Far East Journal of Mathematical Education

Published Online: July 2015
http://dx.doi.org/10.17654/FJMEAug2015_027_033
Volume 15, Number 1, 2015, Pages 27-33 **ISSN: 0973-5631**

THE CONJUGATE DUALITY OF NONLINEAR SEMIDEFINITE PROGRAMMING

Xunzhi Zhu

School of Science
Shandong University of Technology
Zibo City, Shandong Province, 255049
P. R. China

Abstract

Nonlinear semidefinite programming, as a generalization of nonlinear programming, has received considerable attention in recent years. In this paper, we consider the conjugate duality of nonlinear semidefinite programming and its weakly and strongly duality. At last, we present the conjugate dual problem of a semidefinite least-squares problem.

0. Introduction

Semidefinite programming (SDP) is aimed to minimize (or maximize) a linear function of a matrix variable X over an affine subspace of symmetric matrices subject to the constraint that X be positive semidefinite. Recently, SDP has received more and more attention. One of the main reasons is the large variety of applications leading to SDP. However, in many cases, the mathematical models lead to problems, which cannot be formulated as linear, but as nonlinear semidefinite programming problems (NSDP).

Received: March 25, 2015; Accepted: May 2, 2015

Keywords and phrases: nonlinear semidefinite programming, conjugate dual, semidefinite least-squares problem.

Communicated by Mankeun Lee

Duality is an important aspect of optimization theory. There are several dualities, such as the classical Lagrange duality, Fenchel duality and Fenchel-Lagrange duality. Conjugate duality is of great importance for the three types. In this paper, we consider the conjugate dual of NSDP.

1. Conjugate Duality

The nonlinear semidefinite programming (NSDP) is

$$\min\ f(x)$$

$$\text{(NSDP)s.t. } h(x) = 0,$$

$$g(x) \in S_+^n,$$

$$x \in X,$$

where $f : X \to R$, $h : X \to R^m$, $g : X \to S_+^n$ are known functions, $x \in X$ is the decision variables, $g(x) \in S_+^n$ means $g(x)$ is a semidefinite matrix, X is a finite dimensional Hilbert space. The parameterized problem of (NSDP), is

$$\min_{x \in X}\ f(x)$$

$$(P_{a, A})\text{s.t. } h(x) - a = 0,$$

$$g(x) - A \in S_+^n,$$

where $a \in R^m$, $A \in S^n$ are any given parameters. Denote $v(a, A)$ is the optimal solution function of $(P_{a, A})$ on (a, A). Obviously, if $a = 0, A = 0$, $(P_{a, A})$ is the problem (NSDP) and $v(0, 0)$ is the optimal value of (NSDP), denoted P_c^*.

Definition 1. The problem

$$\text{(CD)}\ \sup_{y \in R^m, S \in S^n} \{-v^*(y, S)\}$$

is called the *conjugate dual* of (NSDP), where v^* is the conjugate function of *v*. Denote the optimal solution of (CD) as d_c^*.

Conjugate function of $f(x)$ is defined as $f^*(x) = \sup_{y \in R^n} \{\langle x, y \rangle - f(y)\}$.

We have the following results on conjugate dual of NSDP.

Theorem 1. (i) *Weakly duality holds for conjugate duality*.

(ii) *If* d_c^* *is finite, then the corresponding optimal solution set is the subdifferential of* v^{**}, $\partial v^{**}(0, 0)$.

(iii) *If* $v(a, A)$ *is subdifferentiable at* $(0, 0)$, *then* $P_c^* = d_c^*$ *and the optimal solution set of* (CD) *is* $\partial v(0, 0)$.

(iv) *If* $P_c^* = d_c^*$ *and is finite, then the optimal solution set of* (CD) *is* $\partial v(0, 0)$, (*where the subdifferentiable is defined as* $\partial f(x) = \{\xi \in R^n \mid f(y) \geq f(x) + \langle \xi, y - x \rangle, \forall y \in R^n\}$).

Proof. (i) By the definition of conjugate function, for any $y \in R^m$, $S \in S^n$, we have

$$
\begin{aligned}
-v^*(y, S) &= -\sup_{a \in R^m, A \in S^n} \{\langle y, a \rangle + \langle S, A \rangle - v(a, A)\} \\
&= \inf_{a \in R^m, A \in S^n} \{v(a, A) - \langle y, a \rangle - \langle S, A \rangle\} \\
&= \inf_{a \in R^m, A \in S^n, x \in X} \{f(x) - \langle y, a \rangle - \langle S, A \rangle : h(x) = a, g(x) - A \in S_n^+\} \\
&= \inf_{x \in X} \{f(x) - \langle y, h(x) \rangle - \langle S, g(x) \rangle\} \\
&\quad + \inf_{a \in R^m, A \in S^n, h(x) = a, g(x) - A \in S_n^+} \{\langle h(x) - a, y \rangle + \langle g(x) - A, S \rangle\}
\end{aligned}
$$

$$= \inf_{x \in X} \{L(x, y, S) + \inf_{a \in R^m, g(x) - A \in S_n^+} \{\langle g(x) - A, S \rangle\}\}$$

$$= \begin{cases} \inf_{x \in X} \{L(x, y, S)\}, \text{ if } S \in S_n^+; \\ -\infty, \text{ otherwise,} \end{cases}$$

where $L(x, y, S) = f(x) - \langle y, h(x) \rangle - \langle S, g(x) \rangle$ is Lagrange function.

Then if d_c^* is finite,

$$\sup_{y \in R^m, S \in S^n} \{-v^*(y, S)\} = \sup_{y \in R^m, S \in S^n} \inf_{x \in X} \{L(x, y, S)\} = d_c^*.$$

Notice that

$$\sup_{y \in R^m, S \in S_+^n} \{-v^*(y, S)\} = \sup_{y \in R^m, S \in S_+^n} \{\langle 0, y \rangle + \langle 0, S \rangle - v^*(y, S)\} = v^{**}(0, 0).$$

By the conjugate theorem, $P_c^* = v(0, 0) \geq v^{**}(0, 0) = d_c^*$.

(ii) By the definition of conjugate function and subdifferential, we have

$$d_c^* = v^{**}(0, 0) = \sup_{y \in R^m, S \in S^n} \{-v^*(y, S)\},$$

$$\partial v^{**}(0, 0) = \{(y, S) \in R^m \times S^n : v^{**}(a, A) - v^{**}(0, 0) \geq \langle a, y \rangle + \langle S, A \rangle,$$

$$\forall a \in R^m, \forall A \in S^n\}. \tag{1}$$

Then $(\bar{y}, \bar{S}) \in \partial v^{**}(0, 0)$ is equivalent to

$$(0, 0) \in \arg \min_{a \in R^m, A \in S^n} \{v^{**}(a, A) - \langle \bar{y}, a \rangle - \langle \bar{S}, A \rangle\}$$

$$= \arg\max_{a \in R^m, A \in S^n} \{-v^{**}(a, A) + \langle \bar{y}, a \rangle + \langle \bar{S}, A \rangle\}.$$

Combined with $v^{***}(\bar{y}, \bar{S}) = \sup_{a \in R^m, A \in S^n} \{\langle a, \bar{y} \rangle + \langle A, \bar{S} \rangle - v^{**}(a, A)\}$ and $v^{***} = v^*$, we know that

$$v^{**}(0,\ 0) = -v^{*}(\bar{y},\ \bar{S}) \Leftrightarrow (\bar{y},\ \bar{S}) \in \partial v^{**}(0,\ 0). \tag{2}$$

By (i), (ii) is proved.

(iii) If there exists $\langle \hat{y},\ \hat{S} \rangle \in \partial v(0,\ 0)$, that is,

$$v(a,\ A) - v(0,\ 0) \geq (a,\ \bar{y}) + \langle A,\ \bar{S} \rangle, \quad \forall a \in R^m,\ \forall A \in S^n,$$

which equivalent to $(0,\ 0) \in \arg\max_{a \in R^m,\ A \in S^n} \{\langle \bar{y},\ a \rangle + \langle \bar{S},\ A \rangle - v(a,\ A)\}.$

By the definition of conjugate function, we have

$$v^{*}(\hat{y},\ \hat{S}) = \max_{a \in R^m,\ A \in S^n} \{\langle \bar{y},\ a \rangle + \langle \bar{S},\ A \rangle - v(a,\ A)\}.$$

Then

$$v(0,\ 0) = -v^{*}(\hat{y},\ \hat{S}) \Leftrightarrow (\hat{y},\ \hat{S}) \in \partial v(0,\ 0). \tag{3}$$

Notice that $-v^{*}(\hat{y},\ \hat{S}) \leq v^{**}(0,\ 0)$, and the property of conjugate function $v(0,\ 0) \geq v^{**}(0,\ 0)$, we have $v(0,\ 0) = v^{**}(0,\ 0)$. Then $P_c^{*} = d_c^{*}$ and $(\hat{y},\ \hat{S}) \in \arg\max_{a \in R^m,\ A \in S^n} \{-v^{*}(y,\ S)\}$. By the equivalence of (3), $\partial v(0,\ 0)$ is the optimal solution of (CD).

(iv) If $P_c^{*} = d_c^{*}$ and is finite, by (2) and (3) $\partial v(0,\ 0) = \partial v^{**}(0,\ 0)$. Then (iii) can yield (iv). □

2. Example

Example 1. Semidefinite least-squares problem has a wide application in finance, control and image processing. Here we will compute its conjugate dual:

$$\begin{aligned} & \min \frac{1}{2} \| X - C \|_F^2 \\ \text{(SDLS)} & \text{s.t. } \langle A_i,\ X \rangle = b_i,\ i = 1, \ldots, m, \\ & X \in S_+^n. \end{aligned}$$

It can be easily verified that $f(X) = \frac{1}{2}\| X - C \|_F^2$ is convex function, $h(X) = (\langle A_1, X\rangle, ..., \langle A_m, X\rangle)^T - b$ is affine function, where $b = (b_1, ..., b_m)$, $g(X) = X$ is matrix convex function. So this problem is a convex semidefinite programming. By the proof of Theorem 1, we need to calculate the corresponding Lagrange function:

$$L(X, y, S) = \frac{1}{2}\| X - C \|_F^2 - \Sigma_{i=1}^m y_i(\langle A_i, X\rangle - b_i) - \langle S, X\rangle, \tag{4}$$

$$X \in S_+^n, \quad y \in R^m, \quad S \in S_+^n.$$

As $L(X, y, S)$ is the strictly convex quadratic function on X, its minimizer is obtained at $\nabla_X L(X, y, S) = 0$. Let $0 = \nabla_X L(X, y, S) = X - C - \Sigma_{i=1}^m y_i A_i - S$. Substituting it into (4) yields

$$\begin{aligned} L(X, y, S) &= \frac{1}{2}\| X - C \|_F^2 - \langle \Sigma_{i=1}^m y_i A_i, X\rangle + \langle b, y\rangle - \langle S, X\rangle \\ &= \frac{1}{2}\| X \|_F^2 + \frac{1}{2}\| C \|_F^2 - \left\langle C + \sum_{i=1}^m y_i A_i + S, X \right\rangle + \langle b, y\rangle \\ &= \frac{1}{2}\| X \|_F^2 + \frac{1}{2}\| C \|_F^2 - \langle X, X\rangle + \langle b, y\rangle \\ &= -\frac{1}{2}\| X \|_F^2 + \frac{1}{2}\| C \|_F^2 + \langle b, y\rangle \\ &= -\frac{1}{2}\| S + (C + \Sigma_{i=1}^m y_i A_i) \|_F^2 + \frac{1}{2}\| C \|_F^2 + \langle b, y\rangle. \end{aligned}$$

Then the conjugate dual problem is

$$\begin{aligned} &\max_{y\in R^m, S\in S_+^n} \{\inf_X L(X, y, S)\} \\ &= \max_{y\in R^m, S\in S_+^n} \left\{-\frac{1}{2}\| S + (C + \Sigma_{i=1}^m y_i A_i) \|_F^2 + \frac{1}{2}\| C \|_F^2 + \langle b, y\rangle\right\} \\ &= \min_{y\in R^m} \left\{\frac{1}{2}\| S + (C + \Sigma_{i=1}^m y_i A_i) \|_F^2 - \frac{1}{2}\| C \|_F^2 - \langle b, y\rangle\right\}. \end{aligned}$$

So this problem is equivalent to

$$\min_{y \in R^m} \left\{ \frac{1}{2} \| S + (C + \Sigma_{i=1}^m y_i A_i) \|_F^2 - \langle b, y \rangle \right\}.$$

$\| (a_{ij})_{mn} \|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}$ denotes the Frobenius norm of the matrix.

Acknowledgments

Research of this paper is supported by National Natural Science Foundation of China (11271233), Shandong Province Natural Science Foundation (ZR2012AM016).

References

[1] R. T. Rockafellar and R. J. Wets, Variational Analysis, Springer, Berlin, 1998.

[2] D. Klerk, Aspects of Semidefinite Programming, Dordrecht, Kluwer Academic Publishers, 2002.

[3] J. Balick, A dual approach to semidefinite least-squares problems, SIAM J. Matrix Anal. Appl. 26(1) (2004), 272-284.