Noname manuscript No.
(will be inserted by the editor)

A note on Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms

Yong Xia

Received: date / Accepted: date

Abstract The Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms was posted as an open question in the field of nonlinear analysis and optimization by Hiriart-Urruty ['Question 11' in *SIAM Review* 49, 255-273, (2007)]. Under a convex assumption on the function, it was answered by Zhao [SIAM J. Matrix Analysis & Applications, 31(4), 1792-1811 (2010)]. In this note, we answer the open question without making the convexity assumption.

Keywords Legendre-Fenchel conjugate · quadratic form

PACS 15A48 · 65F15 · 65K05 · 90C25

1 Introduction

The Legendre-Fenchel conjugate of the function $h(y)$ is defined as

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

For example, if $h(y) = \frac{1}{2}y^T Ay$, where A is positive definite, then $h^*(x) = \frac{1}{2}x^T A^{-1}x$. In the field of nonlinear analysis and optimization, J.B. Hiriart-Urruty [1] raised an open question *what is the expression or formula of the conjugate of the product function*

$$f(y) = \frac{1}{4}(y^T Ay)(y^T By),$$

This research was supported by National Natural Science Foundation of China under grants 11001006 and 91130019/A011702, and by the fund of State Key Laboratory of Software Development Environment under grant SKLSDE-2011ZX-15.

Y. Xia
State Key Laboratory of Software Development Environment, LMIB of the Ministry of Education, School of Mathematics and System Sciences, Beihang University, Beijing 100191, P. R. China E-mail: dearyxia@gmail.com

where A, B are two $n \times n$ positive definite matrices.

Recently, Zhao [3] answered the open question under the assumption that $f(y)$ is convex. The main result is as follows.

Theorem 1 *([3]) Let the function $f(y)$ be convex. 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 $x \neq 0$,*

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

where α is any real root of the univariate equation

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

The following gives a sufficient condition under which the real root to the equation (1) is unique.

Theorem 2 *([3]) If $\max\{\kappa(A), \kappa(B)\} \leq 2.5$, where $\kappa(\cdot)$ denotes the condition number, then $f(y)$ is convex and for any $x \neq 0$, there exists a unique real root to the equation (1).*

In this note, we answer the open question without the convexity assumption. Our proof is much shorter than that of Theorem 1 given in [3]. As a corollary, the sufficient condition (Theorem 2) is improved.

2 Main Result

In this section, we study the conjugate $f^*(x) = \sup_{y\in R^n} x^Ty - f(y)$ without assuming that $f(y)$ is convex.

Lemma 1 *The conjugate $f^*(x)$ is finite. $f^*(x) = 0$ if $x = 0$, otherwise if $x \neq 0$, $f^*(x) > 0$.*

Proof. The finiteness of $f^*(x)$ follows from

$$\lim_{\|y\|_2\to\infty} x^Ty - \frac{1}{4}(y^TAy)(y^TBy) \leq \lim_{\|y\|_2\to\infty} x^Ty - \frac{\lambda_{\min}(A)\lambda_{\min}(B)}{4}\|y\|_2^4 = -\infty,$$

where $\lambda_{\min}(A) > 0$ and $\lambda_{\min}(B) > 0$ are the minimum eigenvalues of A and B, respectively. If $x = 0$, it is trivial to verify $f^*(x) = 0$. Now we assume $x \neq 0$. Since

$$\max_{\epsilon}\{x^T(\epsilon x) - f(\epsilon x)\} = \max_{\epsilon}\{x^Tx\epsilon - (x^TAx)(x^TBx)\epsilon^2\} = \frac{(x^Tx)^2}{(x^TAx)(x^TBx)} > 0,$$

we have $f^*(x) > 0$. The proof is complete. □

Lemma 2 *Suppose $x \neq 0$. y is a stationary point of $x^T y - f(y)$ if and only if*

$$x^T y - f(y) = p(\alpha),$$

where α is a real root of the univariate equation (1).

Proof. Suppose y is a stationary point of $x^T y - f(y)$, we have

$$x = \nabla f(y) = \left(\frac{1}{2}(y^T Ay)B + \frac{1}{2}(y^T By)A\right) y.$$

The assumption $x \neq 0$ implies that $y \neq 0$. Therefore,

$$y = 2\left(\beta B + \gamma A\right)^{-1} x, \tag{2}$$

where $\beta = y^T Ay > 0$ and $\gamma = y^T By > 0$. Substituting (2) into $y^T Ay$ and $y^T By$, we have

$$\beta = 4x^T \left(\beta B + \gamma A\right)^{-1} A \left(\beta B + \gamma A\right)^{-1} x, \tag{3}$$

$$\gamma = 4x^T \left(\beta B + \gamma A\right)^{-1} B \left(\beta B + \gamma A\right)^{-1} x. \tag{4}$$

Dividing (3) by (4) yields

$$\frac{\beta}{\gamma} - \frac{x^T(A + \frac{\beta}{\gamma}B)^{-1}A(A + \frac{\beta}{\gamma}B)^{-1}x}{x^T(A + \frac{\beta}{\gamma}B)^{-1}B(A + \frac{\beta}{\gamma}B)^{-1}x} = 0.$$

That is, $\alpha := \frac{\beta}{\gamma}$ is a real root of (1). According to (3) and (4), we have

$$\begin{aligned} & x^T(A + \alpha B)^{-1}x \\ = {} & x^T(A + \alpha B)^{-1}A(A + \alpha B)^{-1}x + \alpha x^T(A + \alpha B)^{-1}B(A + \alpha B)^{-1}x \\ = {} & \frac{1}{4}\beta\gamma^2 + \frac{1}{4}\alpha\gamma^3 \\ = {} & \frac{1}{2}\alpha\gamma^3, \end{aligned}$$

which implies that

$$\gamma = \left(\frac{2x^T(A + \alpha B)^{-1}x}{\alpha}\right)^{1/3}. \tag{5}$$

Therefore, it follows from (2), (5) and the definitions of β, γ and α that

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

On the other hand, let α be a real root of (1). Define γ as in (5). Let $\beta = \alpha\gamma$. Define y as in (2). Then y is a stationary point. The proof is complete. □

Lemma 3 *Suppose $x \neq 0$. $x^T y - f(y)$ has at most $2n - 1$ stationary points, which can be solved to any given precision in polynomial time.*

Proof. According to Lemma 2, the number of stationary points of $x^T y - f(y)$ is equal to the number of the real roots of Equation (1). Since A, B are positive definite, there exists a nonsingular matrix P such that both $P^T AP$ and $P^T BP$ are diagonal matrices. Assume $P^T AP = \mathrm{Diag}(a_1, \ldots, a_n)$ and $P^T AP = \mathrm{Diag}(b_1, \ldots, b_n)$. Then for $i = 1, \ldots, n$, $a_i > 0$ and $b_i > 0$. Let $z = P^T x$. Since P is nonsingular and $x \neq 0$, we have $z \neq 0$. It is trivial to verify that each root of Equation (1) satisfies

$$\left(\prod_{i=1}^{n}(a_i+\alpha b_i)^2\right)\left(\alpha\sum_{i=1}^{n}\frac{b_i z_i^2}{(a_i+\alpha b_i)^2}-\sum_{i=1}^{n}\frac{a_i z_i^2}{(a_i+\alpha b_i)^2}\right)=0, \tag{6}$$

which is a univariate polynomial equation of degree $(2n-1)$. Therefore, Equation (1) has at most $2n-1$ real roots, which can be solved to any given precision in polynomial time [2]. The proof is complete. □

Now we present our main result.

Theorem 3 *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 $x \neq 0$,*

$$f^*(x) = p(\alpha^*),$$

where α^* is the maximizer of $p(\alpha)$, solved in polynomial time.

Proof. Since the maximizer of $x^T y - f(y)$ must be a stationary point, it follows from Lemmas 1 and 2 that $f^*(x) = p(\alpha^*)$, where $\alpha^* = arg\max_{g(\alpha)=0} p(\alpha)$. It is not difficult to verify that the equations $\frac{dp(\alpha)}{d\alpha} = 0$ and $g(\alpha) = 0$ are equivalent. Therefore, $\alpha^* = arg\max p(\alpha)$. Moreover, α^* can be solved by enumerating all the roots of $g(\alpha) = 0$, which is done in polynomial time according to the proof of Lemma 3. □

As a corollary, we show that Theorem 2 is improved.

Theorem 4 *([3]) If $f(y)$ is convex, then for any $x \neq 0$, there exists a unique real root to the equation (1).*

Proof. If $f(y)$ is convex, a stationary point of $x^T y - f(y)$ is also a maximizer. It follows from Lemma 2 and Theorem 3 that any stationary point of $p(\alpha)$ is also a global maximizer. According to Lemma 3 and the equivalence between $\frac{dp(\alpha)}{d\alpha} = 0$ and $g(\alpha) = 0$, $p(\alpha)$ has $k(\leq 2n-1)$ stationary points, denoted by $\alpha_1 < \ldots < \alpha_k$. It is sufficient to show $k = 1$. Suppose this is not true, i.e., $k \geq 2$. Since $p(\alpha_1) = p(\alpha_2)$ are the maximum value of $p(\alpha)$, for any $\alpha \in (\alpha_1, \alpha_2)$, we have $p(\alpha) < p(\alpha_1)$. Then there is a local minimizer in (α_1, α_2), which is a stationary point and can not be a maximizer. Now, we obtain a contradiction. The proof is complete. □

3 Conclusion

The Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms was posted as an open question by Hiriart-Urruty [1]. Under a convex assumption on the function, it was answered by Zhao [3]. In this note, we give an answer to the open question without making the convexity assumption. Our proof is much shorter than that of [3]. Finally, we raise an open question what is the expression of the conjugate of

$$f(y) = \frac{1}{4}(y^T Ay)(y^T By) + \frac{1}{2}y^T Cy,$$

where A, B, C are positive definite matrices.

References

1. J.B. Hiriart-Urruty, Potpourri of conjectures and open questions in nonlinear analysis and optimization, SIAM Review, 49, 255-273, (2007)
2. M. Sagraloff, When Newton meets Descartes: A Simple and Fast Algorithm to Isolate the Real Roots of a Polynomial, in Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation, 297-304, ACM New York, NY, USA, 2012
3. Y.B. Zhao, The Legendre-Fenchel conjugate of the product of two positive-definite quadratic forms, SIAM J. Matrix Analysis & Applications, 31(4), 1792-1811 (2010)