SYMMETRIC DUAL NONDIFFERENTIABLE PROGRAMS

S. CHANDRA AND I. HUSAIN

Symmetric and selfduality results are established for a general class of nonlinear programs which combine differentiable as well as non-differentiable cases appearing in the literature. Many well known results are deduced as special cases and certain natural extensions are discussed.

1. Introduction

Many authors have studied symmetric and selfduality for differentiable and non-differentiable mathematical programs. Dantzig, Eisenberg and Cottle [3] and Mond [7] studied symmetric duality for a certain class of differentiable programs while Mond [9] and Mehndiratta [6] presented symmetric duality results for certain non-differentiable programs which involve square roots of quadratic forms in the objective function. Mond and Cottle [10] gave selfduality results for the class of problems studied in [3] and Mehndiratta [6] examined selfduality for his problem in the spirit of [10]. General symmetric dual programs have also been studied by Mehndiratta [5] and Hanson [4].

In this paper, we not only unify most of these results on symmetric and selfduality but also construct a general class of symmetric dual nonlinear programs, which gives results corresponding to nonlinear extensions of problems studied by Mond [9] and Mehndiratta [6]. The symmetric dual formulations of problems studied by Mond and Schechter [12] and Mond [8] are also mentioned.

Received 5 May 1981.

2. Notations and statement of the problems

We shall make use of the following notations and terminology in this study.

Let R_{+}^{n} and R_{+}^{m} be positive orthants of R^{n} and R^{m} respectively. Let K be a real valued twice continuously differentiable function defined on an open set in R^{n+m} containing $R_{+}^{n} \times R_{+}^{m}$. Then $\nabla_{1}K(x_{0}, y_{0})$ denotes the gradient vector of K with respect to x at the point (x_{0}, y_{0}) ; that is,

$$\nabla_{\mathbf{1}}^{K}(x_{0}, y_{0}) = (\partial K/\partial x_{1}, \partial K/\partial x_{2}, \dots, \partial K/\partial x_{n})|_{(x_{0}, y_{0})}$$

The $n \times n$ matrix of second order partial derivatives with respect to x_i, x_j evaluated at (x_0, y_0) is denoted by $\nabla_{11} K(x_0, y_0)$; that is,

$$\nabla_{\mathbf{1}\mathbf{1}}K(x_0, y_0) = \left(\partial^2 K/\partial x_i \partial x_j\right) \Big|_{(x_0, y_0)}$$

The symbols $\nabla_{12}K(x_0, y_0)$, $\nabla_{21}K(x_0, y_0)$ and $\nabla_{22}K(x_0, y_0)$ are defined similarly. The function K(x, y) will be called convex-concave if it is convex in x for each fixed y and concave in y for each fixed x. In case x and y both are in R^n , then K(x, y) will be called skew symmetric if K(x, y) = -K(x, y).

We now state the following pair of non-differentiable programs and discuss their duality in subsequent sections. Primal (P):

Minimize
$$F(x, y, w) = K(x, y) - y^T \nabla_2 K(x, y) + (x^T B x)^{\frac{1}{2}}$$

(1) subject to: $-\nabla_2 K(x, y) + C w \ge 0$,

$$(3) x \ge 0,$$

 $(4) y \ge 0;$

Dual (D):

Maximize $G(x, y, z) = K(x, y) - x^T \nabla_1 K(x, y) - (y^T C y)^{\frac{1}{2}}$ (5) subject to: $-\nabla_1 K(x, y) - Bz \leq 0$, (6) $z^T Bz \leq 1$,

$$(7) x \ge 0,$$

$$(8) y \ge 0$$

where

(i) $B \in R^{n \times n}$ and $C \in R^{m \times m}$ positive semidefinite, (ii) z and w are vectors in R^n and R^m respectively, and (iii) K is twice continuously differentiable.

3. Symmetric duality

For notational convenience, the sets of feasible solution of the primal and dual programs are denoted by ${\it C}_P$ and ${\it C}_D$ respectively, that is,

$$C_{p} = \left\{ (x, y, w) \mid x \in \mathbb{R}^{n}, y, w \in \mathbb{R}^{m}, -\nabla_{2}K(x, y) + Cw \ge 0, \\ w^{T}Cw \le 1, x \ge 0, y \ge 0 \right\}$$

and

$$C_{D} = \left\{ (x, y, z) \mid x, z \in \mathbb{R}^{n}, y \in \mathbb{R}^{m}, -\nabla_{\underline{1}} K(x, y) - Bz \leq 0, \\ z^{T} Bz \leq 1, x \geq 0, y \geq 0 \right\}.$$

It can be easily seen that if the dual (D) is recast in the minimization form, then its dual is primal (P). Thus the programs (P) and (D) constitute a pair of symmetric dual programs in the sense of [3].

We shall make use of the following generalized Schwarz inequality, which has been extensively referred to in the literature; for example Mond [9],

(9)
$$(x^{T}Ay) \leq (x^{T}Ax)^{\frac{1}{2}} (y^{T}Ay)^{\frac{1}{2}}$$

where $x, y \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$ is positive semidefinite. The equality in (9) holds, if for some $\lambda \ge 0$, $Ax = \lambda Ay$.

We now prove the following duality relations between (P) and (D).

THEOREM 1 (Weak duality). Let K be convex-concave. Then, for any $(x_0, y_0, w_0) \in C_p$ and $(\bar{x}, \bar{y}, \bar{z}) \in C_p$,

$$F(x_0, y_0, w_0) \ge G(\overline{x}, \overline{y}, \overline{z}) .$$

Proof. By noting the implications of $(x_0, y_0, w_0) \in C_p$ and $(\bar{x}, \bar{y}, \bar{z}) \in C_p$, it follows that

$$\bar{y}^T \nabla_2 K(x_0, y_0) - \bar{y}^T C \omega_0 \leq 0$$

and

$$-x_0^T \nabla_{\mathbf{l}} K(\bar{x}, \bar{y}) - x_0^T B\bar{z} \leq 0$$

which, on addition, gives,

(10)
$$\overline{y}^T \nabla_2 K(x_0, y_0) - x_0^T \nabla_1 K(\overline{x}, \overline{y}) \leq \overline{y}^T C w_0 + x_0^T B \overline{z}$$
.

Now, as in [3], by convexity-concavity and differentiability of K,

$$(11) \quad K(\bar{x}, \bar{y}) = \bar{x}^T \nabla_1 K(\bar{x}, \bar{y}) \leq \left(K(x_0, y_0) - y_0^T \nabla_2 K(x_0, y_0) \right) \\ + \left(\bar{y}^T \nabla_2 K(x_0, y_0) - x_0^T \nabla_1 K(\bar{x}, \bar{y}) \right) .$$

Therefore,

$$\begin{split} G(\bar{x}, \bar{y}, \bar{z}) &= K(\bar{x}, \bar{y}) - x^T \nabla_1 K(\bar{x}, \bar{y}) - (\bar{y}^T C \bar{y})^{\frac{1}{2}} \\ &\leq \left(K(x_0, y_0) - y_0^T \nabla_2 K(x_0, y_0) \right) \\ &+ \left(\bar{y}^T \nabla_2 K(x_0, y_0) - x_0^T \nabla_1 K(\bar{x}, \bar{y}) \right) - (\bar{y}^T C \bar{y})^{\frac{1}{2}} \text{ (using (11))} \\ &= \left(K(x_0, y_0) + y_0^T \nabla_2 (x_0, y_0) + \left(x_0^T B x_0 \right)^{\frac{1}{2}} \right) \\ &+ \left(\bar{y}^T \nabla_2 K(x_0, y_0) - x_0^T \nabla_1 K(\bar{x}, \bar{y}) \right) - \left(x_0^T B x_0 \right)^{\frac{1}{2}} - (\bar{y}^T C \bar{y})^{\frac{1}{2}} \end{split}$$

COROLLARY 1. If $(x_0, y_0, w_0) \in C_p$ and $(\bar{x}, \bar{y}, \bar{z}) \in C_p$ such that $F(x_0, y_0, w_0) = G(\bar{x}, \bar{y}, \bar{w})$, then (x_0, y_0, w_0) and $(\bar{x}, \bar{y}, \bar{z})$ are optimal for programs (P) and (D) respectively.

Before proving the main duality theorem, we note that both programs (P) and (D) can be expressed in the form of non-differentiable programs studied by Mond [8]. In particular (P) can be written as

minimize
$$F(\xi) = f(\xi) + (\xi^T \hat{B}\xi)^{\frac{\pi}{2}}$$

subject to: $g(\xi) \ge 0$

m. 1.

where

$$\xi = \begin{bmatrix} x \\ y \\ w \end{bmatrix} \in R^{n+2m}, \quad \hat{B} = \begin{bmatrix} B & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } g(\xi) = \begin{bmatrix} -\nabla_2 K + Cw \\ 1 - w^T Cw \\ x \\ y \end{bmatrix}.$$

Now invoking Fritz John type necessary optimality conditions [1], [2] for the above minimization problem (P), we get the following lemma.

LEMMA 1. If ξ_0 is optimal to (P), then there exist $r_0 \in R$, $\rho_0 \in R^{2m+n+1}$ and $\hat{Z}_0 \in R^{2m+n}$ such that

$$\begin{split} \rho_0^T g\left(\xi_0\right) &= 0 , \\ r_0 \left(\nabla f\left(\xi_0\right) + \hat{B}\hat{Z}_0\right) &= \nabla \rho_0^T g\left(\xi_0\right) , \\ \hat{Z}_0 \hat{B}\hat{Z}_0 &\leq 1 , \\ \left(\xi_0^T \hat{B}\xi_0\right)^{\frac{1}{2}} &= \xi_0 \hat{B}\hat{Z}_0 , \end{split}$$

$$(r_0, \rho_0) = 0$$
,
 $(r_0, \rho_0) \neq 0$,

where $\rho_0^T = (u_0, \eta_0, \mu_0, \nu_0)$ with $u_0, \nu_0 \in \mathbb{R}^m$, $\lambda_0 \in \mathbb{R}$, $\mu_0 \in \mathbb{R}^n$ and $\hat{z}_0 = (z_0, \alpha_0, \beta_0)$ with $z_0 \in \mathbb{R}^n$, $\alpha_0 \in \mathbb{R}^m$, $\beta_0 \in \mathbb{R}^m$. Expanding the above relations, we get the following

(12)
$$u_0 \nabla_2 K(x_0, y_0) = u_0^T C w_0,$$

$$\lambda_0 \left(1 - \omega_0^T C \omega_0 \right) = 0 ,$$

(14)
$$\mu_0 x_0 = 0$$
,

(15)
$$v_0 y_0 = 0$$
,

(16)
$$r_0 \nabla_1 K(x_0, y_0) + (u_0 - r_0 y_0)^T \nabla_{12} K(x_0, y_0) + r_0 B z_0 = \mu_0$$
,

(17)
$$(u_0 - r_0 y_0)^T \nabla_{22} K(x_0, y_0) = v_0 ,$$

(19)
$$z_0 B z_0 \le 1$$
,

(20)
$$\left(x_0^T B x_0 \right)^{\frac{1}{2}} = x_0^T B x_0 ,$$

(21)
$$(r_0; u_0, \lambda_0, \mu_0, \nu_0) \ge 0$$
,

(22)
$$(r_0, u_0, \lambda_0, \mu_0, \nu_0) \neq 0$$
.

THEOREM 2 (Strong duality). If $(x_0, y_0, w_0) \in C_p$ solves (P) and the matrix $\nabla_{22}K(x_0, y_0)$ is negative definite, then there exist $Z_0 \in R^n$ such that $(x_0, y_0, z_0) \in C_p$ with $F(x_0, y_0, w_0) = G(x_0, y_0, w_0)$. If, in addition, K(x, y) is convex-concave then (x_0, y_0, z_0) solves (D) and Min $F(x, y, w) = F(x_0, y_0, w_0) = G(x_0, y_0, z_0) = Max G(x, y, z)$. Proof. Since (x_0, y_0, w_0) solves (P) by Lemma 1, there exists

$$(r_0, \rho_0, z_0)$$
 satisfying (12)-(22). Now from (17) and (15) we obtain

(23)
$$\{u_0^{-r_0}y_0\}^T \nabla_{22} K\{x_0, y_0\}y_0 = 0$$

Also $u_0 \ge 0$ and $r_0 \ge 0$ together with (17) imply

(24)
$$(u_0 - r_0 y_0)^T \nabla_{22} K(x_0, y_0) u_0 = 0$$

Multiplying (23) by r_0 , and then subtracting from (24), we get

$$(u_0 - r_0 y_0)^T \nabla_{22} K(x_0, y_0) (u_0 - r_0 y_0) \ge 0$$

which is contrary to the negative definiteness of the matrix $\nabla_{22}K(x_0, y_0)$ unless $u_0 = r_0 y_0$. Hence

(25)
$$u_0 = r_0 y_0$$
.

Now multiplying (18) by \boldsymbol{w}_0^T , we get

(26)
$$\omega_0^T C u_0 = 2\lambda_0 \omega_0^T C \omega_0 .$$

It is to be observed here that $r_0 > 0$, for otherwise $u_0 = r_0 y_0 = 0$, and (16), (17) and (26) together with (13) readily imply $\mu_0 = 0$, $\nu_0 = 0$ and $\lambda_0 = 0$ respectively, a contradiction to (22). Now equation (18) with the aid of (25) and the fact $r_0 > 0$, gives

(27)
$$y_{0}^{T}Cw_{0} = \left(y_{0}^{T}Cy_{0}\right)^{\frac{1}{2}} \left(w_{0}^{T}Cw_{0}\right)^{\frac{1}{2}}$$

Also from (13), either $\lambda_0 = 0$, and hence $Cy_0 = 2(\lambda_0/r_0)Cw_0 = 0$ or $w_0^T Cw_0 = 1$. In either case (27) gives

(28)
$$y_0^T C \omega_0 = \left(y_0^T C y_0 \right)^{\frac{1}{2}}$$

From relations (7), (8), (16) and (25) together with $\mu_0 \ge 0$ and $r_0 > 0$ we get

$$-\nabla_{12}K(x_{0}, y_{0}) - Bz_{0} \leq 0 ,$$
$$z_{0}^{T}Bz_{0} \leq 1 ,$$
$$x_{0} \geq 0 ,$$
$$y_{0} \geq 0 ,$$

implying that $(x_0, y_0, z_0) \in C_D$.

Multiplying (16) by $x_0 \ge 0$ and using (25), (14) and $r_0 > 0$ in succession, we get

(29)
$$-x_0^T \nabla_1 K(x_0, y_0) = x_0^T B z_0 .$$

Hence

$$F(x_{0}, y_{0}, w_{0}) = K(x_{0}, y_{0}) - y_{0}^{T} \nabla_{2} K(x_{0}, y_{0}) + \left(x_{0}^{T} B x_{0}\right)^{\frac{1}{2}}$$

$$= K(x_{0}, y_{0}) - y_{0}^{T} C w_{0} + \left(x_{0}^{T} B Z_{0}\right)$$
(using (12), (25) with $r_{0} > 0$ and then (20))
$$= K(x_{0}, y_{0}) - x_{0}^{T} \nabla_{1} K(x_{0}, y_{0}) - \left(y_{0}^{T} C y_{0}\right)^{\frac{1}{2}}$$
 (using (28) and (29))
$$= G(x_{0}, y_{0}, z_{0})$$

The rest of the theorem is an immediate consequence of Corollary 1.

4. Selfduality

We now prove the following selfduality theorem for programs (P) and (D), which is very much in the spirit of Mond and Cottle [10].

As in [10], we shall describe (P) and (D) as dual programs if the conclusion of Theorem 2 is true.

```
THEOREM 3 (Selfduality). If
 (i) K is skew symmetric and
 (ii) C = B ,
```

then the programs (P) and (D) are formally identical. Furthermore if (P)

and (D) are dual programs with optimal solutions (x_0, y_0, w_0) and (x_0, y_0, z_0) respectively, then

$$F(x_0, y_0, w_0) = 0 = G(x_0, y_0, z_0)$$
.

Proof. Recasting the dual (D) in the primal form, we have

Minimize
$$-K(u, v) + u^T \nabla_1 K(u, v) + (v^T C v)^{\frac{1}{2}}$$

subject to: $-\nabla_1 K(u, v) - Bz \leq 0$,
 $Z^T Bz \leq 1$,
 $u \geq 0$,
 $v \geq 0$.

But skew symmetry of K implies that $-\nabla_{l}K(u, v) = \nabla_{2}K(v, u)$. When B = C, problem (D) takes the form

Minimize
$$K(v, u) - u^T \nabla_2 K(v, u) + (v^T B v)^{\frac{1}{2}}$$

subject to: $-\nabla_2 K(v, u) + Bz \ge 0$,
 $v \ge 0$,
 $u \ge 0$,

which shows that (P) and (D) are formally identical.

Hence if (x_0, y_0, z_0) is optimal for (D), then (y_0, x_0, w_0) is optimal for (P) and conversely.

Now it remains to show that $F(x_0, y_0, w_0) = 0$; consider $F(x_0, y_0, w_0)$

$$= K(x_{0}, y_{0}) - y_{0}^{T} \nabla_{2} K(x_{0}, y_{0}) + \left[x_{0}^{T} B x_{0}\right]^{T}$$

$$\geq K(x_{0}, y_{0}) - y_{0}^{T} B \omega_{0} + \left[x_{0}^{T} B x_{0}\right]^{\frac{1}{2}} \text{ (using (1) and (4))}$$

$$\geq K(x_{0}, y_{0}) - \left\{y_{0}^{T} B \omega_{0} - \left(y_{0}^{T} B y_{0}\right)^{\frac{1}{2}} \left[\omega_{0}^{T} B \omega_{0}\right]^{\frac{1}{2}}\right\} + \left[x_{0}^{T} B x_{0}\right]^{\frac{1}{2}} - \left[y_{0}^{T} B y_{0}\right]^{\frac{1}{2}} \left[\omega_{0}^{T} B \omega_{0}\right]^{\frac{1}{2}}$$

$$\geq K(x_{0}, y_{0}) + \left[x_{0}^{T} B x_{0}\right]^{\frac{1}{2}} - \left[y_{0}^{T} B y_{0}\right]^{\frac{1}{2}} \text{ (using (9) and (2))}$$

$$= K(x_{0}, y_{0}) \text{ .}$$

Similarly it can be established

$$G(x_0, y_0, z_0) \leq K(x_0, y_0)$$

Hence, by Theorem 2,

$$K(x_0, y_0) \leq F(x_0, y_0, w_0) = G(x_0, y_0, z_0) \leq K(x_0, y_0)$$

which implies that

 $F(x_0, y_0, w_0) = F(y_0, x_0, w_0) = K(x_0, y_0) = K(y_0, x_0) = -K(x_0, y_0) ,$ and therefore $F(x_0, y_0, w_0) = 0$.

5. Special cases

In this section we consider some special cases of the problems (P) and (D) by choosing particular forms of the function K(x, y) and the matrices B and C.

(i) For B = 0 = C, programs (P) and (D) reduce to the symmetric dual pair of Dantzig, Eisenberg and Cottle [3]. The symmetric dual pair of Mond [7] is also obtained under the same condition B = 0 = C because, as observed by Mond and Hanson [11], addition or omission of $y \ge 0$ in (D) and $x \ge 0$ in (P) is not an essential difference.

(ii) For $K(x, y) = p^T x + b^T y - y^T A x$, where $p \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$ the programs (P) and (D) reduce to the symmetric dual pair of Mond [9] and that of Mehndiratta [6].

(iii) For B = 0 = C and $K(x, y) = f(x) + g(y) - y^{T}Ax$, the programs (P) and (D) reduce to general symmetric dual programs of Mehndiratta [5].

6. Certain extensions

This section presents certain generalizations of the symmetric dual pair considered in Section 2. These generalizations can be viewed as nonlinear extensions of problems considered by Mond [8] and Mehndiratta [5] and also as natural symmetric dual formulations for problems studied by Mond [8] and Mond and Schechter [12]. The proofs of duality results are not given here because they follow exactly on the lines of Section 3 and

Section 4 except that the results of Mond and Schecheter [12] will also be required along with the results of Craven and Mond [1], [2].

(i) Symmetric dual pair of Mond's problem: Primal (P₀):

Here functions $h: \mathbb{R}^m \to \mathbb{R}^n$ and $\mathcal{l}: \mathbb{R}^n \to \mathbb{R}^m$ are differentiable convex and concave respectively and remaining symbols have the same meaning as in Section 2. If l = 0 and h = 0, the programs (P₀) and (D₀) reduce to the symmetric dual programs of Section 2.

(ii) Symmetric dual pair of Mond and Schechter's problem:Primal (P₁):

Dual (D1)

Maximize $L(x, y, z) = K(x, y) - x^T \nabla_1 K(x, y) - y^T \mathcal{I}(x)$ $+ x^T [\nabla y^T \mathcal{I}(x)] - ||Ny||_q$

 $\| z \|_p \leq 1 \ ,$ $x, \ y \geq 0 \ .$ Here functions h and l are the same as in (P_0) and (D_0), M and N

Here functions n and l are the same as in (P₀) and (D₀), M and N are $m \times n$ matrices and p-norm is given by

$$\|\alpha\|_{p} = \left(\sum_{i=1}^{m} |\alpha_{i}|^{p}\right)^{1/p}$$

Similarly for q with $p^{-1} + q^{-1} = 1$, we define

subject to: $\nabla y^T l(x) - \nabla_1 K(x, y) + h(y) - Mz \leq 0$,

$$\|\boldsymbol{\beta}\|_{q} = \left(\sum_{i=1}^{m} |\boldsymbol{\beta}_{i}|^{q}\right)^{1/q}$$

References

- B.D. Craven and B. Mond, "Lagrangean conditions for quasidifferentiable optimization", Survey of mathematical programming, Vol. 1, 177-192 (Proc. Ninth Internat. Math. Programming Sympos., Budapest, 1976. North-Holland, Amsterdam, 1979).
- [2] B.D. Craven and B. Mond, "Sufficient Fritz John optimality conditions for nondifferentiable convex programming", J. Austral. Math. Soc. Ser. B 19 (1975-1976), 462-468.
- [3] G.B. Dantzig, E. Eisenberg, and R.W. Cottle, "Symmetric dual nonlinear programs", *Pacific J. Math.* 15 (1965), 809-812.
- [4] M.A. Hanson, "Duality and self-duality in mathematical programming", SIAM J. Appl. Math. 12 (1964), 446-449.
- [5] S.L. Mehndiratta, "General symmetric dual programs", Oper. Res. 14 (1966), 164-172.
- [6] S.L. Mehndiratta, "Symmetric and self-duality in nonlinear programming", Numer. Math. 10 (1967), 103-109.
- [7] Bertram Mond, "A symmetric dual theorem for non-linear programs", Quart. Appl. Math. 23 (1965), 265-269.
- [8] Bertram Mond, "A class of nondifferentiable mathematical programming problems", J. Math. Anal. Appl. 46 (1974), 169-174.
- [9] B. Mond, "Symmetric duality for nonlinear programming", Opsearch 13 (1976), 1-10.
- [10] Bertram Mond and Richard W. Cottle, "Self-duality in mathematical programming", SIAM J. Appl. Math. 14 (1966), 420-423.
- [11] Bertram Mond and Morgan A. Hanson, "Symmetric duality for variational problems", J. Math. Anal. Appl. 23 (1968), 161-172.
- [12] Bertram Mond and Murray Schechter, "A programming problem with an L p norm in the objective function", J. Austral. Math. Soc. Ser. B 19 (1975/1976), 332-342.

Department of Mathematics, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India.