [C_n + n^{-1}\Psi_n^{**}] = n^{m-1} r_1^n \Psi_n^{***},$$

where

$$(C_n)_{jk} = \sum_{c=1}^{M} e^{i(n-m+1)\theta_c}\ P_{jc}\ U_m^{m-1}\ P^{ck} = O(1)$$

and $||\Psi_n^{**}||$ and $||\Psi_n^{***}||$ are bounded as n increases. The last multiplier $P_{jc}\, U_m^{m-1}\ P^{ck}$ is an $m \times m$ matrix with (p, q) element $(P_{jc})_{p1}\ (P^{ck})_{mq}$. For example, if $|\lambda_1| > |\lambda_j|$ for $j > 1$, then $M = 1$ and

$$A^n = \binom{n}{m-1} \lambda_1^{n-m+1} [D + n^{-1}\Psi_n^{****}],$$

where $[D_{jk}]_{pq} = (P_{j1})_{p1}\ (P^{1k})_{mq}$ and $||\Psi_n^{****}||$ is bounded as n increases. If $det(A) \neq 0$ then (4.2) and (4.1) extend to $n = -1, -2, \cdots$, and to complex n. For example,

$$(A^{-1})_{jk} = \sum_{c=1}^{r} P_{jc}\ J_c^{-1}\ P^{ck},$$

where

$$J_m(-\lambda)^{-1} = -\sum_{j=0}^{m-1} \lambda^{-1-j} U_m^j.$$

Example 4.1. *Take*

$$A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{pmatrix}.$$

Then $r = 3,\ J_1 = 1,\ J_2 = 2,\ J_3 = J_2(4)$,

$$\begin{aligned} P &= \begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \\ P^{-1} &= \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \\ J^n &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2^n & 0 & 0 \\ 0 & 0 & 4^n & n4^{n-1} \\ 0 & 0 & 0 & 4^n \end{pmatrix}, \\ A^n &= 4^n(Bn/4 + C) + 2^n D + E = n4^n\ [B/4 + n^{-1}\Phi_n], \end{aligned}$$

where

$$\begin{aligned} B &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \\ C &= \begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \\ D &= \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \\ E &= \begin{pmatrix} 0 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

and $||\Phi_n||$ *is bounded as* n *increases.*

Example 4.2. *Take*

$$A = \frac{1}{9}\begin{pmatrix} 14 & -1 & 2 \\ -3 & 6 & 6 \\ 4 & -3 & 6 \end{pmatrix}.$$

Then $r = 2,\ J_1 = J_2(2),\ J_2 = J_1(1) = 1,$

$$\begin{aligned} P &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \\ P^{-1} &= \frac{1}{18}\begin{pmatrix} -5 & 1 & 7 \\ 7 & -5 & 1 \\ 1 & 7 & -5 \end{pmatrix}, \\ J^n &= \begin{pmatrix} 2^n & n2^{n-1} & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

So, $18A^n = (Bn/2 + C)2^n + D = n2^n\ [B/2 + n^{-1}\Phi_n^*]$, *where*

$$\begin{aligned} B &= \begin{pmatrix} 7 & -5 & 1 \\ 21 & -15 & 3 \\ 14 & -10 & 2 \end{pmatrix}, \\ C &= \begin{pmatrix} 9 & -9 & 7 \\ -8 & -2 & 22 \\ 11 & -13 & 17 \end{pmatrix}, \\ D &= \begin{pmatrix} 3 & 21 & -15 \\ 2 & 14 & -10 \\ 1 & 14 & -5 \end{pmatrix} \end{aligned}$$

and $||\Phi_n^*||$ *is bounded as* n *increases.*

5. Behaviour of $(AA^T)^n$ and Related Forms for Large n

This section gives results similar to those of Section 3 but for $(AA^T)^n$ and $(AA^T)^n A$. These are based on the SVD of A. The SVD of an $m \times n$ matrix A can be written

$$A = \sum_{j=1}^{r} \theta_j l_j r_j^T,$$

where $r = \min(m, n)$, $l_j^T l_k = r_j^T r_k = \delta_{jk}$ and $0 < \theta_1 \leq \cdots \leq \theta_r$. So, for $1 \leq j \leq r$, $Ar_j = \theta_j l_j$ and $A^T l_j = \theta_j r_j$. So,

$$\begin{aligned} AA^T &= \sum_{j=1}^{r} \theta_j^2 l_j l_j^T, \\ A^T A &= \sum_{j=1}^{r} \theta_j^2 r_j r_j^T, \\ AA^T A &= \sum_{j=1}^{r} \theta_j^3 l_j r_j^T. \end{aligned}$$

That is, $\{\theta_j^2\}$ are the non-zero eigenvalues of AA^T and A^TA, and the corresponding eigenvectors are $\{l_j\}$ and $\{r_j\}$. Similarly, we have for $n \geq 1$,

$$\begin{aligned}
(AA^T)^n &= \sum_{j=1}^{r} \theta_j^{2n} l_j l_j^T, \\
(A^TA)^n &= \sum_{j=1}^{r} \theta_j^{2n} r_j r_j^T, \\
(AA^T)^n A &= \sum_{j=1}^{r} \theta_j^{2n+1} l_j r_j^T, \\
(A^TA)^n A^T &= \sum_{j=1}^{r} \theta_j^{2n+1} r_j l_j^T.
\end{aligned}$$

So, if $\max_{1 \leq j \leq s} \theta_j < \theta_{s+1} = \cdots = \theta_r$ then as $n \to \infty$,

$$\begin{aligned}
(AA^T)^n &= \theta_r^{2n} L_{sr} + \theta_s^{2n} \Delta_n = \theta_r^{2n} \Xi_n, \\
(A^TA)^n &= \theta_r^{2n} R_{sr} + \theta_s^{2n} \Delta_n^* = \theta_r^{2n} \Xi_n^*, \\
(AA^T)^n A &= \theta_r^{2n+1} B_{sr} + \theta_s^{2n} \Delta_n^{**} = \theta_r^{2n} \Xi_n^{**}, \\
(A^TA)^n A^T &= \theta_r^{2n+1} C_{sr} + \theta_s^{2n} \Delta_n^{***} = \theta_r^{2n} \Xi_n^{***},
\end{aligned}$$

where $L_{sr} = \sum_{j=s+1}^{r} l_j l_j^T$, $R_{sr} = \sum_{j=s+1}^{r} r_j r_j^T$, $B_{sr} = \sum_{j=s+1}^{r} l_j r_j^T$ and $C_{sr} = \sum_{j=s+1}^{r} r_j l_j^T$ with $||\Delta_n||$, $||\Delta_n^*||$, $||\Delta_n^{**}||$, $||\Delta_n^{***}||$, $||\Xi_n||$, $||\Xi_n^*||$, $||\Xi_n^{**}||$, $||\Xi_n^{***}||$, $||L_{sr}||$, $||R_{sr}||$, $||B_{sr}||$ and $||C_{sr}||$ bounded as n increases.

Acknowledgments

The authors would like to thank the Managing Editor and the referee for carefully reading the paper and for their comments which greatly improved the paper.

References

[1] N. Dunford and J. T. Schwartz, Linear Operators, Part I: General Theory, Interscience, 1958.

[2] D. T. Finkbeiner II, Introduction to Matrices and Linear Transformations, Third Edition, Freeman, 1978.

[3] G. H. Golub and C. F. van Loan, Matrix Computations, Third Edition, Johns Hopkins University Press, Baltimore, 1996.

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

C.S. Withers
Applied Mathematics Group
Industrial Research Limited
Lower Hutt
NEW ZEALAND
c.withers@irl.cri.nz

S. Nadarajah
School of Mathematics
University of Manchester
Manchester M13 9PL
UNITED KINGDOM
mbbsssn2@manchester.ac.uk