Matrix Polynomial Expansion of a Power of a Matrix

By

Kazuaki ANDO and Bunij KONDO

(Received June 30, 1979)

Abstract

Any power of an $n \times n$ matrix can be expanded by a matrix polynomial of the order $n-1$, but the coefficients of expansion are not known in closed form. In this paper, it is shown that the coefficients of expansion are given by the solution of a simultaneous equation of the 1st order, whose coefficients compose the Vandermonde matrix. Using the properties of a generalized Vandermonde determinant, coefficients of an expansion of a power of the matrix are obtained in closed form. As the coefficients thus obtained are homogeneous polynomials of eigen-values, and as every term of the polynomial has the same sign, the upper bounds of the absolute values of the coefficients can be obtained easily, if all the eigen-values are located in a disk centered at the origin.

If all the eigen-values of a transition matrix of a dynamical system are located in a disc with a radius less than 1 and centered at the origin, the dynamical system is exponentially stable. As the reachable subspace of a dynamical system is spanned by input constraint vectors, multiplied by powers of the transition matrix from the left, the results obtained make a bridge to connect the exponential stability property and the structure of the practically reachable subspace of a dynamical system.

1. Introduction

To investigate the practically reachable subspace of a discrete-time linear dynamical system, it is sometimes necessary to evaluate the higher power of a matrix under the condition that all the eigen-values are located in a certain disc centered at the origin[1)].

Whittle derived a formula to evaluate the power of a matrix, using a multiple Laurent expansion method[2)]. As his results were expressed by the elements of the matrix and were very complex, the regions where the elements of the matrix are located are given only in a limiting case.

If A is an $n \times n$ matrix, then the $n+k$ th power of A, A^{n+k}, can be expanded by the polynomial of $I, A, \cdots, A^{n-1}$ either by the Silvester interpolation formula, or by a successive application of the Caley-Hamilton theorem. However, by these methods it is not easy to estimate the regions where the coefficients of expansion are located. By using the expression presented in this paper, the upper bounds

* Department of Electrical Engeering II

of the absolute values of the coefficients can be calculated when all the eigen-values are located in a disc centered at the origin.

2. Fundamental Formula

Suppose that eigen-values of an $n \times n$ matrix are distinct. It is well known that a power of A can be expressed by a polynomial of an order less than $n-1$ as follows:

$$A^{n+k} = \alpha_1(k)A^{n-1} + \alpha_2(k)A^{n-2} + \cdots + \alpha_n(k)I \tag{1}$$

As the eigen-values of A are assumed distinct, A can be diagonalized by a similar transformation. Any power of A is also diagonalized by the same transformation, so there is no loss of generality where A is assumed diagonal to evaluate coefficients of expansion in eq. (1).

If A is diagonal, the following equation is obtained by comparing the diagonal elements of both sides of eq. (1).

$$[\lambda_1^{n+k}, \cdots, \lambda_n^{n+k}] = [\alpha_n(k), \cdots, \alpha_1(k)]\begin{pmatrix} 1, & \cdots, & 1 \\ \lambda_1, & \cdots, & \lambda_n \\ \cdots & \cdots & \cdots \\ \lambda_1^{n-1}, & \cdots, & \lambda^{n-1} \end{pmatrix} \tag{2}$$

Therefore,

$$[\alpha_n(k), \cdots, \alpha_1(k), -1]\begin{pmatrix} 1, & \cdots, & 1 \\ \lambda_1, & \cdots, & \lambda_n \\ \cdots & \cdots & \cdots \\ \lambda_1^{n-1}, & \cdots, & \lambda_n^{n-1} \\ \lambda_1^{n+k}, & \cdots, & \lambda_n^{n+k} \end{pmatrix} = [0, \cdots, 0] \tag{3}$$

Now denote the Vandermonde matrix and the generalized Vandermonde matrix as follows.

$$V(\lambda) \triangleq \begin{pmatrix} 1, & \cdots, & 1 \\ \lambda_1, & \cdots, & \lambda_n \\ \cdots & \cdots & \cdots \\ \lambda^{n-1}, & \cdots, & \lambda_n^{n-1} \end{pmatrix}, \quad V_k(t, \lambda) \triangleq \begin{pmatrix} 1, & 1, & \cdots, & 1 \\ t, & \lambda_1, & \cdots, & \lambda_n \\ t^{n-1}, & \lambda_1^{n-1}, & \cdots, & \lambda_n^{n-1} \\ t^{n+k}, & \lambda_1^{n+k}, & \cdots, & \lambda_n^{n+k} \end{pmatrix} \tag{4}$$

From eqs. (3) and (4), the following equation holds true.

$$\begin{pmatrix} 1 & & 0 & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & 1 & 0 \\ \alpha_n(k), & \cdots, & \alpha_1(k), & -1 \end{pmatrix} V_k(t, \lambda) = \left(\begin{array}{l|c} 1, & \\ t, & V(\lambda) \\ \vdots & \\ t^{n-1}, & \\ \hline \sum_{\sigma=1}^{n-1} \alpha_{n-\sigma}(k)t^{\sigma} - t^{n-k} & 0 \end{array}\right) \tag{5}$$

Taking the determinants of both sides of eq. (5), we have

$$(-1)^{n-1}\left(\sum_{\sigma=1}^{n-1}\alpha_{n-\sigma}t^{\sigma}-t^{n-k}\right)|V(\lambda)| = |V_k(t,\lambda)| \tag{6}$$

From eq. (6) we can get the following lemma.

Lemma 1

The coefficient of expansion $\alpha_{n-\sigma}(k)$ is equal to the coefficient of the t^{σ} term of $|V_k(t,\lambda)|/|V(\lambda)|$ multiplied by $(-1)^{n-1}$.

3. Expansion Formula of the Vandermonde and Generalized Vandermonde Determinants

In the following sections, expansion formulae of the Vandermonde and generalized Vandermonde determinants are used. These formulae are cited below.

Lemma 2 (expansion of the Vandermonde determinant)

$$|V(x)| = (-1)^{n(n-1)/2}\prod_{i>j}^{n}(x_i-x_j) \tag{7}$$

This expansion formula is written in almost every book of matrix theory[3]. The expansion formula of the generalized Vandermonde determinant is given by the following lemma[4].

Lemma 3

$$\begin{vmatrix} x_1^{\nu_1}, & x_2^{\nu_1}, & \cdots, & x_n^{\nu_1} \\ x_1^{\nu_2}, & x_2^{\nu_2}, & \cdots, & x_n^{\nu_2} \\ \cdots & \cdots & \cdots & \cdots \\ x_1^{\nu_n}, & x_2^{\nu_n}, & \cdots, & x_n^{\nu_n} \end{vmatrix} = |V(x)|\cdot\begin{vmatrix} H_{\nu_1}, & H_{\nu_1-1}, & \cdots, & H_{\nu_1-n+1} \\ H_{\nu_2}, & H_{\nu_2-1}, & \cdots, & H_{\nu_2-n+1} \\ & & & \\ H_{\nu_n}, & H_{\nu_n-1}, & \cdots, & H_{\nu_n-n+1} \end{vmatrix} \tag{8}$$

where

$$\left.\begin{aligned} &H_m \triangleq \sum_{i_1+\cdots+i_n=m} x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} \qquad (m\geq 1) \\ &H_0 = 1, \quad H_m = 0 \qquad (m\leq -1) \end{aligned}\right\} \tag{9}$$

4. Evaluation of the Coeffcients of Expansion

By lemma 3, $V_k(t,\lambda)$ is expressed as

$$|V_k(t,\lambda)| = |V(t,\lambda)|\cdot H_k(t,\lambda) \tag{10}$$

where

$$|V(t,\lambda)| \triangleq \begin{vmatrix} 1, & 1, & \cdots, & 1 \\ t, & \lambda_1, & \cdots, & \lambda_n \\ \cdots & \cdots & \cdots & \cdots \\ t^n, & \lambda_1^n, & \cdots, & \lambda_n^n \end{vmatrix} \tag{11}$$

$$H_k(t, \lambda) = \sum_{i_0+\cdots+i_n=k} t^{i_0}\lambda_1^{i_1}\cdots\lambda_n^{i_n} \tag{12}$$

By lemma 2, $|V(t, \lambda)|$ is expanded into the polynomial of t as

$$\begin{aligned} &|V(t, \lambda)| \\ &= (-1)^{n(n-1)/2}\sum_{\mu=0}^{n}(-1)^{n-\mu}\sum_{i<\cdots<i_{n-\mu}}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_{n-\mu}}\sum_{1\leq j<k\leq n}(\lambda_j-\lambda_k)t^{\mu} \\ &= \sum_{\mu=0}^{n}(-1)^{\mu}s_{n-\mu}(\lambda_1, \cdots, \lambda_n)|V(\lambda)|t^{\mu} \end{aligned} \tag{13}$$

where $s_{n-\mu}(\lambda_1, \cdots, \lambda_n)$ is the fundamental symmetric form of the order $n-\mu$, and is given by

$$s_{n-\mu}(\lambda_1, \cdots, \lambda_n) \triangleq \sum_{i_1<\cdots<i_{n-\mu}}\lambda_{i_1}\cdots\lambda_{i_{n-\mu}} \tag{14}$$

From Eqs. (10) and (13),

$$|V_k(t, \lambda)| = \sum_{\sigma=0}^{n+k}\sum_{\mu=\max(0,\sigma-k)}^{n}(-1)^{\mu}s_{n-\mu}(\lambda_1, \cdots, \lambda_n)H_{k-\sigma+\mu}(\lambda)|V(\lambda)|t^{\sigma} \tag{15}$$

Applying lemma 1 to eq. (15), the following relation is obtained.

$$\alpha_{n-\sigma}(k) = (-1)^{n-1}\sum_{\mu=\max(0,\sigma-k)}^{n}(-1)^{\mu}s_{n-\mu}(\lambda)H_{k-\sigma+\mu}(\lambda) \tag{16}$$

Letting $n-\sigma=j$, $\mu=\sigma-\nu$, then

$$\begin{aligned} \alpha_j(k) &= (-1)^{j-1}\sum_{\nu=0}^{\min(k,n-j)}(-1)^{\nu}s_{j+\nu}(\lambda)H_{k-\nu}(\lambda) \\ &= \begin{cases} (-1)^{j-1}\{s_jH_k-s_{j+1}H_{k-1}+\cdots+(-1)^{n-j}s_nH_{k-n+j}\} & (k\geq n-j) \\ (-1)^{j-1}\{s_jH_k-s_{j+1}H_{k-1}+\cdots+(-1)^{k}s_{j+k}H_0\} & (k\leq n-j) \end{cases} \end{aligned} \tag{17}$$

To evaluate the r.h.s. of eq. (17), the following lemma is used.

Lemma 4

The product of s_p and H_q is a homogeneous polynomial of the order $p+q$, and is given by

$$s_p(\lambda)H_q(\lambda) = \sum_{\nu_1+\cdots+\nu_n=p+q}\binom{\beta(\nu_1, \cdots, \nu_n)}{p}\lambda_1^{\nu_1}\cdots\lambda_n^{\nu_n} \tag{18}$$

where $\beta(\lambda_1, \cdots, \lambda_n)$ is the number of the non-zero elements of $\{\nu_1, \cdots, \nu_n\}$.
Proof:

Being a homogeneous polynomial of the order $p+q$, $s_p(\lambda)H_q(\lambda)$ has a form such as

$$s_p(\lambda)H_q(\lambda) = \sum_{\nu_1+\cdots+\nu_n=p+q}K(\nu_1, \cdots, \nu_n)\lambda_1^{\nu_1}\cdots\lambda_n^{\nu_n} \tag{19}$$

As $s_p(\lambda)$ is the sum of the products of p distinct λ_i's, all the non-zero elements of the r.h.s. of eq. (19) have at least p non-zero ν_i's.

Now take a term from the r.h.s. of eq. (19). Without any loss of generality, one can assume that

$$\left.\begin{array}{l}\nu_1, \cdots, \nu_\beta \neq 0 \\ \nu_{\beta+1} = \cdots = \nu_n = 0\end{array}\right\} \tag{20}$$

Let $\{j_1, \cdots, j_p\}$ be an arbitrary subset of $\{1, \cdots, \beta\}$. For every subset $\{j_1, \cdots, j_p\}$, the single term $\lambda_{j1} \cdots \lambda_{jp}$ exists in $s_p(\lambda)$, and at the same time, the single term

$$\lambda_1^{\nu_1-\delta_1}\lambda_2^{\nu_2-\delta_2}\cdots\lambda_\beta^{\nu_\beta-\delta_\beta}, \qquad (\nu_1+\cdots+\nu_\beta = p+q)$$

where

$$\delta_i = \begin{cases} 0 & i \notin \{j_1, \cdots, j_p\} \\ 1 & i \in \{j_1, \cdots, j_p\} \end{cases}$$

exists in $H_q(\lambda)$, and the product of these two terms gives $\lambda_1^{\nu_1}\lambda_2^{\nu_2}\cdots\lambda_\beta^{\nu_\beta}$.

As there are ${}_\beta C_p$ ways of choosing $\{j_1, \cdots, j_p\}$ from $\{1, \cdots, \beta\}$, and no other combination of s_p and H_q gives $\lambda_1^{\nu_1}\cdots\lambda_\beta^{\nu_\beta}$, therefore

$$K(\nu_1, \cdots, \nu_n) = \binom{\beta(\nu_1, \cdots, \nu_n)}{p} \tag{21}$$

Q.E.D.

As

$$\beta(\nu_1, \cdots, \nu_n) \leq \min. (n, j+k) \tag{22}$$

by substituting eq. (18) into eq. (17), the coefficient of $\lambda_1^{\nu_1}\cdots\lambda_n^{\nu_n}$ as the term of the r.h.s. of eq. (17) becomes

$$(-1)^{j-1}\left\{\binom{\beta}{j} - \binom{\beta}{j+1} + \cdots + (-1)^{\beta-j}\binom{\beta}{\beta}\right\} = (-1)^{j-1}\binom{\beta-1}{\beta-j} \tag{23}$$

Therefore we get the following theorem.

Theorem 1

Coefficients of the expansion of A^{n+k} in form eq. (1) are given by

$$\alpha_j(k) = (-1)^{j-1} \sum_{\substack{\nu_1+\cdots+\nu_n=j+k \\ 0\leq\nu_i\leq j+k}} \binom{\beta(\nu_1, \cdots, \nu_n)-1}{\beta(\nu_1, \cdots, \nu_n)-j} \lambda_1^{\nu_1}\cdots\lambda_n^{\nu_n} \tag{24}$$

where $\beta(\nu_1, \cdots, \nu_n)$ is the number of non-zero ν_i's.

Remark: For ${}_{\beta-1}C_{\beta-j}$ to have a non-zero value, β must satisfy

$$j \leq \beta \leq j+k \tag{25}$$

In the above discussion, the eigen-values λ_i's are assumed distinct. When the eigen-values are not distinct, a slight change of A gives distinct eigen-values. As coefficients of expansion vary continuously with the change of matrix A, Theorem 1 holds true when eigen-values are not distinct.

If the minimal polynomial of the matrix is not equal to the characteristic polynomial of A, theorem 1 gives a set of expansion coefficients, though the expansion is not unique. Let the minimal polynomail of A be $\varphi_A(\lambda)$, which is the monic polynomial of order m. In this case, it is apparent that A^{n+k} can be expanded by a polynomial of the order $m-1$. Also, the coefficients are equal to those of the expansion of the $n+k$ th power of an $m\times m$ matrix, whose characteristic polynomial is equal to $\varphi_A(\lambda)$.

5. Upper Bound of Absolute Value of the Coefficient

Suppose that all eigen-values λ_i's are located in a certain disc of radius ρ centered at the origin, i.e.

$$|\lambda_i|\leq\rho \tag{26}$$

By theorem 1, the absolute value of $\alpha_j(k)$ is bounded by

$$|\alpha_j(k)|\leq(-1)^{j-1}\sum_{\nu_1+\cdots+\nu_n=j+k}\binom{\beta(\nu_1,\cdots,\nu_n)-1}{\beta(\nu_1,\cdots,\nu_n)-j}\rho^{j+k} \tag{27}$$

For a fixed value of β which satisfies eq. (25), the number of terms in the r.h.s. of eq. (24) is calculated as follows.
There are ${}_n\mathrm{C}_\beta$ ways to take a set of β kinds of λ_i's from $\{\lambda_1,\cdots,\lambda_n\}$. For every set of β kinds of λ_i's, there are ${}_\beta\mathrm{H}_{j+k-\beta}$ kinds of monomials of the order $j+k$. Therefore,

$$\begin{aligned}|\alpha_j(k)|&\leq\rho^{j+k}\sum_{\beta=j}^{\min(n,j+k)}\binom{n}{\beta}\binom{j+k-1}{j+k-\beta}\binom{\beta-1}{\beta-j}\\&=\rho^{j+k}\binom{j+k-1}{k}\sum_{\beta=j}^{\min(n,j+k)}\binom{n}{\beta}\binom{k}{\beta-j}\end{aligned} \tag{28}$$

To evaluate the r.h.s. of eq. (28), the following lemma is used.

Lemma 5

$$\sum_{r=0}^{p}\binom{n}{r}\binom{m}{p-r}=\binom{n+m}{p}\qquad(p\leq n+m) \tag{29}$$

$$\sum_{r=0}^{m}\binom{n}{\sigma+r}\binom{m}{r}=\binom{m+n}{m+\sigma}\qquad(m+\sigma\leq n) \tag{30}$$

Proof: Eq. (29) is widely known, and is given by the coefficient of x^p, in terms of the following equation.

$$(1+x)^n(1+x)^m = (1+x)^{n+m} \tag{31}$$

When $m+\sigma \leq n$, by eq. (29)

$$\sum_{r=0}^{m}\binom{n}{\sigma+r}\binom{m}{r} = \sum_{r=0}^{n-\sigma}\binom{m}{r}\binom{n}{n-\sigma-r} = \binom{m+n}{m+\sigma} \tag{32}$$

Q.E.D.

Using lemma 5, the upper bound is calculated as follows.

i) In case $n \geq j+k$

$$\begin{aligned} |\alpha_j(k)| &\leq \rho^{j+k}\binom{j+k-1}{k}\sum_{\beta=j}^{j+k}\binom{n}{\beta}\binom{k}{\beta-j} \\ &= \rho^{j+k}\binom{j+k-1}{k}\sum_{\tau=0}^{k}\binom{n}{j+\tau}\binom{k}{\tau} \\ &= \rho^{j+k}\binom{j+k-1}{k}\binom{n+k}{k+j} \end{aligned} \tag{33}$$

ii) In case $n \leq j+k$

$$\begin{aligned} |\alpha_j(k)| &\leq \rho^{j+k}\binom{j+k-1}{k}\sum_{\beta=j}^{n}\binom{n}{\beta}\binom{k}{\beta-j} \\ &= \rho^{j+k}\binom{j+k-1}{k}\sum_{\tau=0}^{n-j}\binom{n}{\tau}\binom{k}{n-j-\tau} \\ &= \rho^{j+k}\binom{j+k-1}{k}\binom{n+k}{k+j} \end{aligned} \tag{34}$$

Both cases give the same results, from which we get the following theorem.

Theorem 2

Suppose all the eigen-values λ_i's of an $n \times n$ matrix A are located in a disc of radius ρ centered at the origin, i.e.,

$$|\lambda_i| \leq \rho \tag{35}$$

Then the absolute value of the coefficient of expansion $\alpha_j(k)$ of A^{n+k} by the form given in eq. (1) is

$$|\alpha_j(k)| \leq \rho^{j+k}\binom{j+k-1}{k}\binom{n+k}{k+j} \tag{36}$$

Remarks: i) Suppose a discrete-time linear dynamical system is given by

$$x(\tau+1) = Ax(\tau) + bu(\tau) \tag{37}$$

If eq. (35) is satisfied, for any positive ε, positive K exists such that

$$||x(\tau)|| < K(\rho+\varepsilon)^{\tau}||x(0)|| \tag{38}$$

where $x(\tau)$ is a zero-input response to an initial value $x(0)$. Therefore, if $\rho<1$, the dynamical system is exponentially stable.

ii) As the reachable subspace of eq. (37) is spanned by vectors of the form $A^i b$ $(i=0, 1, \cdots)$, eq. (36) makes a bridge to connect the exponential stability property, and the structure of the practically reachable subspace of a dynamical system.

6. Conclusions

The coefficient of expansion of an $n\times n$ matrix is evaluated by the form of homogeneous polynomials of eigen-values. Using this expansion formula, the upper bound of the absolute value of the coefficient is also obtained. As the coefficient is expressed by a polynomial of the same sign, this bound is equal to the coefficient of expansion in a case where the matrix is non-cyclic, and all eigen-values are equal and real. Therefore, this bound gives the sharpest bound knowing that all the eigen-values are located in a disc centered at the origin.

The results obtained make a bridge to connect the exponential stability property and the structure of the practically reachable subspace of a dynamical system.

References

1) K. Ando; Preprint of 5th Symposium on Control Theory, p. 73 (1976).
2) P. Whittle; Quart. J. Math. vol. 7, p. 316 (1956).
3) R. Bellman; "Introduction to Matrix Analysis," 3rd edition, McGraw-Hill Book Co., p. 193 (1970).
4) M. Mimura; "Algebra and Geometry," Syokabo Book Co., p. 110 (1956).
5) C. L. Liu; "Introduction to Combinatrial Mathematics," McGraw-Hill Book Co., p. 52 (1968).