Wavelets and Linear Algebra 5(1) (2018) 49- 56

Vali-e-Asr University of Rafsanjan

Wavelets and Linear Algebra

http://wala.vru.ac.ir

Some Results on Convex Spectral Functions: I

AliReza Sattarzadeh[a], Hossein Mohebi[b,*]

[a]*Department of Mathematics, Graduate University of Advanced Technology, Kerman, Islamic Republic of Iran.*
[b] *Department of Mathematics and Mahani Mathematical Research Center, Shahid Bahonar University of Kerman, Kerman, Islamic Republic of Iran.*

Article Info

Article history:
Received 20 June 2017
Accepted 1 August 2017
Available online 3 January 2018
Communicated by Mohammad Ali Dehghan

Keywords:
Spectral function,
Proximal average,
Moreau envelope,
Infimal convolution.

2000 MSC:
15A18, 49J52, 47A75.

Abstract

In this paper, we give a fundamental convexity preserving for spectral functions. Indeed, we investigate infimal convolution, Moreau envelope and proximal average for convex spectral functions, and show that this properties are inherited from the properties of its corresponding convex function. This results have many applications in Applied Mathematics such as semi-definite programmings and engineering problems.

1. Introduction

There has been growing interest in the variational analysis of spectral functions. This growing trend is due to spectral functions that have important applications to some fundamental problems

*Corresponding author
Email addresses: arsattarzadeh@gmail.com (AliReza Sattarzadeh), hmohebi@uk.ac.ir (Hossein Mohebi)

http://doi.org/10.22072/wala.2017.66630.1123

in applied mathematics such as semi-definite programming and engineering problems (see [5, 7], and references therein).
A function F defined on $\mathcal{S}_n$ is called spectral if

$$F(U^T AU) = F(A),\ \forall\, A \in \mathcal{S}_n,\ \forall\, U \in \mathcal{O}_n,$$

where $\mathcal{S}_n$ is the vector space of all $n \times n$ real symmetric matrices and $\mathcal{O}_n$ is the group of all real orthogonal matrices.
One can easily see [7] that every spectral function is the composition of a symmetric function f defined on $\mathbb{R}^n$ and the eigenvalue function $\lambda : \mathcal{S}_n \longrightarrow \mathbb{R}^n$, i.e.,

$$F(A) = (f \circ \lambda)(A),\ \forall\, A \in \mathcal{S}_n.$$

Hence there exists a one-to-one correspondence between the spectral functions F defined on $\mathcal{S}_n$ and the symmetric functions f defined on $\mathbb{R}^n$. In recent years a lot of research shows that the properties of F are inherited from the properties of f, and vice versa [5, 6, 7, 8, 9, 10, 11, 12]. For example, lower semi-continuity and differentiability of F at a point $A \in \mathcal{S}_n$ are inherited from lower semi-continuity and differentiability of f at the point $\lambda(A) \in \mathbb{R}^n$, and vice versa. Moreover, in [7] the conjugate and the subdifferential of F has been characterized in terms of the conjugate and the subdifferential of f.
This paper is devoted to a fundamental convexity preserving for spectral functions. Indeed, we show that the following properties of spectral functions hold:

(1) The infimal convolution of spectral functions $f \circ \lambda$ and $g \circ \lambda$ is spectrally defined in terms of f and g.

(2) The Moreau envelope of spectral functions is equal to spectral functions of Moreau envelope.

(3) The proximal average of $f \circ \lambda$ and $g \circ \lambda$ is spectrally defined in terms of proximal average of f and g.

2. Preliminaries

We denote by $\mathcal{E}$ the Euclidean space $\mathbb{R}^n$ with the inner product $\langle ., . \rangle$ and the induced norm $\|.\|$. For a function $f : \mathcal{E} \longrightarrow \bar{\mathbb{R}} := [-\infty, +\infty]$, define the domain of f by

$$dom(f) := \Big\{x \in \mathcal{E} : f(x) < +\infty\Big\}.$$

We say that f is proper if $dom(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in \mathcal{E}$. The set of all proper lower semi-continuous and convex functions defined on $\mathcal{E}$ with values in $\bar{\mathbb{R}}$ is denoted by $\Gamma_0(\mathcal{E})$. The epigraph of $f : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ is defined by

$$epi(f) := \Big\{(x, \alpha) \in \mathcal{E} \times \mathbb{R} : f(x) \leq \alpha\Big\}.$$

The Fenchel-Moreau conjugate [4, 13] of a function $f : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ is defined by $f^* : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$

$$f^*(x) := \sup_{y \in \mathcal{E}} \Big\{\langle x, y \rangle - f(y)\Big\}, \qquad \forall\, x \in \mathcal{E},$$

and the second conjugate (or bi-conjugate) [4, 13] of f is defined by

$$f^{**}(x) := \sup_{y\in\mathcal{E}}\Big\{\langle x, y\rangle - f^{*}(y)\Big\}, \qquad \forall\, x \in \mathcal{E}.$$

In the following, we state some properties of the conjugate.

Lemma 2.1. [1, Chapter 13]

(*i*) Let $f : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ be a proper function. Then, $f^{*} \in \Gamma_0(\mathcal{E})$.

(*ii*) Let $f, g : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ be proper functions such that $f \le g$. Then, $g^{*} \le f^{*}$.

(*iii*) $f \in \Gamma_0(\mathcal{E})$ if and only if $f(x) = f^{**}(x)$ for all $x \in \mathcal{E}$.

Definition 2.1. [1] Let $f, g : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ be proper functions. The infimal convolution of f and g is defined by $f \oplus g : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$

$$(f \oplus g)(x) := \inf_{y\in\mathcal{E}}\Big\{f(y) + g(x - y)\Big\}, \qquad \forall\, x \in \mathcal{E}.$$

Proposition 2.1. [1]

(*i*) Let $f, g : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ be proper functions. Then, $(f \oplus g)^{*} = f^{*} + g^{*}$.

(*ii*) Let $f, g \in \Gamma_0(\mathcal{E})$. Assume that $epi(f^{*})+epi(g^{*})$ is a closed subset of $\mathcal{E}\times\mathbb{R}$. Then, $f\oplus g \in \Gamma_0(\mathcal{E})$ and $(f + g)^{*} = f^{*} \oplus g^{*}$.

Definition 2.2. [1] Let $f : \mathcal{E} \longrightarrow \bar{\mathbb{R}}$ be a proper function, and let $\gamma > 0$. The Moreau envelope of f with the parameter $\gamma > 0$ is defined by

$$f^{\gamma}(x) := \inf\nolimits_{y\in\mathcal{E}}\Big\{f(y) + \tfrac{1}{2\gamma}\|x - y\|^2\Big\}, \qquad \forall\, x \in \mathcal{E}. \tag{2.1}$$

Definition 2.3. [1, 2] Let f and g be in $\Gamma_0(\mathcal{E})$. The proximal average of f and g is defined by $pav(f, g) : \mathcal{E} \longrightarrow \mathbb{R} \cup \{+\infty\}$

$$pav(f, g)(x) := \frac{1}{2} \inf_{\substack{(y,z)\in\mathcal{E}\times\mathcal{E}\\ y+z=2x}} \Big\{f(y) + g(z) + \frac{1}{4}\|y - z\|^2\Big\}, \qquad \forall\, x \in \mathcal{E}. \tag{2.2}$$

In the following proposition, we state some properties of proximal average of f and g. For more details and its proof, see [1, Chapter 14].

Proposition 2.2. [1] Let f and g be in $\Gamma_0(\mathcal{E})$. Then, the following assertions are true.

(*i*) $pav(f, g) = pav(g, f)$.

(*ii*) $dom(pav(f, g)) = \dfrac{1}{2}dom(f) + \dfrac{1}{2}dom(g)$.

(*iii*) $pav(f, g) \in \Gamma_0(\mathcal{E})$.

(iv) $\left(pav(f,g)\right)^* = pav(f^*, g^*)$.

Let $\mathcal{S}_n$ be the vector space of all $n \times n$ real symmetric matrices. We denote by $\mathcal{O}_n$ the group of all real orthogonal matrices. We endow $\mathcal{S}_n$ with the trace inner product [3]:

$$\langle A, B \rangle := tr(AB), \quad \forall\, A, B \in \mathcal{S}_n.$$

This inner product induces the Frobenius norm [3], i.e., $\|A\|_F = \sqrt{tr(A^2)}$. For any $x \in \mathbb{R}^n$, we denote by the symbol $Diag(x)$ the $n \times n$ matrix with components of x on its diagonal and with zero off the diagonal.
Define the eigenvalue function $\lambda : \mathcal{S}_n \longrightarrow \mathbb{R}^n$ by $\lambda(A) := (\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A))$ for each $A \in \mathcal{S}_n$, where $\lambda_1(A), \lambda_2(A), \ldots, \lambda_n(A)$ are the eigenvalues of A and ordered in a non-increasing order, i.e., $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$. The following theorem due to von Neumann plays a central role in the spectral variation analysis.

Theorem 2.1. [5, 7] For any $A, B \in \mathcal{S}_n$, we have

$$\left\|\lambda(A) - \lambda(B)\right\| \leq \left\|A - B\right\|_F, \tag{2.3}$$

and

$$\langle A, B \rangle \leq \left\langle \lambda(A), \lambda(B) \right\rangle. \tag{2.4}$$

Any $A \in \mathcal{S}_n$ admits a spectral decomposition of the form $A = UDiag(\lambda(A))U^T$ for some $U \in \mathcal{O}_n$. For each $A \in \mathcal{S}_n$, define the set of all orthogonal matrices giving the ordered spectral decomposition of A by

$$\mathcal{O}_A := \left\{U \in \mathcal{O}_n : U^T A U = Diag(\lambda(A))\right\}.$$

It is clear that $\mathcal{O}_A$ is non-empty for each $A \in \mathcal{S}_n$.
A function $F : \mathcal{S}_n \longrightarrow \bar{\mathbb{R}}$ is called spectral if F is $\mathcal{O}_n$-invariant, i.e.,

$$F(U^T A U) = F(A), \ \forall\, A \in dom(F), \ \forall\, U \in \mathcal{O}_n.$$

It is not difficult to see [7] that any spectral function F defined on $\mathcal{S}_n$ can be written as a composition $f \circ \lambda$ for some symmetric function f defined on $\mathbb{R}^n$ (a function $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ is called symmetric if $f(x) = f(Px)$ for all permutation matrices P and for all $x \in \mathbb{R}^n$). For instance, it is well-known that for each $A \in \mathcal{S}_n$,

$$\left\|A\right\|_F^2 = \sum_{i=1}^{n} [\lambda_i(A)]^2 = \left\|\lambda(A)\right\|^2,$$

i.e.,

$$\left\|A\right\|_F = (\|.\| \circ \lambda)(A).$$

The above relation shows that the Frobenius norm is a spectral function defined on $\mathcal{S}_n$ associated with the standard Euclidean norm on $\mathbb{R}^n$.
The following results present convexity and conjugacy of the spectral function.

Theorem 2.2. [7, Corollary 2.4] Let $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ be a proper symmetric function. Then, $f \in \Gamma_0(\mathbb{R}^n)$ if and only if $f \circ \lambda \in \Gamma_0(\mathcal{S}_n)$.

Theorem 2.3. [7, Theorem 2.3] Let $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ be a proper symmetric function. Then,

$$(f \circ \lambda)^*(A) = f^* \circ \lambda(A), \ \forall\, A \in \mathcal{S}_n. \tag{2.5}$$

3. Main Results

In this section we prove the properties (1), (2) and (3) of spectral functions, which given in Page 2. Let $f, g : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ be symmetric functions, and let $\gamma > 0$. It is easy to see that f^γ and $f \oplus g$ are symmetric functions.

Theorem 3.1. Let $f, g \in \Gamma_0(\mathbb{R}^n)$ be symmetric functions and let $\lambda : \mathcal{S}_n \longrightarrow \mathbb{R}^n$ be the eigenvalue function. Assume that $epi(f^*) + epi(g^*)$ is a closed subset of $\mathbb{R}^n \times \mathbb{R}$ and $epi(f^* \circ \lambda) + epi(g^* \circ \lambda)$ is a closed subset of $\mathcal{S}_n \times \mathbb{R}$. Then,

$$(f \oplus g) \circ \lambda(A) = \Big(f \circ \lambda \oplus g \circ \lambda\Big)(A), \ \forall\, A \in \mathcal{S}_n.$$

Proof: Let $A \in \mathcal{S}_n$ be arbitrary. In view of Theorem 2.3 and Proposition 2.1(i), one has

$$\Big((f \oplus g) \circ \lambda\Big)^*(A) = (f \oplus g)^* \circ \lambda(A) = (f^* + g^*) \circ \lambda(A) = \Big(f^* \circ \lambda + g^* \circ \lambda\Big)(A). \tag{3.1}$$

Therefore, it follows from (3.1), Lemma 2.1(iii), Proposition 2.1(ii) and Theorem 2.3 that

$$\begin{aligned} (f \oplus g) \circ \lambda(A) \quad &= \Big((f \oplus g) \circ \lambda\Big)^{**}(A) = \Big(f^* \circ \lambda + g^* \circ \lambda\Big)^*(A) \\ &= \Big((f^* \circ \lambda)^* \oplus (g^* \circ \lambda)^*\Big)(A) = \Big(f \circ \lambda \oplus g \circ \lambda\Big)(A), \end{aligned}$$

which completes the proof. ■

Theorem 3.2. Let $f : \mathbb{R}^n \longrightarrow \bar{\mathbb{R}}$ be a proper symmetric function, $\lambda : \mathcal{S}_n \longrightarrow \mathbb{R}^n$ be the eigenvalue function, and let $\gamma > 0$. Then,

$$(f \circ \lambda)^\gamma(A) = f^\gamma \circ \lambda(A), \ \forall\, A \in \mathcal{S}_n. \tag{3.2}$$

Proof: Let $A \in \mathcal{S}_n$ be arbitrary. First, note that it follows from (2.1) that

$$(f \circ \lambda)^\gamma(A) = \inf_{B \in \mathcal{S}_n} \Big\{f(\lambda(B)) + \frac{1}{2\gamma}\|B - A\|_F^2\Big\},$$

and

$$(f^\gamma \circ \lambda)(A) = \inf_{y \in \mathbb{R}^n} \Big\{f(y) + \frac{1}{2\gamma}\|y - \lambda(A)\|^2\Big\}.$$

Now, let $B \in \mathcal{S}_n$ be arbitrary. In view of (2.3), one has

$$f(\lambda(B)) + \frac{1}{2\gamma}\|B - A\|_F^2 \geq f(\lambda(B)) + \frac{1}{2\gamma}\|\lambda(B) - \lambda(A)\|^2 \geq f^\gamma \circ \lambda(A).$$

Taking infimum over all $B \in \mathcal{S}_n$, we conclude that

$$(f \circ \lambda)^\gamma(A) \geq f^\gamma \circ \lambda(A).$$

For the converse of (3.2), let $y \in \mathbb{R}^n$ be arbitrary and let $A \in \mathcal{O}_A$. Consider

$$\begin{aligned}
f(y) + \frac{1}{2\gamma}\left\|y - \lambda(A)\right\|^2 \quad &= f\Big(\lambda(Diag(y))\Big) + \frac{1}{2\gamma}\left\|Diag(y) - Diag(\lambda(A))\right\|_F^2 \\
&= f\Big(\lambda(Diag(y))\Big) + \frac{1}{2\gamma}\left\|Diag(y) - U^TAU\right\|_F^2 \\
&= f\Big(\lambda(Diag(y))\Big) + \frac{1}{2\gamma}\left\|Diag(y) - A\right\|_F^2 \\
&\geq (f \circ \lambda)^\gamma(A).
\end{aligned}$$

Now, taking infimum over all $y \in \mathbb{R}^n$, we obtain

$$f^\gamma \circ \lambda(A) \geq (f \circ \lambda)^\gamma(A),$$

and the proof is complete. ■

Remark 3.1. Let $f, g \in \Gamma_0(\mathbb{R}^n)$ be symmetric functions. It is clear that $pav(f, g)$ is a symmetric function. Also, if $F, G \in \Gamma_0(\mathcal{S}_n)$ are spectral functions, then, $pav(F, G)$ is a spectral function. Indeed, let $A \in \mathcal{S}_n$ and $U \in \mathcal{O}_n$ be given. Consider

$$\begin{aligned}
pav(F, G)(UAU^T) \quad &= \frac{1}{2} \inf_{\substack{(B,C)\in\mathcal{S}_n\times\mathcal{S}_n \\ B+C=2UAU^T}} \left\{F(B) + G(C) + \frac{1}{4}\|B - C\|_F^2\right\} \\
&= \frac{1}{2} \inf_{\substack{(B,C)\in\mathcal{S}_n\times\mathcal{S}_n \\ \underbrace{U^TBU}_{\widetilde{B}}+\underbrace{U^TCU}_{\widetilde{C}}=2A}} \left\{F(B) + G(C) + \frac{1}{4}\|B - C\|_F^2\right\} \\
&= \frac{1}{2} \inf_{\substack{(\widetilde{B},\widetilde{C})\in\mathcal{S}_n\times\mathcal{S}_n \\ \widetilde{B}+\widetilde{C}=2A}} \left\{F(\widetilde{B}) + G(\widetilde{C}) + \frac{1}{4}\|\widetilde{B} - \widetilde{C}\|_F^2\right\} \\
&= pav(F, G)(A)
\end{aligned}$$

Hence, $pav(F, G)$ is a spectral function.

Theorem 3.3. Let $f, g \in \Gamma_0(\mathbb{R}^n)$ be symmetric functions and let $\lambda : \mathcal{S}_n \longrightarrow \mathbb{R}^n$ be the eigenvalue function. Then,

$$pav(f, g) \circ \lambda(A) = pav(f \circ \lambda, g \circ \lambda)(A), \; \forall\, A \in \mathcal{S}_n.$$

Proof: Let $A \in \mathcal{S}_n$ be arbitrary. We first observe from (2.2) that

$$pav(f,g) \circ \lambda(A) = \frac{1}{2} \inf_{\substack{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n \\ x+y=2\lambda(A)}} \left\{ f(x) + g(y) + \frac{1}{4}\|x-y\|^2 \right\},$$

and

$$pav(f \circ \lambda, g \circ \lambda)(A) = \frac{1}{2} \inf_{\substack{(X,Y)\in\mathcal{S}_n\times\mathcal{S}_n \\ X+Y=2A}} \left\{ f(\lambda(X)) + g(\lambda(Y)) + \frac{1}{4}\|X-Y\|_F^2 \right\}$$

Let $x, y \in \mathbb{R}^n$ be such that $x + y = 2\lambda(A)$, and let $U \in \mathcal{O}_A$. Consider

$$\begin{aligned} f(x) + g(y) + \frac{1}{4}\left\|x-y\right\|^2 &= f\big(\lambda(Diag(x))\big) + g\big(\lambda(Diag(y))\big) + \frac{1}{2}\left\|Diag(x) - Diag(y)\right\|^2 \\ &\geq pav(f \circ \lambda, g \circ \lambda)\big(Diag(\lambda(A)\big) \\ &= pav(f \circ \lambda, g \circ \lambda)(U^T AU) = pav(f \circ \lambda, g \circ \lambda)(A). \end{aligned}$$

(Note that the above inequality follows from this fact that $Diag(x) + Diag(y) = 2Diag(\lambda(A))$). Now, taking infimum over all $x, y \in \mathbb{R}^n$ with $x + y = 2\lambda(A)$. We get

$$pav(f,g) \circ \lambda(A) \geq pav(f \circ \lambda, g \circ \lambda)(A). \tag{3.3}$$

Now, we conclude from (3.3) and Lemma 2.1(*ii*) that

$$(pav(f,g) \circ \lambda)^*(A) \leq \big(pav(f \circ \lambda, g \circ \lambda)\big)^*(A).$$

Hence, in view of Proposition 2.2(*iv*) and Theorem 2.2, one has

$$pav(f^*, g^*) \circ \lambda(A) \leq pav(f^* \circ \lambda, g^* \circ \lambda)(A). \tag{3.4}$$

Since (3.4) holds for all symmetric functions $f, g \in \Gamma_0(\mathbb{R}^n)$, we can replace f by f^* and g by g^* in (3.4). Therefore, it follows from Lemma 2.1(*iii*) that

$$pav(f,g) \circ \lambda(A) \leq pav(f \circ \lambda, g \circ \lambda)(A). \tag{3.5}$$

Hence, in view of (3.3) and (3.5) the proof is complete. ■

References

[1] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operators Theory in Hilbert Spaces*, Springer, New York, 2011.

[2] H.H. Bauschke, R. Goebel, Y. Lucet and X. Wang, The proximal average: Basic theory, *SIAM J. Optim.*, **19**(2) (2008), 766–785.

[3] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.

[4] J.M. Borwein and A.S. Lewis, *Convex Analysis and Nonlinear Optimization*, Springer, New York, 2006.

[5] J.M. Borwein and Q. Zhu, *Techniques of Variational Analysis*, CMS/Springer, New York, 2005.

[6] A. Daniilidis, A.S. Lewis, J. Malick and H. Sendov, Prox-regularity of spectral functions and spectral sets, *J. Convex Anal.*, **15**(3) (2008), 547–560.

[7] A.S. Lewis, Convex analysis on the Hermitian matrices, *SIAM J. Optim.*, **6**(1) (1996), 164–177.

[8] A.S. Lewis, Derivatives of spectral functions, *Math. Oper. Res.*, **21**(3) (1996), 576–588.

[9] A.S. Lewis, Nonsmooth analysis of eigenvalues, *Math. Program.*, **84**(1-24) (1999), 1–24.

[10] A.S. Lewis, Twice differentiable of spectral functions, *SIAM J. Matrix Anal. Appl.*, 23(2) (2001), 368–386.

[11] H. Mohebi, Analysis of symmetric matrix valued functions I, *Nonlinear Anal., Theory Methods Appl.*, **69**(1) (2008), 110–125.

[12] H. Mohebi and A. Salemi, Analysis of symmetric matrix valued functions, *Numer. Funct. Anal. Optim.*, **28**(5-6) (2007), 691–715.

[13] A.M. Rubinov, *Abstract Convexity and Global Optimization,* Kluwer Academic Publishers, Dordrecht-Boston-London, 2000.