J. Math. Anal. Appl. 433 (2016) 1459–1485

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Derivatives of compound matrix valued functions ☆

Brendan P.W. Ames [a,*], Hristo S. Sendov [b]

[a] Department of Mathematics, The University of Alabama, Tuscaloosa, AL 35487, USA
[b] Department of Statistical and Actuarial Sciences, The University of Western Ontario, 1151 Richmond Str., London, ON, N6A 5B7, Canada

ARTICLE INFO

Article history:
Received 5 June 2015
Available online 14 August 2015
Submitted by A. Daniilidis

Keywords:
Primary matrix functions
Spectral functions
Compound matrix
Differentiability
Multiplicative compound matrix
Additive compound matrix

ABSTRACT

Primary matrix functions and spectral functions are two classes of orthogonally invariant functions on a symmetric matrix argument. Many of their properties have been investigated thoroughly and find numerous applications both theoretical and applied in areas ranging from engineering, image processing, optimization and physics. We propose a family of maps that provide a natural connection and generalization of these two classes of functions. The family of maps also contains the well-known multiplicative and additive compound matrices. We explain when each member of this family is a differentiable function and exhibit a formula for its derivative.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Denote by $\mathbf{R}^{n\times n}$ the space of all $n\times n$ matrices with real entries and by S^n the Euclidean space of all $n\times n$ symmetric matrices with inner product $\langle X,Y\rangle := \mathrm{tr}\,(XY)$ and corresponding norm $\|X\| := \sqrt{\mathrm{tr}\,(X^2)}$. Let O^n denote the set of $n\times n$ orthogonal matrices. That is, $X\in O^n$ if $XX^T = I$. Let $\mathbf{R}^n_{\geq}$ be the convex cone in $\mathbf{R}^n$ of all vectors with nonincreasing components, i.e., $x\in\mathbf{R}^n_{\geq}$ implies $x_1\geq x_2\geq\cdots\geq x_n$. We denote by $\lambda(X)\in\mathbf{R}^n_{\geq}$ the vector of eigenvalues of $X\in S^n$ in nonincreasing order:

$$\lambda_1(X)\geq\lambda_2(X)\geq\cdots\geq\lambda_n(X).$$

For a vector $x\in\mathbf{R}^n$, denote by $\mathrm{Diag}\,x$ the $n\times n$ matrix with x on its main diagonal and zeros everywhere else. For a matrix $X\in S^n$, denote by $\mathrm{diag}\,X$ the vector in $\mathbf{R}^n$ of diagonal entries of X.

☆ The work presented in this paper was carried out while B.P.W Ames was a postdoctoral fellow at the Institute for Mathematics and its Applications during the IMA's annual program on Mathematics of Information (supported by National Science Foundation (NSF) award DMS-0931945) and while B.P.W. Ames was a Von Karman instructor at the California Institute of Technology supported by Joel Tropp under Office of Naval Research (ONR) award N00014-11-1002.

* Corresponding author.
E-mail addresses: bpames@ua.edu (B.P.W. Ames), hssendov@stats.uwo.ca (H.S. Sendov).

http://dx.doi.org/10.1016/j.jmaa.2015.08.029
0022-247X/© 2015 Elsevier Inc. All rights reserved.

In the last several decades, interest in functions on the space of symmetric matrices has exploded. Two of the more prominent classes of functions are the following primary matrix functions and the spectral functions.

Definition 1.1. Every function $f : \mathbf{R} \to \mathbf{R}$ defines a *primary matrix function* $F : S^n \to S^n$ by

$$F(X) = U\big(\mathrm{Diag}\,(f(\lambda_1(X)), \dots, f(\lambda_n(X)))\big)U^T,$$

where $U \in O^n$ is such that $X = U(\mathrm{Diag}\,\lambda(X))U^T$.

Extensive treatment of the primary matrix functions is given in [9, Chapter 6] and in [1, Chapter V]. The primary matrix functions are also known as Lowener operators. Different variational properties of the primary matrix functions have been of main interest. The classical works [4] and [3], when restricted to the space of real symmetric matrices, describe the k-th order derivative of the function $t \in \mathbf{R} \to F(X(t))$, where $X(t) \in S^n$ is a k-times continuously differentiable curve of symmetric matrices. This is known as the Daleckii–Krein formula. One can find a surprisingly different approach in [14], where Theorem 4.1 removes the requirement that $\frac{d^k}{dt^k}X(t)$ be continuous in order to conclude that $t \in \mathbf{R} \to F(X(t))$ is k-times differentiable. Naturally, treating the whole matrix X as an independent variable, brings larger generality. We say a function is *symmetric* if its value is invariant under permutations of its arguments.

Definition 1.2. A real-valued function $F : S^n \to \mathbf{R}$ is a *spectral function* if there exists symmetric function $f : \mathbf{R}^n \to \mathbf{R}$ such that $F(X) = (f \circ \lambda)(X)$ for all $X \in S^n$.

Spectral functions have received much attention in the last twenty years and find numerous applications both theoretical and applied. The literature on the subject is vast and goes out of the scope of this paper to survey it thoroughly. In engineering they often appear as constraints in feasibility problems (for example, in the design of tight frames [22] or in the design of low-rank controllers [17] in control). Among the more notable optimization related works are [2] and [20]. For recent theoretical developments see [5] and [6]. Works dealing with higher-order differentiability of spectral functions are [23,21,19,18]. In engineering literature, spectral functions are also known as isotropic functions.

One of the main results in [10], among other places, states the following.

Theorem 1.1. *The spectral function $F(X) = (f \circ \lambda)(X)$ is differentiable at X if and only if f is differentiable at $\lambda(X)$ and in that case*

$$\nabla F(X) = U^T\big(\nabla f(\lambda(X))\big)U^T, \tag{1}$$

where $U \in O^n$ is such that $X = U(\mathrm{Diag}\,\lambda(X))U^T$.

The type of orthogonal invariance possessed by the primary matrix functions and the spectral functions indicates that there must be a natural connection between them. One such connection is a consequence of Theorem 1.1. Indeed, consider the separable symmetric function

$$f(x_1, \dots, x_n) = g(x_1) + \cdots + g(x_n),$$

where $g : \mathbf{R} \to \mathbf{R}$. Then, the derivative of the corresponding spectral function $F(X) = (f \circ \lambda)(X)$ is

$$\nabla F(X) = U\big(\mathrm{Diag}\,(g'(\lambda_1(X)), \dots, g'(\lambda_n(X)))\big)U^T, \tag{2}$$

Download English Version:

https://daneshyari.com/en/article/4614937

Download Persian Version:

https://daneshyari.com/article/4614937

Daneshyari.com