МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2006 MATHEMATICS AND EDUCATION IN MATHEMATICS, 2006 Proceedings of the Thirty Fifth Spring Conference of the Union of Bulgarian Mathematicians Borovets, April 5–8, 2006

FRECHÉT DERIVATIVES OF RATIONAL POWER MATRIX FUNCTIONS^{*}

Mihail M. Konstantinov, Juliana K. Boneva, Petko H. Petkov

Let A be a positive definite real or complex matrix. We characterize the Frechét derivatives of matrix valued functions $X \mapsto X^p$ at the point A, where p is a rational number, as special types of Lyapunov operators.

Introduction. In this paper we derive explicit expressions for the Frechét derivatives of the matrix valued function $f_p : \mathbb{S}^{n \times n}_+ \to \mathbb{K}^{n \times n}$, defined by $f_p(X) = X^p$, where $p \in \mathbb{Q}$ and $\mathbb{S}^{n \times n}_+ \subset \mathbb{K}^{n \times n}$ is the set of Hermitian positively definite $n \times n$ matrices. We note that although X^p is well defined for $X \in \mathbb{S}^{n \times n}_+$ and all $p \in \mathbb{R}$, only for rational values of p one may determine reasonably the derivatives of f_p .

The Frechét derivative $f'_p(A)$ of f_p at a given point $A \in \mathbb{S}^{n \times n}_+$ may be used to compute the absolute condition number of the problem $B = f_p(A)$ as the induced norm of the operator $f'_p(A)$. This is important in analyzing the sensitivity of the corresponding computational problem and finding error estimates for the computed solution. We note that when using the Frobenius matrix norm, the induced norm of the operator $f'_p(A)$ is equal to the 2–norm of its matrix.

The Frechét derivatives and condition numbers for the function $X \mapsto a_0 I + a_1 X + a_2 X^2 + \cdots, X \in \mathbb{K}^{n \times n}$, defined by convergent power series, has been studied in [2]. Structured condition numbers (with restrictions on the matrix argument) for such series have been considered in [1].

Further on we use the following general notations: $\mathbb{K}^{n \times n}$ – the space of $n \times n$ matrices over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$; I_n – the identity $n \times n$ matrix; \overline{A} – the complex conjugate of the matrix A; A^{\top} – the transpose of A; $A^{\mathrm{H}} = \overline{A}^{\top}$ – the transposed complex conjugate of A; spect $(A) = \{\lambda_1(A), \lambda_2(A), \dots, \lambda_n(A)\}$ – the full spectrum of the matrix $A \in \mathbb{K}^{n \times n}$, i.e. the collection of its eigenvalues $\lambda_i(A)$, counted according to their algebraic multiplicity; $\mathcal{U}(n) \subset \mathbb{C}^{n \times n}$ – the group of unitary matrices $U \in \mathbb{C}^{n \times n}$ ($U^{\mathrm{H}}U = I_n$); $\mathbb{S}^{n \times n}_+ \subset \mathbb{K}^{n \times n}$ – the set of Hermitian positively definite $n \times n$ matrices; $\lambda_{\max}(A) \ge \lambda_{\min}(A) > 0$ – the maximal and minimal eigenvalues of $A \in \mathbb{S}^{n \times n}_+$; $\operatorname{vec}(A) \in \mathbb{K}^{n^2}$ – the column-wise vector representation of the matrix $A \in \mathbb{K}^{n \times n}$; $P_{n^2} \in \mathbb{R}^{n^2 \times n^2}$ – the vec–permutation matrix

²⁰⁰⁰ Mathematics Subject Classification: 15A24.

This work is supported by Contract 64/16.05.2005 with the Scientific Research Center of UACEG, 1046 Sofia, Bulgaria

such that $\operatorname{vec}(A^{\top}) = P_{n^2}\operatorname{vec}(A), A \in \mathbb{K}^{n \times n}; A \otimes B$ – the Kronecker product of the matrices A and B; $\|\cdot\|_2$ – the Euclidean norm in \mathbb{K}^n or the spectral norm in $\mathbb{K}^{n \times n}; \|\cdot\|_{\mathrm{F}}$ – the Frobenius norm in $\mathbb{K}^{n \times n}$.

Matrix operators. Let $\mathcal{L}(n, \mathbb{K})$ be the space of linear operators $\mathcal{F} : \mathbb{K}^{n \times n} \to \mathbb{K}^{n \times n}$ and $\operatorname{Mat}(\mathcal{F}) \in \mathbb{K}^{n^2 \times n^2}$ be the matrix of the linear operator $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ such that $\operatorname{vec}(\mathcal{F}(A)) = \operatorname{Mat}(\mathcal{F})\operatorname{vec}(A), A \in \mathbb{K}^{n \times n}$. Denote by \mathcal{I} the identity operator on $\mathcal{L}(n, \mathbb{K})$, i.e. $\operatorname{Mat}(\mathcal{I}) = I_{n^2}$. For $M, N \in \mathbb{K}^{n \times n}$ and $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ we denote by $M\mathcal{F}N$ the linear operator, defined by $(M\mathcal{F}N)(X) = M\mathcal{F}(X)N$. Using the identity $\operatorname{vec}(MYN) = (N^{\top} \otimes M)\operatorname{vec}(Y)$, we see that $\operatorname{Mat}(M\mathcal{F}N) = (N^{\top} \otimes M)\operatorname{Mat}(\mathcal{F})$.

For any $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ with Sylvester index r (see [3]) there exist 2r matrices A_k, B_k such that $\mathcal{F}(X) = \sum_{k=1}^r A_k X B_k$ and, hence, $\operatorname{Mat}(\mathcal{F}) = \sum_{r=1}^r B_k^\top \otimes A_k$. More general linear matrix operators $\mathbb{K}^{m \times n} \to \mathbb{K}^{m \times n}$ may be defined by the same formula with $A_k \in \mathbb{K}^{m \times m}$ and $B_k \in \mathbb{K}^{n \times n}$.

For $\mathcal{F}, \mathcal{G} \in \mathcal{L}(n, \mathbb{K})$ let $\mathcal{F} \circ \mathcal{G} \in \mathcal{L}(n, \mathbb{K})$ be the composition of \mathcal{F} and \mathcal{G} , i.e. $\mathcal{F} \circ \mathcal{G}(X) = \mathcal{F}(\mathcal{G}(X))$. We have $\operatorname{Mat}(F \circ \mathcal{G}) = \operatorname{Mat}(\mathcal{F})\operatorname{Mat}(\mathcal{G})$ and $\operatorname{Mat}(\mathcal{F}^{-1}) = (\operatorname{Mat}(\mathcal{F}))^{-1}$ whenever \mathcal{F} is invertible. The norm of the operator \mathcal{F} may be defined as $\|\mathcal{F}\| := \max\{\|\mathcal{F}(X)\|_{\mathrm{F}} : \|X\|_{\mathrm{F}} = 1\} = \max\{\|\operatorname{vec}(\mathcal{F}(X))\|_{2} : \|\operatorname{vec}(X)\|_{2} = 1\} = \max\{\|\operatorname{Mat}(\mathcal{F})\operatorname{vec}(X)\|_{2} : \|\operatorname{vec}(X)\|_{2} = 1\} = \|\operatorname{Mat}(\mathcal{F})\|_{2}.$

The relative condition number of the invertible operator $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ is the quantity $\operatorname{cond}(\mathcal{F}) := \|\mathcal{F}\| \, \|\mathcal{F}^{-1}\| = \|\operatorname{Mat}(\mathcal{F})\|_2 \|(\operatorname{Mat}(\mathcal{F}))^{-1}\|_2$. Hence, $\operatorname{cond}(\mathcal{F}) = \operatorname{cond}_2(\operatorname{Mat}(\mathcal{F}))$.

The nonzero $n \times n$ matrix X is called an *eigenmatrix* of the operator $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ if there is a scalar α such that $\mathcal{F}(X) = \alpha X$. The quantity α is an *eigenvalue* of this operator. Note that for $\mathcal{F} \in \mathcal{L}(n, \mathbb{R})$ the *eigenpair* (α, X) of \mathcal{F} may be complex. The eigenvalues of \mathcal{F} are the eigenvalues of its matrix $Mat(\mathcal{F})$ and, if X is an eigenmatrix of \mathcal{F} , then vec(X) is an eigenvector of $Mat(\mathcal{F})$.

The full spectrum (the collection of eigenvalues counted according to algebraic multiplicity) of the linear matrix operator $\mathcal{F}: \mathbb{K}^{m \times n} \to \mathbb{K}^{m \times n}$, defined by $\mathcal{F}(X) = \sum_{i,j} A^i X B^j$,

where $A \in \mathbb{K}^{m \times m}$ and $B \in \mathbb{K}^{n \times n}$ are fixed matrices, is

(1)
$$\left\{\sum_{i,j}\lambda^{i}\mu^{j}:\lambda\in\operatorname{spect}(A),\ \mu\in\operatorname{spect}(B)\right\}.$$

We recall [3, 4] that $\mathcal{F} \in \mathcal{L}(n, \mathbb{K})$ is a Lyapunov operator if $(\mathcal{F}(A))^{\mathrm{H}} = \mathcal{F}(A^{\mathrm{H}})$, $A \in \mathbb{K}^{n \times n}$. The necessary and sufficient condition for \mathcal{F} to be a Lyapunov operator is $P_{n^2} \operatorname{Mat}(\mathcal{F}) = \operatorname{Mat}(\mathcal{F})P_{n^2}$. We note that the sets of Hermitian and skew–Hermitian (resp. symmetric and skew–symmetric) matrices are invariant sets for complex (resp. real) Lyapunov operators.

Problem statement. Let $A \in \mathbb{S}^{n \times n}_+$ and $p \in \mathbb{R}$. Then the matrix A^p is correctly defined as follows. There exists a matrix $U \in \mathcal{U}(n)$ such that $A = U\Lambda U^{\mathrm{H}}$, where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ are the eigenvalues of A. Here Λ is the (diagonal) Schur form of A. Now we may define the power p of A as $A^p := U\Lambda^p U^{\mathrm{H}}$, $\Lambda^p := \operatorname{diag}(\lambda_1^p, \lambda_2^p, \ldots, \lambda_n^p)$. Here λ_k^p is the positive (real) p-th degree of $\lambda_k > 0$. 170 In a more general setting suppose that the matrix A has positive eigenvalues and diagonal Jordan form $\Lambda = V^{-1}XV$, where V is an invertible matrix. Then we may define $A^p := V\Lambda^p V^{-1}$.

In order to obtain meaningful results we restrict ourselves to the case when $p = \pm r/s$ is a rational number, where $r, s \in \mathbb{N}$ are coprime. To define correctly the matrix quantity $(A + E)^p$, where $A \in \mathbb{S}^{n \times n}_+$ and $E \in \mathbb{K}^{n \times n}$ is a given increment, we may suppose that Eis Hermitian and $A + E \in \mathbb{S}^{n \times n}_+$. In particular, we may assume that E varies over the set $\mathcal{E}_{\alpha} \subset \mathbb{K}^{n \times n}$ of all Hermitian matrices with $\|E\|_2 < \alpha := \lambda_{\min}(A)$.

Both conditions $E = E^{\mathrm{H}}$ and $||E||_2 < \alpha$ are in certain sense necessary in order to define $(A+E)^p$. Indeed, there exist non-Hermitian and arbitrarily small matrices E such that A + E has complex eigenvalues and there arises the problem to choose a suitable branch of the root $(A+E)^{1/s}$ when $1 < s \in \mathbb{N}$. Let for example $A = I_2$ and $E = [e_{ik}]$ be anti-diagonal with $e_{12} = -e_{21} = \varepsilon$, where $\varepsilon > 0$ is a small parameter. Then A + E has complex eigenvalues $1 \pm i\varepsilon$, where $i^2 = -1$. Conversely, if $||E||_2 \ge \alpha$, then the matrix A + E may have zero eigenvalues and its negative powers will not exist.

Under the above restrictions we may define the Frechét derivative $\mathcal{F}(p, A) := f'_p(A)$ of the function f_p at the point $A \in \mathbb{S}^{n \times n}_+$ as the linear operator such that $f_p(A + E) = (A + E)^p = A^p + \mathcal{F}(p, A)(E) + O(||E||^2), E \to 0, E \in \mathcal{E}_{\alpha}$. Note that for non-integer values of p the operator $\mathcal{F}(p, A)$ is defined only on the set \mathcal{E}_{α} .

Main results. We shall consider successively the five cases p = r, p = -r, p = 1/s, p = r/s and p = -r/s, where $r, s \in \mathbb{N}$, giving explicit expressions for the operator $\mathcal{F}(p, A)$. For the first two cases we also derive expressions for the eigenvalues and condition numbers of this operator.

(i) The case
$$p = r, r \in \mathbb{N}$$
. Since $(A + E)^r = A^r + \sum_{k=0}^{r-1} A^{r-1-k} E A^k + O(||E||^2)$ it

follows that $\mathcal{F}(r, A)(E) = \sum_{k=0}^{r-1} A^{r-1-k} E A^k$ and hence

(2)
$$\mathcal{F}(r,A) = \sum_{k=0}^{r-1} A^{r-1-k} \mathcal{I} A^k.$$

Note that $\mathcal{F}(r, A) \in \mathcal{L}(n, \mathbb{K})$, i.e. $\mathcal{F}(r, A)(E)$ is defined for all $E \in \mathbb{K}^{n \times n}$.

For r = 1 we have the trivial result $\mathcal{F}(1, A)(E) = E$, i.e. $\mathcal{F}(1, A)$ does not depend on A and is equal to \mathcal{I} . Since the matrix A is Hermitian $(A^{\mathrm{H}} = A)$, we see that for r = 2 the operator $\mathcal{F}(2, A)$, given by $\mathcal{F}(2, A)(E) = AE + EA = A^{\mathrm{H}}E + EA$, is the standard Lyapunov operator arising in the theory of linear continuous-time systems. Moreover, $\mathcal{F}(r, A)$ is a Lyapunov operator for all $r \in \mathbb{N}$.

For $r, s \in \mathbb{N}$ we have

$$(A+E)^{r+s} = A^{r+s} + A^r \mathcal{F}(s,A)(E) + \mathcal{F}(r,A)(E)A^s + \mathcal{O}(||E||^2), \ E \to 0$$

Hence

(3)
$$\mathcal{F}(r+s,A) = A^r \mathcal{F}(s,A) + \mathcal{F}(r,A)A^s = A^s \mathcal{F}(r,A) + \mathcal{F}(s,A)A^r.$$

The matrix of the operator $\mathcal{F}(r,A)$ is given by

(4)
$$\operatorname{Mat}(\mathcal{F}(r,A)) = \sum_{k=0}^{r-1} (A^{\top})^k \otimes A^{r-1-k} = \sum_{k=0}^{r-1} \overline{A}^k \otimes A^{r-1-k}.$$
171

Since $A \in \mathbb{S}^{n \times n}_+$ the matrix $Mat(\mathcal{F}(r, A))$ is Hermitian. The collection of its eigenvalues (i.e. the eigenvalues of $\mathcal{F}(r, A)$) according to (1) is

spect
$$(\mathcal{F}(r,A)) = \left\{ \sum_{k=0}^{r-1} \lambda_i^k \lambda_j^{r-1-k} : i, j = 1, 2, \dots, n \right\}$$

Suppose that the spectrum of A consists of m pairwise distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ with algebraic multiplicities k_1, k_2, \ldots, k_m , respectively. Then the full spectrum spect $(\mathcal{F}(r, A))$ of $\mathcal{F}(r, A)$ consists of n^2 elements, namely of k_i^2 quantities $r\lambda_i^{r-1}$, $i = 1, 2, \ldots, m$ and $2k_ik_j$ quantities $(\lambda_i^r - \lambda_j^r)/(\lambda_i - \lambda_j)$, $i = 1, 2, \ldots, r-1$, $j = i+1, i+2, \ldots, r$. In particular

 $\lambda_{\min}(\operatorname{Mat}(\mathcal{F}(r,A))) = r(\lambda_{\min}(A))^{r-1}, \ \lambda_{\max}(\operatorname{Mat}(\mathcal{F}(r,A))) = r(\lambda_{\max}(A))^{r-1}$ and hence

 $\operatorname{cond}(\mathcal{F}(r,A)) = \operatorname{cond}_2(\operatorname{Mat}(\mathcal{F}(r,A))) = (\lambda_{\max}(A)/\lambda_{\min}(A))^{r-1} = (\operatorname{cond}_2(A))^{r-1}.$

The composition of two operators $\mathcal{F}(r, A)$ and $\mathcal{F}(s, A)$ is a Lyapunov operator (for $A = A^{\mathrm{H}}$) and has the form

$$\mathcal{F}(r,A) \circ \mathcal{F}(s,A) = \mathcal{F}(s,A) \circ \mathcal{F}(r,A) = \sum_{k=0}^{t-1} A^k \mathcal{F}(r+s-1-2k,A) A^k, \ t := \min\{r,s\}.$$

Consider the matrix valued function $X \mapsto g(X) := \sum_{r=0}^{\infty} a_r X^r$, $||X|| < \rho$. The Frechét derivative $g'(A) \in \mathcal{L}(n, \mathbb{K})$ of G at any point A, where A is arbitrary with $||A|| < \rho$, is $g'(A) = \sum_{r=1}^{\infty} a_r \mathcal{F}(r, A)$. Under the above assumptions on the spectrum of A it may be shown that the full spectrum of g'(A) consists of n^2 elements, namely of the k_i^2 numbers $g'(\lambda_i), i = 1, 2, \ldots, m$ and the $2k_i k_j$ numbers $(g(\lambda_i) - g(\lambda_j))/(\lambda_i - \lambda_j), i = 1, 2, \ldots, r-1, j = i + 1, i + 2, \ldots, r$.

(ii) The case $p = -r, r \in \mathbb{N}$. We have

$$(A+E)^{p} = A^{-r} - A^{-r} \mathcal{F}(r,A)(E)A^{-r} + \mathcal{O}(||E||^{2}), \ E \to 0.$$

Therefore, $\mathcal{F}(-r,A)(E) = -A^{-r} \mathcal{F}(r,A)(E)A^{-r} = -\sum_{r=1}^{r-1} A^{-1-k} E A^{k-r}$, or

(5) $\mathcal{F}(-r,A) = -A^{-r}\mathcal{F}(r,A)A^{-r} = \mathcal{F}(-1,A^r) \circ \mathcal{F}(r,A).$

In particular, we have the relations $\mathcal{F}(-1, A) = -A^{-1}\mathcal{I}A^{-1}$, $\mathcal{F}(-1, A^{-1}) = -A\mathcal{I}A$ and $\mathcal{F}(-1, A) \circ \mathcal{F}(-1, A^{-1}) = \mathcal{I}$.

The operator $\mathcal{F}(-r, A)$ is a Lyapunov operator. Its matrix and its full spectrum are given by $\operatorname{Mat}(\mathcal{F}(-r, A)) = -\sum_{k=0}^{r-1} \overline{A}^{k-r} \otimes A^{-1-k}$ and

spect
$$(\mathcal{F}(-r,A)) = \left\{ -\sum_{k=0}^{r-1} \lambda_i^{k-r} \lambda_j^{-1-k} : i, j = 1, 2, \dots, n \right\}.$$

In particular, we have

 $\lambda_{\min}(\operatorname{Mat}(\mathcal{F}(-r,A))) = r(\lambda_{\max}(A))^{-r-1}, \lambda_{\max}(\operatorname{Mat}(\mathcal{F}(-r,A))) = r(\lambda_{\min}(A))^{-r-1}.$ 172

Therefore,

$$\operatorname{cond}(\mathcal{F}(-r,A)) = \operatorname{cond}_2(\operatorname{Mat}(\mathcal{F}(-r,A)))$$
$$= (\lambda_{\max}(A)/\lambda_{\min}(A))^{r+1} = (\operatorname{cond}(A))^{r+1}.$$

(iii) The case p = 1/s, $s \in \mathbb{N}$. Here we may take the *s*-th power from both sides of the equality $(A + E)^{1/s} = A^{1/s} + \mathcal{F}(1/s, A)(E) + O(||E||^2)$, $E \to 0$, in order to obtain $A + E = A + \mathcal{F}(s, A^{1/s})(\mathcal{F}(1/s, A)(E)) + O(||E||^2)$, $E \to 0$. Hence, $\mathcal{F}(s, A^{1/s}) \circ \mathcal{F}(1/s, A) = \mathcal{I}$ and

(6)
$$\mathcal{F}(1/s, A) = (\mathcal{F}(s, A^{1/s}))^{-1}.$$

In this case the operator $\mathcal{F}(1/s, A)$, being the inverse of the Lyapunov operator $\mathcal{F}(s, A^{1/s})$, is again a Lyapunov operator, see [3].

(iv) The case
$$p = r/s$$
; $r, s \in \mathbb{N}$. It follows from the previous results that $(A+E)^{r/s} = A^{r/s} + \mathcal{F}(r, A^{1/s})(\mathcal{F}(1/s, A)(E)) + O(||E||^2).$

Hence,

(7)
$$\mathcal{F}(r/s, A) = \mathcal{F}(r, A^{1/s}) \circ \mathcal{F}(1/s, A) = \mathcal{F}(r, A^{1/s}) \circ (\mathcal{F}(s, A^{1/s}))^{-1}.$$

(v) The case p = -r/s; $r, s \in \mathbb{N}$. Since

$$(A+E)^{-r/s} = A^{-r/s} - A^{-r/s} \mathcal{F}(r/s, A)(E) A^{-r/s} + \mathcal{O}(||E||^2),$$

we obtain

(8) $\mathcal{F}(-r/s, A) = -A^{-r/s}\mathcal{F}(r/s, A)A^{-r/s} = \mathcal{F}(-1, A^{r/s}) \circ \mathcal{F}(r/s, A).$ Thus we have proved the following result.

Theorem 1. For all cases of rational degrees p described in (i), (ii), (iii), (iv) and (v), the operator $\mathcal{F}(p, A)$ is given by the relations (2), (5), (6), (7) and (8), respectively.

REFERENCES

 P. DAVIES. Structured conditioning of matrix functions. *Elect. J. Linear Algebra*, **11** (2004), 132-161, PDF text available at www.emis.de/journals/ELA/ela-articles/11.html.

[2] C. KENNEY, A. LAUB. Condition estimates for matrix functions. *SIAM J. Matrix Anal. Appl.*, **10** (1989), 191–209.

[3] M. KONSTANTINOV, V. MEHRMANN, P. PETKOV. On properties of general Sylvester and Lyapunov operators. *Linear Algebra Appl.*, **312** (2000), 35–71.

[4] M. KONSTANTINOW, D. GU, V. MEHRMANN, P. PETKOV. Perturbation Theory for Matrix Equations. North–Holland Publ. Co., Amsterdam, 2003.

Mihail Konstantinov	Juliana Boneva
UACEG	UACEG
1046 Sofia, Bulgaria	1046 Sofia, Bulgaria
e-mail:mmk_fte@uacg.bg	e-mail: boneva_fte@uacg.bg

Petko Petkov TU - Sofia 1756 Sofia, Bulgaria e-mail: php@tu-sofia.bg

ПРОИЗВОДНИ НА ФРЕШЕ НА РАЦИОНАЛНО–СТЕПЕННИ МАТРИЧНИ ФУНКЦИИ

Михаил М. Константинов, Юлиана К. Бонева, Петко Х. Петков

Нека A е положително дефинитна реална или комплексна матрица. Описани са производните на Фреше на матричните функции $X \mapsto X^p$ в точката A, където p е рационално число, като специален тип оператори на Ляпунов.