UNIVERSITY OF BIRMINGHAM
University of Birmingham
Research at Birmingham

The Legendre–Fenchel Conjugate of the Product of Two Positive Definite Quadratic Forms

Zhao, Yun-Bin

DOI:
10.1137/090753711

Document Version
Peer reviewed version

Citation for published version (Harvard):
Zhao, Y-B 2010, 'The Legendre–Fenchel Conjugate of the Product of Two Positive Definite Quadratic Forms', *S I A M Journal on Matrix Analysis and Applications*, vol. 31, no. 4, pp. 1792-. https://doi.org/10.1137/090753711

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 19. Apr. 2024

(Published in *SIAM J. Matrix Analysis & Applications*, 31(2010), no.4, pp.1792-1811.)

THE LEGENDRE-FENCHEL CONJUGATE OF THE PRODUCT OF TWO POSITIVE-DEFINITE QUADRATIC FORMS

YUN-BIN ZHAO *

Abstract. It is well-known that the Legendre-Fenchel conjugate of a positive-definite quadratic form can be explicitly expressed as another positive-definite quadratic form, and that the conjugate of the sum of several positive-definite quadratic forms can be expressed via inf-convolution. However, the Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms is not clear at present. Jean-Baptiste Hiriart-Urruty posted it as an open question in the field of nonlinear analysis and optimization [‘Question 11’ in *SIAM Review* 49 (2007), 255-273]. From convex analysis point of view, it is interesting and important to address such a question. The purpose of this paper is to answer this question and to provide a formula for the conjugate of the product of two positive-definite quadratic forms. We prove that the computation of the conjugate can be implemented via finding a root to certain univariate polynomial equation. Furthermore, we show that the conjugate can be explicitly expressed as a single function in some situations. Our analysis shows that the relationship between the matrices of quadratic forms plays a vital role in determining whether or not the conjugate can be expressed explicitly, and our analysis also sheds some light on the computational complexity of the Legendre-Fenchel conjugate for the product of quadratic forms.

Key words. Convex analysis, matrix theory, quadratic form, Legendre-Fenchel conjugate

AMS subject classifications. 15A48, 65F15, 65K05, 90C25.

1. Introduction.

Given a function $h : R^n \to R$, its Legendre-Fenchel conjugate is defined as

$$h^*(x) = \sup_{y\in R^n} x^T y - h(y),$$

which is also widely referred to as the Legendre-Fenchel transformation of $h(y)$ in the literature (e.g. [1, 2, 3, 5, 10, 12, 13, 14, 25, 26, 27]). Like the familiar Fourier and Laplace transforms, the Legendre-Fenchel transformation takes in a function h and creates another function denoted by h^*. Lasserre [19] pointed out that the Legendre-Fenchel transformation is the analogue in the “min-plus” algebra of the Laplace transformation in the “plus-prod” algebra (see e.g. [4]).

Throughout this paper, we use the term ‘conjugate’ as a short for ‘Legendre-Fenchel conjugate (transformation)’. The conjugate has a significant impact in many areas. It plays an essential role in developing the convex optimization theory and methods (e.g. [3, 5, 6, 14, 25]). It is widely used in matrix analysis and eigenvalue optimization [20, 21, 22]. For example, in [20, 21], the conjugate was employed to establish a Fenchel-dual type theorem of a spectral function that can be viewed as an analogous result to von Neumann’s Theorem on unitarily invariant matrix norms. The conjugate is also commonly used in thermodynamics and in the theory of nonlinear differential equations of first order, e.g. in solving a class of Hamilton-Jacobi equation with explicit formulas [10, 16, 18, 23]. In addition, the conjugate of so-called ‘log-exp’ function is the well-known Shannon’s entropy function [28] which has been widely used in the field of information science, and in so many fields ranging from image enhancement to economics and from statistical mechanics to nuclear physics [8]. Recently, conjugate functions have been used to establish the smooth-convex-dual

*School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom (zhaoyy@maths.bham.ac.uk).

problems of the L_1 spline models which in their original forms are non-differential (see [9, 29] for example). Also, a broad class of robust optimization problems with general uncertainty sets can be represented as finite mathematical programming problems via conjugate functions (see [30]).

Although various properties of the conjugate function are presented (e.g. [3, 6, 25, 26]), given a function the expression of its conjugate is not always straightforward. In fact, the question of whether or not the conjugate of a given function can be (explicitly) expressed remains open in many situations. This stimulates many recent investigations on the expression or computation of the conjugate (see for example, [7, 9, 12, 13, 15, 19, 24, 29, 30]).

Let us consider the functions generated by quadratic forms. Let $q_A(y)$ and $q_B(y)$ be the quadratic forms:

$$q_A(y) = (1/2)y^T Ay, \quad q_B(y) = (1/2)y^T By \tag{1.1}$$

where A, B are two $n \times n$ positive definite matrices (in this case $q_A(y)$ and $q_B(y)$ are called positive-definite quadratic forms). For a single quadratic form, it is well-known (see e.g. [25, 26]) that its conjugate is also a quadratic form:

$$q_A^*(x) = (1/2)x^T A^{-1}x, \quad q_B^*(x) = (1/2)x^T B^{-1}x. \tag{1.2}$$

It is also well-known that the conjugate of the sum of quadratic forms (or more generally, the sum of convex functions) can be expressed via inf-convolution of conjugates of its individual components (e.g. [25, 26])). However, for the product of two positive-definite quadratic forms

$$f(y) = q_A(y)q_B(y) = (1/4)(y^T Ay)(y^T By), \tag{1.3}$$

the formula for its conjugate is not clear at present. So a natural question is: *What is the expression or formula of the conjugate of the product function $f(y)$?* From fast computation and practical application point of view, it is interesting and important to answer the above question. Jean-Baptiste Hiriart-Urruty posted it as an open question in the field of nonlinear analysis and optimization ([11], 'Problem 11').

In this paper, we will address this question and provide a formula for the conjugate of the function (1.3). We will show that if $f(y)$ is convex, its conjugate is finite at any point in R^n and the formula for the conjugate can be obtained. It should be stressed that the convexity of $f(y)$ is the only assumption required here. Since the product function loses convexity in general, we first establish a sufficient condition for the convexity of the product function in the next section. However, the main purpose of this paper is to derive an expression or formula for the conjugate. To achieve this goal, we will develop a series of technical results on the existence and representation of the solution to some nonlinear system of equations. These results will eventually lead to a formula of the conjugate (see Theorem 3.6 for details), based on which some other equivalent expressions will be obtained as well. In addition, we prove that a completely explicit expression of the conjugate can be obtained in some situations.

It should be mentioned that in one-dimensional case, Hiriart-Urruty et al [13] proved that the conjugate of the product of two both increasing or both decreasing positive convex functions can be expressed. Their formulas show that even for one-dimension functions, the conjugate of the product cannot be expressed in a simple way (see 'Theorem 15' in [13]). Their results dependent on monotonicity seen difficult to be generalized to the product of two general quadratic forms due to the loss of such monotonicity in higher dimensional spaces.

This paper is organized as follows. In Section 2, a sufficient convexity condition for the product of two positive-definite quadratic forms is developed. In Section 3, we establish a series of technical results and derive some formulas for the conjugate of the function (1.3). These formulas show that evaluating the conjugate amounts to finding a root to a univariate polynomial equation to which the root is shown, in Section 4, to be unique under some conditions. The completely explicit expression of the conjugate is discussed in Section 5. Conclusions are given in the last section.

Notation: Throughout this paper, $M \succ 0 (\succeq 0)$ denotes a positive definite (positive semi-definite) matrix. $\kappa(M)$ denotes the condition number of M, i.e., the ratio of its largest and smallest eigenvalues: $\lambda_{\max}(M)/\lambda_{\min}(M)$. Let $q_A(y)$, $q_B(y)$ and $f(y)$ be defined as (1.1) and (1.3) respectively, and their conjugates are denoted by $q_A^*(x), q_B^*(x)$ and $f^*(x)$ respectively. $q_A^*(x)$ and $q_B^*(x)$ are given as (1.2).

2. Convexity conditions for the product $f(y)$. To ensure that the conjugate is well-defined, a natural assumption is that f is convex. However, the product of two convex functions is not convex in general. Therefore, before we start to express the conjugate of $f(y)$, let us first develop some convexity conditions for $f(y)$. Notice that the gradient and Hessian matrix of f are given respectively as

$$\nabla f(y) = q_A(y)By + q_B(y)Ay, \tag{2.1}$$

$$\nabla^2 f(y) = q_A(y)B + q_B(y)A + Ayy^T B + Byy^T A. \tag{2.2}$$

When A and B are positive definite, a simple observation can be made immediately from (2.2): *If there exists a positive number $\gamma > 0$ such that $A = \gamma B$ (in which case A and B are called positively linearly dependant), the matrix $Ayy^T B + Byy^T A$ is positive semi-definite at any point $y \in R^n$. Hence, the function $f(y)$ is convex.* But the condition "$A = \gamma B$" is too restrictive, which basically implies that $f(y)$ is the square of a quadratic form. In what follows, we develop a general sufficient condition for the convexity of $f(y)$.

Let $\Im$ denote the following class of positive definite matrices:

$$\Im = \left\{ M \succ 0 : \ \ \kappa(M) = \lambda_{\max}(M)/\lambda_{\min}(M) \le (\sqrt{5}+1)/(\sqrt{5}-1) \right\}. \tag{2.3}$$

The condition $\kappa(M) \le (\sqrt{5}+1)/(\sqrt{5}-1)$ can be equivalently written as

$$(\lambda_{\max}(M) - \lambda_{\min}(M)) / (\lambda_{\max}(M) + \lambda_{\min}(M)) \le \sqrt{5}/5. \tag{2.4}$$

Notice that if $M \in \Im$ then $\gamma M \in \Im$ for any $\gamma > 0$. So $\Im$ is cone. This implies that if $M \in \Im$, so are all matrices which are positively linearly dependant with M. Moreover, by Weyl Theorem (Theorem 4.3.1 in [17]), it is easy to show that $\Im$ is also convex.

Clearly, $2.5 < (\sqrt{5}+1)/(\sqrt{5}-1)$. Thus, $\{M \succ 0 : \ \kappa(M) \le 2.5\} \subset \Im$. This fact will be used at the end of Section 4. The following lemma is used to prove Theorem 2.2 which is the main result of this section.

Lemma 2.1 ([17], Theorem 7.4.34). *Let M be a given $n \times n$ positive definite matrix with eigenvalues $0 < \lambda_{\min}(M) \le \cdots \le \lambda_{\max}(M)$. Then*

$$\left(x^T My\right)^2 \le \left(\frac{\lambda_{\max}(M) - \lambda_{\min}(M)}{\lambda_{\max}(M) + \lambda_{\min}(M)}\right)^2 (x^T Mx)(y^T My)$$

for every pair of orthogonal vectors $x, y \in R^n$.

The following result claims that the product of two quadratic forms is convex if the matrices A and B fall into the category $\Im$ defined by (2.3).

Theorem 2.2. *For any $n \times n$ matrices $A, B \in \Im$, the following inequality holds*

$$4(x^TAy)(x^TBy) + x^TAxy^TBy + y^TAyx^TBx \geq 0, \quad \textit{for all } x, y \in R^n, \tag{2.5}$$

and hence the function $f(y)$, defined by (1.3), is convex.

Proof. Denote

$$\vartheta(x,y) = 4(x^TAy)(x^TBy) + x^TAxy^TBy + y^TAyx^TBx.$$

To prove (2.5), it is sufficient to show that for every given $x \in R^n$, the inequality $\vartheta(x,y) \geq 0$ holds for any $y \in R^n$. There is nothing to prove when $x = 0$. Thus, in the remainder of this proof, let $0 \neq x \in R^n$ be an arbitrarily given point, and denote by $\mathcal{L}_x$ the (one-dimensional) subspace generated by $\{x\}$ and by $\mathcal{L}_x^\perp$ the orthogonal subspace of $\mathcal{L}_x$, i.e.

$$\mathcal{L}_x = \{tx : \ t \in R\}, \ \ \mathcal{L}_x^\perp = \{y : \ x^Ty = 0, \ y \in R^n\}.$$

For any vector $y \in R^n$, either $y \in \mathcal{L}_x^\perp$ or $y \notin \mathcal{L}_x^\perp$. We now prove that each of these cases implies that $\vartheta(x,y) \geq 0$ when $A, B \in \Im$.

Case 1: $y \in \mathcal{L}_x^\perp$. Denote by

$$\chi(A) = \frac{\lambda_{\max}(A) - \lambda_{\min}(A)}{\lambda_{\max}(A) + \lambda_{\min}(A)}, \ \ \chi(B) = \frac{\lambda_{\max}(B) - \lambda_{\min}(B)}{\lambda_{\max}(B) + \lambda_{\min}(B)}.$$

In this case, since y is orthogonal with x, by Lemma 2.1 we have

$$(\chi(A))^2 \, x^TAxy^TAy \geq (x^TAy)^2, \quad (\chi(B))^2 \, x^TBxy^TBy \geq (x^TBy)^2.$$

Therefore, we have

$$\begin{aligned}
&\chi(A)\chi(B)\left(x^TAxy^TBy + y^TAyx^TBx\right) \\
&\geq \chi(A)\chi(B)\left(2\sqrt{(x^TAxy^TBy)(y^TAyx^TBx)}\right) \\
&= 2\sqrt{(\chi(A))^2\,(x^TAxy^TAy)\,(\chi(B))^2\,(x^TBxy^TBy)} \\
&\geq 2\sqrt{(x^TAy)^2(x^TBy)^2} = 2|x^TAyx^TBy)|.
\end{aligned} \tag{2.6}$$

Since $A, B \in \Im$, by (2.4) we have $\chi(A) \leq \sqrt{5}/5$ and $\chi(B) \leq \sqrt{5}/5$. It follows from (2.6) that

$$(1/5)(x^TAxy^TBy + y^TAyx^TBx) \geq 2|x^TAyx^TBy|,$$

and hence

$$\begin{aligned}
&\vartheta(x,y) \\
&= \frac{3}{5}(x^TAxy^TBy + y^TAyx^TBx) + \left[\frac{2}{3}(x^TAxy^TBy + y^TAyx^TBx) + 4x^TAyx^TBy\right] \\
&\geq \frac{3}{5}(x^TAxy^TBy + y^TAyx^TBx) + 4(|x^TAyx^TAy| + (x^TAy)(x^TAy)) \geq 0.
\end{aligned}$$

Case 2: $y \notin \mathcal{L}_x^\perp$. In this case, y can be represented as $y = \tilde{x} + \tilde{y}$ for some $\tilde{x} \in \mathcal{L}_x$ and $\tilde{y} \in \mathcal{L}_x^\perp$. By the definition $\mathcal{L}_x$, $\tilde{x} = tx$ for some $t \in R$. Clearly $t \neq 0$ (since otherwise $y = \tilde{y} \in \mathcal{L}_x^\perp$). Therefore, by the definition of $\vartheta(x,y)$, we have

$$\vartheta(x,y) = \vartheta(x, tx + \tilde{y}) = \vartheta(x, t\,(x + \tilde{y}/t)) = t^2\vartheta\,(x, x + \tilde{y}/t)\,.$$

Thus, to prove that $\vartheta(x,y)\geq 0$, it is sufficient to show that $\vartheta(x,x+\tilde{y}/t)\geq 0$. Notice that $\tilde{y}/t\in\mathcal{L}_x^{\perp}$. It is sufficient to prove that

$$\vartheta(x,x+z)\geq 0,\quad \text{for all } z\in\mathcal{L}_x^{\perp}. \tag{2.7}$$

In fact, from the proof of '*Case 1*', we have actually proved the following inequality:

$$(1/5)\left(x^TAxz^TBz+z^TAzx^TBx\right)\geq 2|x^TAzx^TBz|,\quad \text{for any } z\in\mathcal{L}_x^{\perp}.$$

Thus, for any $z\in\mathcal{L}_x^{\perp}$, it follows from the above inequality that

$$\begin{aligned}
&\vartheta(x,x+z)\\
&=4(x^TA(x+z))x^TB(x+z)+(x^TAx)(x+z)^TB(x+z)+(x+z)^TA(x+z)(x^TBx)\\
&=6(x^TAx)(x^TBx)+6(x^TAx)(x^TBz)+6(x^TAz)(x^TBx)\\
&\quad +4(x^TAz)(x^TBz)+(x^TAx)(z^TBz)+(z^TAz)(x^TBx)\\
&\geq 6(x^TAx)(x^TBx)+6(x^TAx)(x^TBz)+6(x^TAz)(x^TBx)\\
&\quad +\frac{3}{5}(x^TAxz^TBz+z^TAzx^TBx)+\left[\frac{2}{5}(x^TAxz^TBz+z^TAzx^TBx)+4x^TAzx^TAz\right]\\
&\geq 6(x^TAx)(x^TBx)+6(x^TAx)(x^TBz)+6(x^TAz)(x^TBx)\\
&\quad +(3/5)(x^TAxz^TBz+z^TAzx^TBx)\\
&=6(x^TAx)(x^TBx)\eta(x,z)
\end{aligned}$$

where $\eta(x,z)$ is denoted by

$$\eta(x,z)=1+\frac{x^TAz}{x^TAx}+\frac{x^TBz}{x^TBx}+\frac{1}{10}\left(\frac{z^TBz}{x^TBx}+\frac{z^TAz}{x^TAx}\right).$$

Thus, it suffices to prove that $\eta(x,z)$ is nonnegative under our assumption. Since $x^Tz=0$, by Lemma 2.1, we see that

$$\frac{|x^TAz|}{x^TAx}\leq\left(\frac{\lambda_{\max}(A)-\lambda_{\min}(A)}{\lambda_{\max}(A)+\lambda_{\min}(A)}\right)\sqrt{\frac{z^TAz}{x^TAx}}\leq\frac{\sqrt{5}}{5}\sqrt{\frac{z^TAz}{x^TAx}}.$$

Similarly, we have

$$\frac{|x^TBz|}{x^TBx}\leq\frac{\sqrt{5}}{5}\sqrt{\frac{z^TBz}{x^TBx}}.$$

Therefore,

$$\begin{aligned}
\eta(x,z)&\geq 1-\frac{\sqrt{5}}{5}\sqrt{\frac{z^TAz}{x^TAx}}-\frac{\sqrt{5}}{5}\sqrt{\frac{z^TBz}{x^TBx}}+\frac{1}{10}\left(\frac{z^TBz}{x^TBx}+\frac{z^TAz}{x^TAx}\right)\\
&=\left(\frac{1}{2}-\frac{\sqrt{5}}{5}\sqrt{\frac{z^TAz}{x^TAx}}+\frac{1}{10}\frac{z^TAz}{x^TAx}\right)+\left(\frac{1}{2}-\frac{\sqrt{5}}{5}\sqrt{\frac{z^TBz}{x^TBx}}+\frac{1}{10}\frac{z^TBz}{x^TBx}\right)\geq 0,
\end{aligned}$$

where the last inequality follows from the fact that the quadratic function $\frac{1}{2}-\frac{\sqrt{5}}{5}\beta+\frac{1}{10}\beta^2\geq 0$ for any $\beta\in R$. Therefore, the inequality (2.7) holds as desired.

It follows from (2.2) that for any $x,y\in R^n$, $x^T\nabla^2 f(y)x=\frac{1}{2}\vartheta(x,y)$. Under our condition, $\vartheta(x,y)$ is nonnegative for any $x,y\in R^n$. Thus the Hessian $\nabla^2 f(y)$ is positive semi-definite at any point $y\in R^n$, and hence $f(y)$ is convex. □

Remark 2.3: We have given an affirmative answer to the question: *Can we develop certain convexity condition for the product function $f(y)$?* We believe that the condition in Theorem 2.2 can be relaxed in certain directions, and thus the convex cone $\Im$ may be enlarged without damaging the conclusion of the theorem. Such an extension is worthwhile in the sense that the developed condition can ensure the algebraic inequality (2.5) which may find some applications in optimization and numerical analysis. However, we are not intending to make such an extension here, since our major goal is to derive the formula for the conjugate.

3. Expression of the conjugate. In this section, we derive the formula for the conjugate $f^*(x)$. First, let us develop some useful technical results. Most of them cope with such issues as existence and representation of the solution to certain nonlinear system of equations associated with the pair of matrices (A, B). Through these results the final expression of the conjugate will be obtained (see Theorem 3.6 in this section).

Lemma 3.1. *Let A, B be positive definite. For any given $0 \neq x \in R^n$, let $g : (0, \infty) \to R$ be a function defined as*

$$g(\alpha) = \alpha - \frac{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x}, \quad \alpha > 0. \tag{3.1}$$

Then the equation $g(\alpha) = 0$ always has a (solution) root α in $(0, \infty)$, and every root to $g(\alpha) = 0$ is located in the interval

$$\left[\lambda_{\min}\left(B^{-1/2}AB^{-1/2}\right), \lambda_{\max}\left(B^{-1/2}AB^{-1/2}\right)\right]. \tag{3.2}$$

Proof. Clearly, $g(\alpha)$ is continuous (in fact, continuously differentiable, see Section 4 for detail). Notice that

$$\lim_{\alpha \to 0^+} g(\alpha) = -\left(x^TA^{-1}x\right) / \left(x^TA^{-1}BA^{-1}x\right) < 0,$$

and

$$\begin{aligned}\lim_{\alpha\to\infty} \frac{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} &= \lim_{\alpha\to\infty} \frac{x^T(\frac{1}{\alpha}A+B)^{-1}A(\frac{1}{\alpha}A+B)^{-1}x}{x^T(\frac{1}{\alpha}A+B)^{-1}B(\frac{1}{\alpha}A+B)^{-1}x} \\ &= \left(x^TB^{-1}AB^{-1}x\right)/(x^TB^{-1}x)\end{aligned}$$

which implies that $g(\alpha) \to \infty$ as $\alpha \to \infty$. Thus, by continuity the equation $g(\alpha) = 0$ must admit a root in $(0, \infty)$.

Denote by $z = B^{1/2}(A + \alpha B)^{-1}x$. Then for any root α of $g(\alpha) = 0$, we have

$$\begin{aligned}\alpha &= \frac{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} \\ &= \frac{x^T(A+\alpha B)^{-1}B^{1/2}(B^{-1/2}AB^{-1/2})B^{1/2}(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B^{1/2}B^{1/2}(A+\alpha B)^{-1}x} \\ &= \frac{z^T(B^{-1/2}AB^{-1/2})z^T}{z^Tz}\end{aligned}$$

which indicates that any root α must be located in between the largest and smallest eigenvalues of the matrix $B^{-1/2}AB^{-1/2}$, and thus in the interval (3.2). □

The next result plays a vital role in expressing the conjugate of $f(y)$.

Lemma 3.2. *Let A, B be positive definite, and let $0 \neq x \in R^n$ be an arbitrarily given vector. Then there exist two vectors $0 \neq x^{(1)}, 0 \neq x^{(2)}$ in R^n such that $(x^{(1)}, x^{(2)})$ is a solution to the following nonlinear system of equations:*

$$x = x^{(1)} + x^{(2)}, \tag{3.3}$$

$$\left[\frac{2}{(x^{(1)})^T B^{-1} A B^{-1} x^{(1)}}\right]^{1/3} B^{-1} x^{(1)} = \left[\frac{2}{(x^{(2)})^T A^{-1} B A^{-1} x^{(2)}}\right]^{1/3} A^{-1} x^{(2)}. \tag{3.4}$$

Moveover, any solution $(x^{(1)} \neq 0, x^{(2)} \neq 0)$ of the above system can be represented as

$$x^{(1)} = \left(I + \frac{1}{\alpha} AB^{-1}\right)^{-1} x, \quad x^{(2)} = \frac{1}{\alpha} AB^{-1} \left(I + \frac{1}{\alpha} AB^{-1}\right)^{-1} x \tag{3.5}$$

where $\alpha > 0$ is a positive scalar such that $g(\alpha) = 0$, i.e.,

$$\alpha = \frac{x^T (A + \alpha B)^{-1} A (A + \alpha B)^{-1} x}{x^T (A + \alpha B)^{-1} B (A + \alpha B)^{-1} x}. \tag{3.6}$$

Proof. Let $x \neq 0$ be given. First, by Lemma 3.1 there exists some scalar $\alpha > 0$ such that it satisfies (3.6). Let $\alpha^* > 0$ be such a scalar. Define

$$\bar{x}^{(1)} = \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-1} x, \quad \bar{x}^{(2)} = \frac{1}{\alpha^*} AB^{-1} \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-1} x. \tag{3.7}$$

Clearly, $\bar{x}^{(1)} \neq 0, \bar{x}^{(2)} \neq 0$ since $x \neq 0$. We now verify that $(\bar{x}^{(1)}, \bar{x}^{(2)})$ is a solution to the system of (3.3) and (3.4). Notice that the matrix $\begin{bmatrix} I & I \\ AB^{-1} & -\alpha^* I \end{bmatrix}$ is nonsingular, and its inverse is given by

$$\begin{bmatrix} I & I \\ AB^{-1} & -\alpha^* I \end{bmatrix}^{-1} = \begin{bmatrix} \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-1} & \frac{1}{\alpha^*} \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-1} \\ \frac{1}{\alpha^*} AB^{-1} \left[I + \frac{1}{\alpha^*} AB^{-1}\right]^{-1} & (\frac{1}{\alpha^*})^2 AB^{-1} \left[I + \frac{1}{\alpha^*} AB^{-1}\right]^{-1} - \frac{1}{\alpha^*} I \end{bmatrix}. \tag{3.8}$$

It is evident that $(\bar{x}^{(1)}, \bar{x}^{(2)})$ given by (3.7) can be rewritten as

$$\begin{pmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \end{pmatrix} = \begin{bmatrix} I & I \\ AB^{-1} & -\alpha^* I \end{bmatrix}^{-1} \begin{pmatrix} x \\ 0 \end{pmatrix},$$

i.e.,

$$\begin{bmatrix} I & I \\ AB^{-1} & -\alpha^* I \end{bmatrix} \begin{pmatrix} \bar{x}^{(1)} \\ \bar{x}^{(2)} \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix},$$

which simply is

$$\begin{cases} \bar{x}^{(1)} + \bar{x}^{(2)} = x, \\ B^{-1} \bar{x}^{(1)} = \alpha^* A^{-1} \bar{x}^{(2)}. \end{cases} \tag{3.9}$$

By (3.7), we have

$$\begin{aligned} (\bar{x}^{(1)})^T B^{-1} A B^{-1} \bar{x}^{(1)} &= x^T \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-T} (B^{-1} A B^{-1}) \left(I + \frac{1}{\alpha^*} AB^{-1}\right)^{-1} x \\ &= x^T \left(B + \frac{1}{\alpha^*} A\right)^{-1} A \left(B + \frac{1}{\alpha^*} A\right)^{-1} x \\ &= (\alpha^*)^2 x^T (A + \alpha^* B)^{-1} A (A + \alpha^* B)^{-1} x, \end{aligned} \tag{3.10}$$

and

$$
\begin{aligned}
&(\bar{x}^{(2)})^T A^{-1} B A^{-1} \bar{x}^{(2)} \\
&= x^T \left[\frac{1}{\alpha^*} AB^{-1} \left[I + \frac{1}{\alpha^*} AB^{-1} \right]^{-1} \right]^T A^{-1} B A^{-1} \left[\frac{1}{\alpha^*} AB^{-1} \left[I + \frac{1}{\alpha^*} AB^{-1} \right]^{-1} \right] x \\
&= x^T (A + \alpha^* B)^{-1} B (A + \alpha^* B)^{-1} x.
\end{aligned}
\tag{3.11}
$$

Thus, by the definition of α^* we have

$$
\left[\frac{(\bar{x}^{(1)})^T B^{-1} A B^{-1} \bar{x}^{(1)}}{(\bar{x}^{(2)})^T A^{-1} B A^{-1} \bar{x}^{(2)}} \right]^{\frac{1}{3}} = \left[\frac{(\alpha^*)^2 x^T (A + \alpha^* B)^{-1} A (A + \alpha^* B)^{-1} x}{x^T (A + \alpha^* B)^{-1} B (A + \alpha^* B)^{-1} x} \right]^{\frac{1}{3}} = [(\alpha^*)^3]^{\frac{1}{3}} = \alpha^*.
$$

This together with (3.9) implies that $(\bar{x}^{(1)}, \bar{x}^{(2)})$ is a solution to the nonlinear system (3.3) and (3.4). Thus, the system (3.3) and (3.4) does have a solution given as (3.7).

We now prove that any nonzero solution $(x^{(1)} \neq 0, x^{(2)} \neq 0)$ of the system (3.3) and (3.4) can be represented as the form of (3.5) with $\alpha > 0$ satisfying (3.6). The argument is basically the reverse of the above proof. Suppose that $(x^{(1)} \neq 0, x^{(2)} \neq 0)$ is a solution to the system (3.3) and (3.4). Set

$$
\widetilde{\alpha} = \left(\frac{(x^{(1)})^T B^{-1} A B^{-1} x^{(1)}}{(x^{(2)})^T A^{-1} B A^{-1} x^{(2)}} \right)^{1/3} > 0. \tag{3.12}
$$

Since (3.4) is equivalent to $B^{-1} x^{(1)} = \widetilde{\alpha} A^{-1} x^{(2)}$, i.e., $AB^{-1} x^{(1)} = \widetilde{\alpha} x^{(2)}$, the system (3.3) and (3.4) can be rewritten as

$$
\begin{bmatrix} I & I \\ AB^{-1} & -\widetilde{\alpha} I \end{bmatrix} \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}.
$$

The coefficient matrix of the left-hand side is non-singular, and its inverse is given by (3.8) provided that α^* in (3.8) is replaced by $\widetilde{\alpha}$. Therefore,

$$
\begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} I & I \\ AB^{-1} & -\widetilde{\alpha} I \end{bmatrix}^{-1} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} \left(I + \frac{1}{\widetilde{\alpha}} AB^{-1} \right)^{-1} x \\ \frac{1}{\widetilde{\alpha}} AB^{-1} \left(I + \frac{1}{\widetilde{\alpha}} AB^{-1} \right)^{-1} x \end{bmatrix} \tag{3.13}
$$

which is exactly the form of (3.5). We now prove that $\widetilde{\alpha}$ given by (3.12) satisfies (3.6). In fact, by (3.13), we have the following equalities which can be verified by the same way as (3.10) and (3.11):

$$
\begin{aligned}
(x^{(1)})^T B^{-1} A B^{-1} x^{(1)} &= \widetilde{\alpha}^2 x^T (A + \widetilde{\alpha} B)^{-1} A (A + \widetilde{\alpha} B)^{-1} x, \\
(x^{(2)})^T A^{-1} B A^{-1} x^{(2)} &= x^T (A + \widetilde{\alpha} B)^{-1} B (A + \widetilde{\alpha} B)^{-1} x.
\end{aligned}
$$

Combining (3.12) and the above two equalities, we have

$$
\widetilde{\alpha} = \left(\frac{(x^{(1)})^T B^{-1} A B^{-1} x^{(1)}}{(x^{(2)})^T A^{-1} B A^{-1} x^{(2)}} \right)^{1/3} = \left(\frac{\widetilde{\alpha}^2 x^T (A + \widetilde{\alpha} B)^{-1} A (A + \widetilde{\alpha} B)^{-1} x}{x^T (A + \widetilde{\alpha} B)^{-1} B (A + \widetilde{\alpha} B)^{-1} x} \right)^{1/3}
$$

which implies that $\widetilde{\alpha}$ satisfies (3.6). $\square$

The next result gives some equivalent descriptions for the solution to $g(\alpha)=0$.

Lemma 3.3. *Let A, B be positive definite. For any given $0 \neq x \in R^n$, each of the following univariate equations in α is equivalent to (3.6):*

$$2\alpha = \frac{x^T(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x}, \tag{3.14}$$

$$2 = \frac{x^T(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}. \tag{3.15}$$

Proof. Note that for any $\alpha > 0$ we have

$$\begin{aligned} & x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x \\ & = x^T(A+\alpha B)^{-1}(A+\alpha B-\alpha B)(A+\alpha B)^{-1}x \\ & = x^T(A+\alpha B)^{-1}x - \alpha x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x. \end{aligned} \tag{3.16}$$

Dividing both sides of the above by $x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x$, we get

$$\frac{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} = \frac{x^T(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} - \alpha$$

which implies that α satisfies (3.6) if and only if it satisfies (3.14). Thus, (3.6) and (3.14) are equivalent. Similarly, dividing both sides of (3.16) by $x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x$, we have

$$1 = \frac{x^T(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x} - \alpha\frac{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}A(A+\alpha B)^{-1}x}$$

which implies that α satisfies (3.6) if and only if it satisfies (3.15). □

Lemma 3.4. *(i) For any given $x^{(1)} \neq 0$ in R^n, the nonlinear equation*

$$q_A(y)By = x^{(1)} \tag{3.17}$$

has a unique solution which is given by

$$y^{(1)} = \left(\frac{2}{(x^{(1)})^T B^{-1}AB^{-1}x^{(1)}}\right)^{1/3} B^{-1}x^{(1)}. \tag{3.18}$$

(ii) Similarly, for any given $x^{(2)} \neq 0$ in R^n, the nonlinear equation

$$q_B(y)Ay = x^{(2)} \tag{3.19}$$

has a unique solution which is given by

$$y^{(2)} = \left(\frac{2}{(x^{(2)})^T A^{-1}BA^{-1}x^{(2)}}\right)^{1/3} A^{-1}x^{(2)}. \tag{3.20}$$

Proof. (i) It is easy to verify that (3.18) is a solution to the equation (3.17). We now prove that it is unique to (3.17). Let $\widetilde{y}$ be the unique solution to $By = x^{(1)}$, i.e., $\widetilde{y} = B^{-1}x^{(1)} \neq 0$ since $x^{(1)} \neq 0$. Assume that y is an arbitrary solution to (3.17). Thus, $B(q_A(y)y) = x^{(1)}$ and thus $q_A(y)y = \widetilde{y} \neq 0$, which indicates that $y \neq 0$. Denote by $\beta = 1/q_A(y)$. Then $y = \beta\widetilde{y}$. Since y is a solution to (3.17), substituting $y = \beta\widetilde{y}$ into (3.17) yields $x^{(1)} = q_A(\beta\widetilde{y})B(\beta\widetilde{y}) = \beta^3 q_A(\widetilde{y})B\widetilde{y}$. Since $B\widetilde{y} = x^{(1)} \neq 0$, it follows that $\beta^3 q_A(\widetilde{y}) = 1$, i.e., $\beta = 1/\left(q_A(\widetilde{y})\right)^{1/3}$. Therefore,

$$y = \frac{1}{\left(q_A(\widetilde{y})\right)^{1/3}}\widetilde{y} = \left(\frac{2}{\widetilde{y}^T A\widetilde{y}}\right)^{1/3}\widetilde{y} = \left(\frac{2}{(x^{(1)})^T B^{-1}AB^{-1}x^{(1)}}\right)^{1/3} B^{-1}x^{(1)} = y^{(1)}.$$

Therefore, the solution to (3.17) is unique and given by (3.18).

Similarly, we can prove (ii) by exchanging the role of A and B. □

The vectors $y^{(1)}, y^{(2)}$ given as (3.18) and (3.20) can be equal if $x^{(1)}$ and $x^{(2)}$ satisfy certain conditions. In fact, notice that (3.4) is nothing but $y^{(1)} = y^{(2)}$. Lemma 3.2 claims that there exist vectors $x^{(1)}, x^{(2)}$ such that $y^{(1)} = y^{(2)}$ in which case by Lemma 3.4 the system (3.17) and (3.19) have a common unique solution. Clearly, we have the following result:

Corollary 3.5. *Given $x^{(1)} \neq 0$ and $x^{(2)} \neq 0$ in R^n, the system*

$$\begin{cases} q_A(y)By = x^{(1)} \\ q_B(y)Ay = x^{(2)} \end{cases}$$

has a solution if and only if $y^{(1)} = y^{(2)}$, where $y^{(1)}$, $y^{(2)}$ are defined by (3.18) and (3.20) respectively. If this system has a solution, it must have a unique solution $y = y^{(1)} = y^{(2)}$.

We now prove the main result of this section.

Theorem 3.6. *Let A, B be positive definite and the function $f(y)$ be convex. Then at any point $x \in R^n$ the value of the conjugate $f^*(x)$ is finite, and $f^*(x) = 0$ if $x = 0$, otherwise if $x \neq 0$,*

$$f^*(x) = 3\alpha^{1/3}\left(\frac{x^T(A+\alpha B)^{-1}x}{4}\right)^{2/3} \tag{3.21}$$

where α is a root of the univariate equation $g(\alpha) = 0$ at x, i.e., α satisfies (3.6).

Proof. Clearly, $f^*(0) = 0$. We only need to prove the case $x \neq 0$. In fact, for any given $0 \neq x \in R^n$, by Lemma 3.2 there exist two vectors $x^{(1)} \neq 0, x^{(2)} \neq 0$ such that $(x^{(1)}, x^{(2)})$ is a solution to the system (3.3) and (3.4), and $x^{(1)}, x^{(2)}$ can be represented by (3.5) where α is a roof of $g(\alpha) = 0$. Notice that the equality (3.4) simply means $y^{(1)} = y^{(2)}$ (defined by (3.18) and (3.20)). Thus, by Lemma 3.4 or Corollary 3.5, there exists a unique vector denoted by y^* such that

$$q_A(y^*)By^* = x^{(1)}, \quad q_B(y^*)Ay^* = x^{(2)}. \tag{3.22}$$

The unique vector is equal to $y^{(1)}$ and $y^{(2)}$, i.e.,

$$y^* = \left(\frac{2}{(x^{(1)})^T B^{-1}AB^{-1}x^{(1)}}\right)^{1/3} B^{-1}x^{(1)} = \left(\frac{2}{(x^{(2)})^T A^{-1}BA^{-1}x^{(2)}}\right)^{1/3} A^{-1}x^{(2)}.$$

Since $(x^{(1)}, x^{(2)})$ is the solution to (3.3) and (3.4), substituting (3.22) into (3.3) yields $x = q_A(y^*)By^* + q_B(y^*)Ay^*$, which by (2.1) implies that,

$$x = \nabla f(y^*). \tag{3.23}$$

Since f is convex, the function $x^T y - f(y)$ is concave with respect to y, and its gradient with respect to y is given by $x - \nabla f(y)$. Thus, (3.23) implies that the gradient of this concave function at y^* is equal to zero, and hence its maximum value (over R^n) attains at y^*. Therefore, by the definition of the conjugate

$$f^*(x) = \sup_{y\in R^n} (x^T y - f(y)) = x^T y^* - f(y^*). \tag{3.24}$$

Since $f(y)$ is positively homogenuous of degree 4, i.e., $f(\lambda y) = \lambda^4 f(y)$ for any $\lambda > 0$. Thus, by Euler's formula, it follows form (3.23) that $x^T y^* = (\nabla f(y^*))^T y^* = 4f(y^*)$ (where the second equality above can be directly verified without using Euler's formula). Therefore, (3.24) implies that

$$\begin{aligned} f^*(x) &= x^T y^* - \frac{1}{4}x^T y^* = \frac{3}{4}x^T y^* \\ &= \frac{3}{4}\left(\frac{2}{(x^{(1)})^T B^{-1}AB^{-1}x^{(1)}}\right)^{1/3} x^T B^{-1}x^{(1)} \qquad (3.25) \\ &= \frac{3}{4}\left(\frac{2}{(x^{(2)})^T A^{-1}BA^{-1}x^{(2)}}\right)^{1/3} x^T A^{-1}x^{(2)}. \qquad (3.26) \end{aligned}$$

Since $x^{(1)}$ and $x^{(2)}$ are represented as (3.5), we can eliminate $x^{(1)}$ from (3.25) to get the formula (3.21). Similarly, eliminating $x^{(2)}$ from (3.26) yields the same result. Let us eliminate $x^{(1)}$ from (3.25). By simply substituting (3.5) into (3.25), we have

$$\begin{aligned} f^*(x) &= \frac{3}{4}\left[\frac{2}{x^T\left(I+\frac{1}{\alpha}AB^{-1}\right)^{-T} B^{-1}AB^{-1}\left(I+\frac{1}{\alpha}AB^{-1}\right)^{-1}x}\right]^{\frac{1}{3}} x^T B^{-1}\left[I+\frac{1}{\alpha}AB^{-1}\right]^{-1} x \\ &= \frac{3}{4}\left(\frac{2}{x^T\left(B+\frac{1}{\alpha}A\right)^{-1} A\left(B+\frac{1}{\alpha}A\right)^{-1}x}\right)^{\frac{1}{3}} x^T\left(B+\frac{1}{\alpha}A\right)^{-1} x \\ &= \frac{3}{4}\left(\frac{2}{\alpha^2(x^T(\alpha B+A)^{-1}A(\alpha B+A)^{-1}x}\right)^{\frac{1}{3}} \alpha x^T(\alpha B+A)^{-1}x \\ &= \frac{3}{4}\left(\frac{2\alpha}{x^T(\alpha B+A)^{-1}A(\alpha B+A)^{-1}x}\right)^{\frac{1}{3}} x^T(\alpha B+A)^{-1}x. \end{aligned}$$

Note that α is a solution to $g(\alpha) = 0$. By Lemma 3.3, α satisfies (3.15), and hence the above formula can be further simplified as

$$f^*(x) = \frac{3}{4}\left(\frac{4\alpha}{x^T(\alpha B+A)^{-1}x}\right)^{1/3} x^T(\alpha B+A)^{-1}x = 3\alpha^{1/3}\left(\frac{x^T(\alpha B+A)^{-1}x}{4}\right)^{2/3}.$$

If we eliminate $x^{(2)}$ from (3.26), then (3.14) will be used instead of (3.15). □

From Theorem 3.6, the conjugate of $f(y)$ is, roughly speaking, determined by a positive linear combination of their matrices. This is not surprising since both quadratic forms contribute equally and interrelatedly to the product function. However, in order to see what role the individual quadratic form plays in the conjugate of the product, we may rewrite the formula of $f^*(x)$ as follows.

Theorem 3.7. *Let A, B be positive definite and f be convex. Then $f^*(x) = 0$ if $x = 0$; otherwise for $x \neq 0$,*

$$f^*(x) = 3\alpha^{\frac{1}{3}}\left[\frac{2q_A^*(x) - \alpha x^T(AB^{-1}A+\alpha A)^{-1}x}{4}\right]^{\frac{2}{3}} = 3\alpha^{\frac{1}{3}}\left[\frac{2q_B^*(x) - x^T(\alpha BA^{-1}B+B)^{-1}x}{4\alpha}\right]^{\frac{2}{3}}$$

where $\alpha > 0$ is a root to $g(\alpha) = 0$ at x.

Proof. Notice that for any $\alpha > 0$ the inverse of $A + \alpha B$ can be given as the following two forms:

$$(A + \alpha B)^{-1} = A^{-1} - \alpha A^{-1}(B^{-1} + \alpha A^{-1})^{-1}A^{-1}, \tag{3.27}$$

$$(A + \alpha B)^{-1} = (1/\alpha)B^{-1} - (1/\alpha^2)B^{-1}\left(A^{-1} + (1/\alpha)B^{-1}\right)^{-1} B^{-1}. \tag{3.28}$$

Therefore, by (3.27), we have

$$\begin{aligned} x^T(A + \alpha B)^{-1}x &= x^T A^{-1}x - \alpha x^T A^{-1}(B^{-1} + \alpha A^{-1})^{-1}A^{-1}x \\ &= x^T A^{-1}x - \alpha x^T (AB^{-1}A + \alpha A)^{-1}x \\ &= 2q_A^*(x) - \alpha x^T (AB^{-1}A + \alpha A)^{-1}x. \end{aligned} \tag{3.29}$$

On the other hand, by (3.28) we have

$$\begin{aligned} x^T(A + \alpha B)^{-1}x &= (1/\alpha)x^T B^{-1}x - (1/\alpha^2)x^T(BA^{-1}B + (1/\alpha)B)^{-1}x \\ &= (1/\alpha)\left(2q_B^*(x) - x^T(\alpha BA^{-1}B + B)^{-1}x\right). \end{aligned} \tag{3.30}$$

Substituting (3.29) and (3.30) into (3.21) yields the desired formula. □

From the above result, f^* has two equal expressions in which either q_A^* or q_B^* is involved. An immediate consequence from this result is the following expression which is symmetric in the sense that both q_A^* and q_B^* are involved.

Corollary 3.8. *Under the same condition of Theorem 3.7, the conjugate $f^*(x) = 0$ if $x = 0$, otherwise for $x \neq 0$,*

$$f^*(x) = \frac{3}{2}\alpha^{\frac{1}{3}} \left\{ \left[\frac{2q_A^*(x) - \alpha x^T(AB^{-1}A + \alpha A)^{-1}x}{4}\right]^{\frac{2}{3}} + \left[\frac{2q_B^*(x) - x^T(B + \alpha BA^{-1}B)^{-1}x}{4\alpha}\right]^{\frac{2}{3}} \right\}$$

where α is a solution to $g(\alpha) = 0$ at x.

Remark 3.9. For any function $h(y)$ which is positive homogeneous of p-degree, its conjugate function h^* must be homogeneous of q-degree such that $(1/p)+(1/q) = 1$ (See Lasserre [19]). Since the product of two quadratic forms is positive homogeneous of 4-degree, i.e., $p = 4$, its conjugate f^* must be homogeneous of $\left(\frac{4}{3}\right)$-degree. This can be easily verified from (3.21) by noting that for any positive number $\lambda > 0$, the value of α does not change if x is replaced by λx.

Basically, at every $x \in R^n$ the conjugate f^* can be expressed in terms of the conjugate of the individual quadratic form, and the conjugate of certain combination of these two quadratic forms. The computation of $f^*(x)$ can be implemented via finding a root for the univariate equation $g(\alpha) = 0$ at x. It should be stressed that the aforementioned analysis does not rely on the uniqueness of the root of $g(\alpha) = 0$ at x. This implies that the conjugate f^* can be evaluated by any real root of $g(\alpha) = 0$. When the real root to $g(\alpha) = 0$ is not unique, all these roots yield the same value of $f^*(x)$. While the uniqueness of α is not required in the computation of the value of f^*, the question about whether the real root α is unique or not is interesting and worthy of further consideration. In the next section, we will prove that if the condition number of matrices involved are not too large, the root of $g(\alpha) = 0$ is always unique.

4. Uniqueness of the root to $g(\alpha)$ =0. In this section, we prove that the real root of $g(\alpha) = 0$ is unique under certain condition. We will see that all matrices in the convex cone $\Im$ defined in Section 2 satisfy this condition. First, let us calculate the derivative of $g(\alpha)$.

Lemma 4.1. *Let A, B be positive definite. For any given $0 \neq x \in R^n$, the function $g(\alpha)$, defined as (3.1), is continuously differentiable and its first derivative is given by $g'(\alpha) = 3 - \frac{w(\alpha)w''(\alpha)}{(w'(\alpha))^2}$ where $\alpha \in (0, \infty)$ and $w(\alpha) := x^T(A + \alpha B)^{-1}x$.*

Proof. For simplicity, denote

$$g_1(\alpha) = x^T(A + \alpha B)^{-1}A(A + \alpha B)^{-1}x, \ g_2(\alpha) = x^T(A + \alpha B)^{-1}B(A + \alpha B)^{-1}x.$$

For any given $\alpha > 0$ and letting t be sufficiently small, by (3.27) it is easy to see that

$$\begin{aligned} w(\alpha + t) &= x^T[A + (\alpha + t)B]^{-1}x \\ &= x^T[(A + \alpha B) + tB]^{-1}x^T \\ &= x^T\left\{(A + \alpha B)^{-1} - t(A + \alpha B)^{-1}\left[(B^{-1} + t(A + \alpha B)^{-1}\right]^{-1}(A + \alpha B)^{-1}\right\}x \\ &= w(\alpha) - tx^T\left\{(A + \alpha B)^{-1}\left[(B^{-1} + t(A + \alpha B)^{-1}\right]^{-1}(A + \alpha B)^{-1}\right\}x \end{aligned}$$

which implies that for any $\alpha > 0$,

$$w'(\alpha) = -x^T(A + \alpha B)^{-1}B(A + \alpha B)^{-1}x = -g_2(\alpha). \tag{4.1}$$

Let $Q = AB^{-1}A + 2\alpha A + \alpha^2 B$. Since for any $\alpha > 0$ we have

$$(A + \alpha B)^{-1}B(A + \alpha B)^{-1} = \left(AB^{-1}A + 2\alpha A + \alpha^2 B\right)^{-1} = Q^{-1}.$$

Thus, $g_2(\alpha) = x^TQ^{-1}x$. At any $\alpha > 0$, by (3.27) we have for any sufficiently small t,

$$\begin{aligned} g_2(\alpha + t) &= x^T(A + (\alpha + t)B)^{-1}B(A + (\alpha + t)B)^{-1}x \\ &= x^T[AB^{-1}A + 2(\alpha + t)A + (\alpha + t)^2B]^{-1}x \\ &= x^T[Q + t(2A + (2\alpha + t)B)]^{-1}x \\ &= x^T\left\{Q^{-1} - tQ^{-1}[(2A + (2\alpha + t)B)^{-1} + tQ^{-1}]^{-1}Q^{-1}\right\}x \\ &= g_2(\alpha) - tx^TQ^{-1}[(2A + (2\alpha + t)B)^{-1} + tQ^{-1}]^{-1}Q^{-1}x, \end{aligned}$$

which implies that

$$\begin{aligned} g_2'(\alpha) &= -x^TQ^{-1}(2A + 2\alpha B)Q^{-1}x \\ &= -2x^T(A + \alpha B)^{-1}B(A + \alpha B)^{-1}B(A + \alpha B)^{-1}x. \end{aligned}$$

This together with (4.1) implies that $w(\alpha)$ is twice continuously differentiable, and its second derivative is given as

$$w''(\alpha) = -g_2'(\alpha) = 2x^T(A + \alpha B)^{-1}B(A + \alpha B)^{-1}B(A + \alpha B)^{-1}x. \tag{4.2}$$

We now consider the derivative of $g_1(\alpha)$. By (3.16) and (4.1), we have

$$g_1(\alpha) = w(\alpha) - \alpha g_2(\alpha) = w(\alpha) + \alpha w'(\alpha) \tag{4.3}$$

which implies that

$$g_1'(\alpha) = 2w'(\alpha) + \alpha w''(\alpha). \tag{4.4}$$

Therefore, by (4.1), (4.2), (4.3) and (4.4), the derivative of $g(\alpha)$ is given by

$$g'(\alpha) = (\alpha - g_1(\alpha)/g_2(\alpha))' = 3 - w(\alpha)w''(\alpha)/(w'(\alpha))^2.$$

The proof is completed. □

Theorem 4.2. *Let A, B be positive definite. If the condition number*

$$\kappa(B^{-1/2}AB^{-1/2}) < 3 + 2\sqrt{3}, \tag{4.5}$$

then for any given $0 \neq x \in R^n$, there exists a unique $\alpha \in (0, \infty)$ such that $g(\alpha) = 0$.

Proof. By Lemma 3.1, all solutions of $g(\alpha) = 0$ on $(0, \infty)$ are in the interval (3.2). To prove the uniqueness of the solution, it suffices to show that the function $g(\alpha)$ is strictly increasing over the interval (3.2). Denote by $z = B^{\frac{1}{2}}(A + \alpha B)^{-1}x$. Let $w(\alpha)$ be defined as in Lemma 4.1. By (4.1) and (4.2), we have

$$\begin{aligned} \frac{w(\alpha)}{w'(\alpha)} &= -\frac{x^T(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} \\ &= -\frac{x^T(A+\alpha B)^{-1}B^{\frac{1}{2}}\left(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}\right)B^{\frac{1}{2}}(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} \\ &= -\frac{z^T\left(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}\right)z}{z^Tz} \end{aligned}$$

and

$$\begin{aligned} \frac{w''(\alpha))}{w'(\alpha)} &= -\frac{2x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x}{x^T(A+\alpha B)^{-1}B(A+\alpha B)^{-1}x} \\ &= -\frac{2z^TB^{\frac{1}{2}}(A+\alpha B)^{-1}B^{\frac{1}{2}}z}{z^Tz} = -\frac{2z^T\left(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}\right)^{-1}z}{z^Tz}. \end{aligned}$$

Denote by $P = B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$. By Lemma 4.1 and Kantorovich's inequality (See Theorem 7.4.41 in [17]), we have

$$\begin{aligned} g'(\alpha) &= 3 - w(\alpha)w''(\alpha)/(w'(\alpha))^2 \\ &= 3 - 2\left(\frac{z^T\left(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}\right)z}{z^Tz}\right)\left(\frac{z^T\left(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}\right)^{-1}z}{z^Tz}\right) \\ &\geq 3 - 2\left(\frac{\left(\lambda_{\max}(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}}) + \lambda_{\min}(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}})\right)^2}{4\lambda_{\max}(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}})\lambda_{\min}(B^{-\frac{1}{2}}(A+\alpha B)B^{-\frac{1}{2}})}\right) \\ &= 3 - \frac{[(\lambda_{\max}(P)+\alpha) + (\lambda_{\min}(P)+\alpha)]^2}{2(\lambda_{\max}(P)+\alpha)(\lambda_{\min}(P)+\alpha)} \\ &= \frac{2\alpha^2 + 2(\lambda_{\max}(P)+\lambda_{\min}(P))\alpha + 4\lambda_{\max}(P)\lambda_{\min}(P) - \lambda_{\max}(P)^2 - \lambda_{\min}(P)^2}{2(\lambda_{\max}(P)+\alpha)(\lambda_{\min}(P)+\alpha)}. \end{aligned}$$

To show that $g'(\alpha) > 0$ on the interval (3.2), it is sufficient to show the following quadratic function in α is positive over the interval (3.2):

$$\delta(\alpha) = 2\alpha^2 + 2(\lambda_{\max}(P)+\lambda_{\min}(P))\alpha + 4\lambda_{\max}(P)\lambda_{\min}(P)) - \lambda_{\max}(P)^2 - \lambda_{\min}(P)^2.$$

In fact, it is easy to verify that the quadratic equation $\delta(\alpha) = 0$ has at least one root, the largest one is given as

$$r = \frac{-(\lambda_{\max}(P)+\lambda_{\min}(P)) + \sqrt{3}(\lambda_{\max}(P)-\lambda_{\min}(P))}{2}.$$

It is easy to see that $r < \lambda_{\min}(P)$ if and only if $\kappa(P) = \lambda_{\max}(P)/\lambda_{\min}(P) < 3 + 2\sqrt{3}$. Thus, the condition of the theorem implies that $r < \lambda_{\min}(P)$. Hence $\delta(\alpha)$ is positive on the interval (3.2) since the interval is on the right of the largest root of $\delta(\alpha)$. $\square$

We now consider the set $\{M \succ 0 : \ \kappa(M) \leq 2.5\}$ which is a subset of $\Im$ (see Section 2). It is not difficult to show that for any $A, B \in \{M \succ 0 : \ \kappa(M) \leq 2.5\}$ the condition (4.5) is satisfied. In fact, let $0 \neq z \in R^n$ be an eigenvector of $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$ associated with $\lambda_{\max}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})$. Then

$$\lambda_{\max}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}) = \left(\frac{z^T B^{-\frac{1}{2}}AB^{-\frac{1}{2}}z}{z^T B^{-\frac{1}{2}}B^{-\frac{1}{2}}z}\right)\left(\frac{z^T B^{-1}z}{z^T z}\right) \leq \frac{\lambda_{\max}(A)}{\lambda_{\min}(B)}.$$

Let $u \in R^n$ be an eigenvector of $B^{-\frac{1}{2}}AB^{-\frac{1}{2}}$ associated with $\lambda_{\min}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}})$. Then

$$\lambda_{\min}(B^{-\frac{1}{2}}AB^{-\frac{1}{2}}) = \left(\frac{u^T B^{-\frac{1}{2}}AB^{-\frac{1}{2}}u}{u^T B^{-\frac{1}{2}}B^{-\frac{1}{2}}u}\right)\left(\frac{u^T B^{-1}u}{u^T u}\right) \geq \frac{\lambda_{\min}(A)}{\lambda_{\max}(B)}.$$

Therefore,

$$\kappa(B^{-1/2}AB^{-1/2}) = \frac{\lambda_{\max}\left(B^{-1/2}AB^{-1/2}\right)}{\lambda_{\min}\left(B^{-1/2}AB^{-1/2}\right)} \leq \left(\frac{\lambda_{\max}(A)}{\lambda_{\min}(B)}\right)\left(\frac{\lambda_{\max}(B)}{\lambda_{\min}(A)}\right) = \kappa(A)\kappa(B).$$

If $A, B \in \{M \succ 0 : \ \kappa(M) \leq 2.5\}$, the above inequality implies that $\kappa(B^{-1/2}AB^{-1/2}) \leq (2.5)^2 < 3 + 2\sqrt{3}$. By Theorem 4.2, the solution to $g(\alpha) = 0$ is unique. Moveover, since $\{M \succ 0 : \kappa(M) \leq 2.5\} \subset \Im$, by Theorem 2.2 the function $f(y)$ is also convex. Hence, combining Theorems 2.2, 3.6 and 4.2 leads to the following result:

Theorem 4.3. *Let $A \succ 0, B \succ 0, \kappa(A) \leq 2.5$ and $\kappa(B) \leq 2.5$. Then $f^*(0) = 0$, and for any $x \neq 0$*

$$f^*(x) = 3\alpha^{1/3}\left(x^T(A + \alpha B)^{-1}x/4\right)^{2/3},$$

where $\alpha > 0$ is the unique root to $g(\alpha) = 0$ at x.

From Theorems 3.6-3.8 and 4.3, we see that if the root of $g(\alpha) = 0$ can be given explicitly in terms of x, a completely explicit expression of f^* can be available by eliminating α from the formula of f^*. The next section is devoted to this discussion.

5. Completely explicit expression of the conjugate. In this section, we take a further step to prove that in some special cases a completely explicit representation of the conjugate can be available. Our analysis shows that whether a completely explicit expression of f^* can be obtained or not depends how close the connection between the two quadratic forms is. Let A be an $n \times n$ positive definite matrix and let $k = 0, 1, ..., n-1$. Consider the following cones of positive definite matrices:

$$C_A^{(k)} :=$$
$$\left\{\gamma A + \sum_{j=1}^{k} \tau_j u^j (u^j)^T : \ \gamma > 0, \tau_j > 0, u^j \in R^n, \|u^j\| = 1, j = 1, ..., k, \ (u^i)^T u^j = 0 \text{ for } i \neq j\right\}.$$

For example, when $k = 0, 1, 2$, we have $C_A^{(0)} = \{\gamma A : \ \ \gamma > 0\}$ and

$$C_A^{(1)} = \{\gamma A + \tau uu^T : \ \gamma > 0, \ \tau > 0, \ u \in R^n, \ \|u\| = 1\},$$
$$C_A^{(2)} = \{\gamma A + \tau_1 uu^T + \tau_2 vv^T : \ \gamma, \tau_1, \tau_2 > 0, u, v \in R^n, \|u\| = \|v\| = 1, u^T v = 0\}.$$

Thus, given a set of mutually orthogonal unit vectors $u^1, ..., u^k$ and positive numbers $\gamma > 0$ and $\tau_j > 0 (j = 1, ..., k)$, we may generate an element $M \in C_A^{(k)}$ by setting $M = \gamma A + \sum_{j=1}^k \tau_j u^j (u^j)^T$ which is a positive definite matrix. Clearly, each set $C_A^{(k)} (k = 0, 1, ..., n-1)$ is a cone. The next lemma shows that the union of these cones is exactly the whole positive-definite cone denoted by $\mathcal{S}_{++}^n = \{M : M \succ 0\}$.

Lemma 5.1. *Let A be an $n \times n$ positive definite matrix. Then any $n \times n$ positive definite matrix M must belong to some $C_A^{(k)}$, thus $\mathcal{S}_{++}^n = \bigcup_{k=0}^{n-1} C_A^{(k)}$.*

Proof. Clearly, any element of $C_A^{(k)}$ is positive definite, and thus the right-hand side of the above is contained in $\mathcal{S}_{++}^n$. It is sufficient to prove that any $n \times n$ positive definite matrix M must belong to a $C_A^{(k)}$. In fact, for any matrix $M \succ 0$, set $\gamma = \frac{1}{\lambda_{\max}(M^{-1/2} A M^{-1/2})}$. Then

$$M - \gamma A = M^{1/2} \left(I - \frac{1}{\lambda_{\max}(M^{-1/2} A M^{-1/2})} (M^{-1/2} A M^{-1/2}) \right) M^{1/2} \succeq 0$$

and at least one of the eigenvalues of $M - \gamma A$ is zero. If all its eigenvalues are zero, then $M = \gamma A$, i.e., $M \in C_A^{(0)}$. Otherwise, let $\tau_1 > 0, ..., \tau_k > 0$ be all its nonzero eigenvalues $(k \leq n-1)$, and let $u^1, ..., u^k$ be the eigenvectors corresponding to these eigenvalues respectively. These vectors can be chosen such that they are mutually orthogonal unit vectors. Therefore, $M - \gamma A$ can be represented as $M - \gamma A = \sum_{k=1}^k \tau_j u^j (u^j)^T$, and hence $M \in C_A^{(k)}$ where $k \leq n-1$. □

The next result claims that if $B \in C_A^k$ where $k = 0, 1$ then a completely explicit expression of the conjugate is available.

Theorem 5.2. *Let $A \succ 0$, $B \succ 0$ and f be convex. Then*

(i) *If $B \in C_A^{(0)}$, then $f^*(x) = 3 \left(\frac{\lambda_{\max}(B^{-\frac{1}{2}} A B^{-\frac{1}{2}})}{16} \right)^{1/3} \left(\frac{x^T A^{-1} x}{2} \right)^{2/3}$.*

(ii) *If $B \in C_A^{(1)}$, then $f^*(0) = 0$ and at every $x \neq 0$, $f^*(x)$ can be completely and explicitly expressed, i.e., $f^*(x) = 3(\alpha(x))^{1/3} \left(\frac{x^T (A + \alpha(x) B)^{-1} x}{4} \right)^{2/3}$, where $\alpha(x)$ is the real root (which can be explicitly given in terms of x) of the the cubic polynomial equation in α: $\alpha^3 + c_1(x)\alpha^2 + c_2(x)\alpha + c_3(x) = 0$, where $c_1(x), c_2(x), c_3(x)$ are explicitly given in terms of x and some known data determined by A and B.*

Proof. (i) Notice that $B \in C_A^{(0)}$ if and only if there exists a $\gamma > 0$ such that $B = \gamma A$, and hence $B^{-1/2} A B^{-1/2} = (1/\gamma) I$. Thus,

$$1/\gamma = \lambda_{\max}(B^{-1/2} A B^{-\frac{1}{2}}) = \lambda_{\min}(B^{-\frac{1}{2}} A B^{-1/2})$$

in which case the interval (3.2) in Lemma 3.1 is reduced to a single point, and thus the solution to $g(\alpha) = 0$ is $\alpha = \lambda_{\max}(B^{-1/2} A B^{-1/2}) = 1/\gamma$, which can be verified also by subtituting $B = \gamma A$ into (3.6). Thus, by Theorem 3.6, the conjugate function is given

$$f^*(x) = 3 \left(\frac{1}{\gamma} \right)^{\frac{1}{3}} \left(\frac{x^T (A + \frac{1}{\gamma}(\gamma A))^{-1} x}{4} \right)^{\frac{2}{3}} = 3 \left(\frac{\lambda_{\max}(B^{-\frac{1}{2}} A B^{-\frac{1}{2}})}{16} \right)^{\frac{1}{3}} \left(\frac{x^T A^{-1} x}{2} \right)^{\frac{2}{3}}.$$

(ii) Let $B \in C_A^1$. We now prove that the equation $g(\alpha) = 0$ can be written as a cubic polynomial equation. Thus its real solution can be expressed explicitly. Since

$B \in C_A^1$, the matrix B can be represented as $B = \gamma A + \tau uu^T$ where $\gamma > 0, \tau > 0$ and $\|u\| = 1$. This implies that

$$I - \gamma B^{-1/2} A B^{-1/2} = \tau (B^{-1/2} u)(B^{-1/2} u)^T$$

which is a rank-one and positive semi-definite matrix. Therefore, it is easy to see that $\gamma = 1/\lambda_{\max}\left(B^{-\frac{1}{2}} A B^{-\frac{1}{2}}\right)$. Since $B - \gamma A = \tau uu^T$ and $\|u\| = 1$, τ must be the unique nonzero eigenvalue of $B - \gamma A$. Therefore, we have

$$\tau = \lambda_{\max}(B - \gamma A) = \lambda_{\max}\left(B - \frac{1}{\lambda_{\max}\left(B^{-\frac{1}{2}} A B^{-\frac{1}{2}}\right)} A\right),$$

and the vector u with $\|u\| = 1$ is an eigenvector of $B - \gamma A$ associated with the eigenvalue τ. Thus, when $B \in C_A^{(1)}$, the data (γ, τ, u) can be completely given.

By Sherman-Morrison formula, we have

$$\begin{aligned} (A + \alpha B)^{-1} &= \left(A + \alpha(\gamma A + \tau uu^T)\right)^{-1} = \left((1 + \alpha\gamma)A + \alpha\tau uu^T\right)^{-1} \\ &= [(1 + \alpha\gamma)A]^{-1} - \frac{[(1 + \alpha\gamma)A]^{-1}(\alpha\tau uu^T)[(1 + \alpha\gamma)A]^{-1}}{1 + \alpha\tau u^T[(1 + \alpha\gamma)A]^{-1}u} \\ &= \frac{1}{1 + \alpha\gamma}\left(A^{-1} - \frac{\alpha\tau A^{-1}uu^T A^{-1}}{(1 + \alpha\gamma) + \alpha\tau u^T A^{-1}u}\right). \end{aligned} \tag{5.1}$$

Therefore, if α is the root to $g(\alpha) = 0$, it must satisfy that

$$\begin{aligned} & x^T(A + \alpha B)^{-1} A (A + \alpha B)^{-1} x \\ & = \alpha x^T (A + \alpha B)^{-1} B (A + \alpha B)^{-1} x \\ & = \alpha x^T (A + \alpha B)^{-1} (\gamma A + \tau uu^T)(A + \alpha B)^{-1} x \\ & = \alpha\gamma x^T (A + \alpha B)^{-1} A (A + \alpha B)^{-1} x + \alpha\tau x^T (A + \alpha B)^{-1} uu^T (A + \alpha B)^{-1} x. \end{aligned}$$

Thus,

$$(1 - \gamma\alpha) x^T (A + \alpha B)^{-1} A (A + \alpha B)^{-1} x = \alpha\tau (x^T (A + \alpha B)^{-1} u)^2.$$

By (3.15), the above equality implies that

$$(1 - \alpha\gamma) x^T (A + \alpha B)^{-1} x = 2\alpha\tau (x (A + \alpha B)^{-1} u)^2. \tag{5.2}$$

By (5.1) we have

$$x^T (A + \alpha B)^{-1} x = \frac{1}{1 + \alpha\gamma}\left(x^T A^{-1} x - \frac{\alpha\tau (x^T A^{-1} u)^2}{(1 + \alpha\gamma) + \alpha\tau u^T A^{-1} u}\right)$$

and

$$\begin{aligned} x^T (A + \alpha B)^{-1} u &= \frac{1}{1 + \alpha\gamma}\left(x^T A^{-1} u - \frac{\alpha\tau (x^T A^{-1} u)(u^T A^{-1} u)}{(1 + \alpha\gamma) + \alpha\tau u^T A^{-1} u}\right) \\ &= \frac{x^T A^{-1} u}{(1 + \alpha\gamma) + \alpha\tau u^T A^{-1} u}. \end{aligned}$$

Substituting the last two equalities into (5.2) yields

$$\frac{1-\alpha\gamma}{1+\alpha\gamma}\left(x^TA^{-1}x-\frac{\alpha\tau(x^TA^{-1}u)^2}{1+\alpha(\gamma+\tau u^TA^{-1}u)}\right)=\frac{2\alpha\tau(x^TA^{-1}u)^2}{(1+\alpha(\gamma+\tau u^TA^{-1}u))^2}.$$

Denote by $D(x)=(x^TA^{-1}u)^2/(x^TA^{-1}x)$ and $\beta=\gamma+\tau u^TA^{-1}u$. Then the equation above can be written as

$$\frac{1-\alpha\gamma}{1+\alpha\gamma}\left(1-\frac{\alpha\tau D(x)}{1+\alpha\beta}\right)=\frac{2\alpha\tau D(x)}{(1+\alpha\beta)^2}.$$

Multiplying both sides by $(1+\alpha\gamma)(1+\alpha\beta)^2$ and rearranging the terms, we have

$$[\gamma\beta(\tau D(x)-\beta)]\alpha^3+\left[\beta^2-2\beta\gamma-\tau(\beta+\gamma)D(x)\right]\alpha^2+[2\beta-\gamma-3\tau D(x)]\alpha+1=0.$$

Since $(x^TA^{-1}u)^2\le(x^TA^{-1}x)(u^TA^{-1}u)$, it is easy to see that the coefficient of α^3 is nonzero. In fact, $\gamma\beta(\tau D(x)-\beta)<0$. Let

$$c_1(x)=\frac{\beta^2-2\beta\gamma-\tau(\beta+\gamma)D(x)}{\gamma\beta(\tau D(x)-\beta)},\ c_2(x)=\frac{2\beta-\gamma-3\tau D(x)}{\gamma\beta(\tau D(x)-\beta)},\ c_3(x)=\frac{1}{\gamma\beta(\tau D(x)-\beta)},$$

Then, the equation (3.6) is eventually written as $\alpha^3+c_1(x)\alpha^2+c_2(x)\alpha+c_3(x)=0$ which is a cubic polynomial. Its real root can be expressed explicitly in terms of its coefficients c_1,c_2 and c_3. Since α can be explicitly given, the desired explicit formula of f^* can be obtained by substituting $\alpha(x)$ into (3.21). □

When $n\le 2$, it follows from Lemma 5.1 that $\mathcal{S}^n_{++}$ is equal to the union of at most two subcones: $C_A^{(0)}$ and $C_A^{(1)}$. In both case, by Theorem 5.2, f^* can be completely and explicitly expressed. Thus, the following result is an immediate consequence of Theorem 5.2.

Corollary 5.3. *(i) If $n=1$ in which case $f(t)=\left(\frac{1}{2}\gamma_1t^2\right)\left(\frac{1}{2}\gamma_2t^2\right)$ where $\gamma_1>0$ and $\gamma_2>0$ are two constants, then the conjugate is given by $f^*(t)=\frac{3}{4}\left(\frac{1}{\gamma_1\gamma_2}\right)^{1/3}t^{4/3}$.*

(ii) *Let A,B be any 2×2 positive definite matrices and $f(y)$ be convex. Then its conjugate f^* can always be completely and explicitly expressed.*

Proof. (i) When $n=1$, the matrices $A\succ 0, B\succ 0$ reduce to two scalars $\gamma_1>0$ and $\gamma_2>0$ respectively. Clearly, for this case, $\lambda_{\max}(B^{-1/2}AB^{-1/2})=\frac{\gamma_1}{\gamma_2}$. By Theorem 5.2 (i), we have

$$f^*(t)=3\left(\frac{\gamma_1}{16\gamma_2}\right)^{1/3}\left(\frac{t^2}{2\gamma_1}\right)^{2/3}=\frac{3}{4}\left(\frac{1}{\gamma_1\gamma_2}\right)^{1/3}t^{4/3}.$$

(ii) For any 2×2 matrices $A\succ 0$ and $B\succ 0$, we must either have $B\in C_A^{(0)}$ or $B\in C_A^{(1)}$. Both cases, by Theorem 5.2, imply that the conjugate f^* can be completely and explicitly expressed. □

We now consider the more general cases: $n>2$ and $B\in C_A^k$ where $2\le k\le n-1$. Notice that in this case, the matrix B can be represented $B=\gamma A+\tau_1u^1(u^1)^T+...+\tau_ku^k(u^k)^T$. By applying several times of Sherman-Morrison-formula, it is not difficult to see that the equation (3.6) can be represented explicitly. For example, when $k=2$, the matrix $B-\gamma A$ is a rank-two matrix. By applying two times of Sherman-morrison formula, $(A+\alpha B)^{-1}$ can be represented explicitly in terms of the inverse of A^{-1}, and the equation (3.6) can be written explicitly as a quintic polynomial. We have the following general observation:

Observation: *If* $B \in C_A^k$ $(k = 0, 1, ..., n-1)$, *then* $g(\alpha) = 0$ *can be equivalently written as a polynomial equation (in* α*) with degree* $2k + 1$.

When $k = 2$, $g(\alpha) = 0$ can be written as quintic polynomial equation whose solution cannot be expressed explicitly by basic arithmetic operations. Thus, when $k \geq 2$, some numerical methods should be employed to find a root to such a polynomial equation. Clearly, the classification of the positive definite matrices into n subclasses such as $C_A^k(k = 0, ..., n-1)$ can actually be regarded as a way for measuring the connection (or distance) between matrices: $B \in C_A^k$ means $B - \gamma A$ is a rank-k matrix. The larger the number k is, the weaker the relationship between A and B will be, and the less likely a completely explicit expression of the conjugate will be available.

Before we close this section, let us state one more result concerning the explicit expression of the conjugate f^*.

Theorem 5.4. *Let* $A \succ 0, B \succ 0$ *be two* $n \times n$ *matrices such that* f *is convex. If* A, B *can be simultaneously diagonalizable by congruence, i.e., there exists a nonsingular matrix* U *such that*

$$A = U^T diag(\sigma_1(A), \cdots, \sigma_n(A))U, \quad B = U^T diag(\sigma_1(B), \cdots, \sigma_n(B))U, \tag{5.3}$$

where $\sigma_i(A)$ *and* $\sigma_i(B)$ *satisfy the following properties:*

(a) $\sigma_1(A) \neq \sigma_2(A) = \cdots = \sigma_n(A)$,

(b) $\sigma_2(B) = \cdots = \sigma_n(B)$ *and* $\sigma_1(B)$ *can be any positive number.*

Then the conjugate f^* *can be completely and explicitly expressed.*

Proof. For any given $x \neq 0$, notice that the solution α to $g(\alpha) = 0$ can be obtained by solving the equation (3.15), i.e.,

$$x^T(A + \alpha B)^{-1}x = 2x^T(A + \alpha B)^{-1}A(A + \alpha B)^{-1}x.$$

Substituting (5.3) into the above, and using (a) and (b), we have

$$\frac{\rho_1^2}{\sigma_1(A) + \alpha\sigma_1(B)} + \frac{(n-1)\rho_2^2}{\sigma_2(A) + \alpha\sigma_2(B)} = 2\left(\frac{\sigma_1(A)\rho_1^2}{(\sigma_1(A) + \alpha\sigma_1(B))^2} + \frac{(n-1)\sigma_2(A)\rho_2^2}{(\sigma_2(A) + \alpha\sigma_2(B))^2}\right)$$

where $\rho_j = \left(U^{-T}x\right)_j$, the jth component of the vector $U^{-T}x$. By multiplying both sides of the above by $(\sigma_1(A) + \alpha\sigma_1(B))^2(\sigma_2(A) + \alpha\sigma_2(B))^2$, the above equation is reduced to a cubic polynomial equation in α. Its solution α can be explicitly expressed, and thus by (3.21) f^* can be completely and explicitly given. □

Conclusions. In this paper, we have derived formulas for the Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms, and thus the open 'question 11' in [*SIAM Review* 49 (2007), 255-273] has been addressed. We start with developing a sufficient convexity condition for the product function which to our knowledge is the first sufficient convexity condition for this class of functions. Following that, we develop a series of technical results on the existence and uniqueness of the solution to certain nonlinear system of equations. These technical results were employed to prove our main results (Theorems 3.6) concerning the expression of the conjugate of the product function. We have proved that (i) if the product is convex, at any point in R^n its conjugate function is finite, and can be given by the formula (3.21); (ii) the computation of the conjugate can be implemented via solving a univariate polynomial equation with odd-degree to which a root always exists (Lemma 3.1 and discussion in Section 5), and the root is unique if the condition numbers of the matrices involved are not too large (Theorem 4.2); and (iii) when the connection of two positive definite matrices are strong enough ($B \in C_A^{(k)}, k = 0, 1$), a completely explicit expression of the conjugate can be obtained (Theorem 5.2). Particularly, in one- and two-dimensional spaces, a completely explicit expression of the conjugate of the product function is always available (Corollary 5.3) provided the product is convex.

REFERENCES

[1] H. Attouch, R.J.-B. Wets, Isometries for the Legendre-Fenchel transform, *Trans. Amer. Math. Soc.*, 296 (1986), pp.33-60.

[2] H. Attouch, R.J.-B. Wets, Another isometry for the Legendre-Fenchel transform, *J. Math. Anal. Appl.,* 131 (1988), pp.404-411.

[3] J. P. Aubin, *Optima and Equilibria: An intoduction to Nonlinear Analysis*, Springer-Verlag, 1993.

[4] F. Bacelli, G. Cohen, G. Olsder and J. Quadrat, *Synchronization and lienarity*, Wiley, 1992.

[5] D.P. Bertsekas, A. Nedić and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, 2003.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[7] Y. Brenier, A fast algorithm for the computation of the discrete Legendre-Fenchel transform, *C. R. Acad. Sci. Paris Sér. I Math.*, 308 (1989), pp. 587–589.

[8] B. Buck and V.A. Macaulay, *Maximum Entropy in Action: A Collection of Expository Essays*, Oxford University Press, London, UK, 1991.

[9] H. Cheng, S.C. Fang and J.E. Lavery, A geometric programming framework for univariate cubic L_1 smoothing spline, *Ann. Oper. Res.*, 133 (2005), pp. 229-248.

[10] L. Corrias, Fast Legendre-Cenchel transform and applications to hamilton-Jacobi equations and conservation laws, *SIAM J. Numer. Anal.*, 33 (1996), pp. 1534-1558.

[11] J.B. Hiriart-Urruty, Potpourri of conjectures and open questions in nonlinear analysis and optimization, *SIAM Review*, 49 (2007), pp. 255-273.

[12] J.B. Hiriart-Urruty, *A note on the Legendre-Fenchel transform of convex composite functions*, in Nonsmooth Mechanics and Analysis (Book series of Advances in Matheamtics and Mechanics, Springer), Vol. 12, 2006, pp. 34-56.

[13] J.B. Hiriart-Urruty and J.E. Martinez-Legaz, New formulas for the Legendre-Fenchel transform, *J. Math. Anal. Appl.,* 288 (2003), pp. 544-555.

[14] J.B. Hiriart-Urruty and C. Lemarechal, *Fundamentals of Convex Analysis*, Grundlehren Text Ed., Springer-Verlag, Berlin, 2001.

[15] J.B. Hiriart-Urruty and Y. Lucet, Parametric computation of the Legendre-Fenchel conjugate with application to the computation of the Moreau envelope, *J. Convex Anal.*, 14 (2007), pp.657-666.

[16] E. Hope, Generalized solutions of nonlinear equations of the first order, *J. Math. Mech.*, 14 (1965), pp. 951-973.

[17] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1986.

[18] A. V. Igorevich, *Mathematical Methods of Classical Mechanics* (second edition), Springer, 1989.

[19] J.B. Lasserre, Homogeneous functions and conjugacy, *J. Convex Anal.*, 5 (1998), pp. 397-403.

[20] A.S. Lewis, The mathematics of eigenvalue optimization, *Math. Program.*, Ser.B, 97 (2003), pp. 155-176.

[21] A.S. Lewis, Convex analysis on the Hermitian matrices, *SIAM J. Optim.*, 6 (1996), pp. 164-177.

[22] A.S. Lewis and M.L. Overton, Eigenvalue optimization, *Acta Numerica*, 5 (1996), pp. 149-190.

[23] P.L. Lions, *Generalized Solutions of Hamilton-Jacobi Equations*, Pitman, London, 1982.

[24] Y. Lucet, A fast computational algorithm for the Legendre-Fenchel transform, *Comput. Optim. Appl.*, 6 (1996), pp. 27-57.

[25] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[26] R.T. Rockafellar, *Conjugate Duality and Optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics No.16, SIAM, Philadelphia, 1974.

[27] R.T. Rochafellar and R.J.B. Wets, *Variational Analysis*, Springer, 1998.

[28] C.E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, 27 (1948), pp. 379-423 and 623-656.

[29] Y.B. Zhao, S.C. Fang and J.E. Lavery, Geometric dual formulation for first-derivative-based univariate cubic L_1 splines, *J. Global Optim.*, 40(2008), pp.589-621.

[30] I. Averbakh and Y.B. Zhao, Explicit reformulations of robust optimization problens with general uncertainty sets, *SIAM J. Optim.*, 18 (2007), pp. 1436-1466.