Z. Wahrscheinlichkeitstheorie verw. Geb. 14, 257–263 (1970)

Dual Nonlinear Programming Problems in Partially Ordered Banach Spaces

K. RITTER

Summary. The concept of duality plays an important role in mathematical programming and has been studied extensively in a finite dimensional Eucledian space, (see e.g. [13, 4, 6, 8]). More recently various dual problems with functionals as objective functions have been studied in infinite dimensional vector spaces [5, 7, 1, 10, 12].

In this note we consider a nonlinear minimization problem in a partially ordered Banach space. It is assumed that the objective function of this problem is given by a (nonlinear) operator and that its feasible domain is defined by a system of (nonlinear) operator inequalities. In analogy to the finite dimensional case we associate with this minimization problem a dual maximization problem which is defined in the Cartesian product of certain Banach spaces. It is shown that under suitable assumptions the main results of the finite dimensional duality theory can be extended to this general case. This extension is based on optimality conditions obtained in [11].

1. Statement of the Problem

Let $X, Y_0, Y_1, \ldots, Y_m$ be real Banach spaces. Suppose Y_0 is reflexive. For each $i, j \in \{0, 1, \ldots, m\}$ let L_{ij} denote the space of bounded linear operators mapping Y_i into Y_j.

A subset K of Y_j is a convex cone (with vertex 0) if $K+K \subset K$ and $\lambda K \subset K$ for every positive number λ.

For $j \in \{0, 1, \ldots, m\}$ let $K_j \subset Y_j$ be a closed convex cone. It is assumed that, for $j \neq 0$, K_j has interior points and that K_0 has the following properties:

1) If $f(K_0)=0$ then $f=0$ for every $f \in Y_0^*$, where Y_0^* denotes the space of continuous linear functionals over Y_0.

2) There is a $\delta>0$ such that $\|y_1+y_2\| \geqq \delta$ whenever $y_1, y_2 \in K_0$ and $\|y_1\| = \|y_2\| = 1$.

The cone K_j can be used to define a partial ordering for Y_j. We write $y \geqq 0$ if and only if $y \in K_j$. Hence, $y_1 \geqq y_2$ if and only if $y_1 - y_2 \in K_j$. If K_j has interior points we write $y > 0$ if and only if y is an interior point of K_j.

It should be noted that property 2) of K_0 implies that $y=0$ whenever $y \in K_0$ and $-y \in K_0$. However, this need not be true for the other cones K_j.

An element $A \in L_{ij}$ is said to be positive, $A \geqq 0$, if $A(K_i) \subset K_j$. Let $F: X \to Y_0$ and $g_j: X \to Y_j$ for $j=1, \ldots, m$. If $F(x)$ or $g_j(x)$ is Fréchet-differentiable at x_0 we denote its Fréchet derivative at x_0 by $F'(x_0)$, respectively $g_j'(x_0)$.

A mapping $h(x)$ of X into Y_j is said to be convex if for every $x_1, x_2 \in X$

$$\lambda h(x_1) + (1-\lambda) h(x_2) \geqq h(\lambda x_1 + (1-\lambda) x_2)$$

for every $\lambda \in E^1$ such that $0 \leqq \lambda \leqq 1$.

If for every $y \in Y_j$ the set

$$\{x \in X \mid h(x) \geqq y\}$$

is convex, then $h(x)$ is called quasi-concave.

Suppose $h(x)$ is Fréchet-differentiable. According to [9] $h(x)$ is called pseudo-convex if for every $x_1, x_2 \in X$

$$h'(x_1)(x_2 - x_1) \geqq 0 \quad \text{implies} \quad h(x_2) \geqq h(x_1).$$

It is not difficult to show that $h(x)$ is pseudo-convex if it is convex and Fréchet-differentiable. But a pseudo convex mapping need not be convex.

Let

$$R = \{x \in X \mid g_j(x) \geqq 0, j = 1, \ldots, m\}.$$

We consider the following minimization problem which is referred to as the primal problem, PP.

Determine an $x_0 \in R$ such that

$$F(x) \geqq F(x_0)$$

for all $x \in R$.

Any $x \in R$ is said to be a feasible solution to the PP. If $x_0 \in R$ and

$$F(x) \geqq F(x_0) \qquad \text{for all } x \in R$$

then x_0 is called an optimal solution of the PP.

With this minimization problem we associate a maximization problem, which is referred to as the dual problem, DP, as follows.

Let

$$\Omega = \{(x, T_1, \ldots, T_m) \in X \times L_{10} \times \cdots \times L_{m0} \mid F'(x) - \sum_{j=1}^{m} T_j \circ g_j'(x) = 0,\ T_j \geqq 0, j = 1, \ldots, m\}.$$

Determine an $(x_1, T_{11}, \ldots, T_{m1}) \in \Omega$ such that

$$F(x_1) - \sum_{j=1}^{m} T_{j1} \circ g_j(x_1) \geqq F(x) - \sum_{j=1}^{m} T_j \circ g_j(x) \tag{1}$$

for all $(x, T_1, \ldots, T_m) \in \Omega$.

Any $(x, T_1, \ldots, T_m) \in \Omega$ is said to be a feasible solution to the DP. If $(x_1, T_{11}, \ldots, T_{m1}) \in \Omega$ and (1) holds for all $(x, T_1, \ldots, T_m) \in \Omega$, then $(x_1, T_{11}, \ldots, T_{m1})$ is called an optimal solution of the DP.

As in the finite dimensional case we have to assume that the PP satisfies a certain constraint qualification which can be formulated in various ways. For our purposes it is convenient to use the following

Constraint Qualification for the PP. Suppose x_0 is a feasible solution to the PP. Then x_0 is said to satisfy the constraint qualification if for all $T_j \in L_{j0}, j = 1, \ldots, m$, with the property

$$T_j y \geqq 0 \quad \text{whenever} \quad y \in \{K_j - g_j(x_0)\} \tag{$*$}$$

the equation

$$\sum_{j=1}^{m} T_j \circ g_j'(x_0) = 0$$

implies

$$T_j \circ g_j'(x_0) = 0 \qquad \text{for } j = 1, \ldots, m.$$

It should be observed that, if $g_j(x_0) > 0$ for some j, it follows from (*) that $T_j \circ g_j'(x_0) = 0$. Hence, it depends only on the constraints which are active at x_0 whether x_0 satisfies the constraint qualification.

In particular, if $Y_1 = \cdots = Y_m = E^1$, the constraint qualification reduces to the assumption that the zero element of X^* is not in the convex hull of the gradients of the constraints which are active at x_0.

Furthermore, we need a constraint qualification for the DP which we state in the following way:

Constraint Qualification for the DP. A feasible solution $(x_1, T_{11}, \ldots, T_{m1})$ of the dual problem is said to satisfy the constraint qualification if there exists a continuously differentiable mapping $x = h(T_1, \ldots, T_m)$ of some open neighborhood N of $(T_{11}, \ldots, T_{m1})$ into X such that $x_1 = h(T_{11}, \ldots, T_{m1})$ and $(h(T_1, \ldots, T_m), T_1, \ldots, T_m)$ is dual feasible for $(T_1, \ldots, T_m) \in N \cap \{(T_1, \ldots, T_m) | T_j \geqq 0, j = 1, \ldots, m\}$.

The following theorem will be needed in deriving the duality theorems in the next section. It can be obtained as a special case of Theorem (3.2) and (3.3) in [11]. We state it here for the sake of completeness.

Theorem (1.1). *Suppose $F(x)$ and $g_j(x)$, $j = 1, \ldots, m$, are continuously Fréchet-differentiable.*

1. *Let x_0 be an optimal solution of the PP which satisfies the constraint qualification and for which either*

$$g_j(x_0) > 0 \quad or \quad \mathscr{R}(g_j'(x_0)) = Y_j, \qquad j = 1, \ldots, m,$$

holds[1]. *Then there are $T_j \in L_{j0}$ such that*

α) $F'(x_0) = \sum_{j=1}^{m} T_j \circ g_j'(x_0)$,

β) $y \in (K_j - g_j(x_0))$ *implies* $T_j y \geqq 0$ *for* $j = 1, \ldots, m$.

2. *Suppose $F(x)$ is pseudo-convex and $g_j(x)$, $j = 1, \ldots, m$ is quasi-concave. If x_1 is a feasible solution of the PP and if there are $T_j \in L_{j0}$ such that α) and β) hold, then x_1 is an optimal solution of the PP.*

2. The Duality Theorems

First we give a generalization of a theorem which, in the finite dimensional case, is often called the Weak Duality Theorem.

Theorem (2.1). *Suppose $F(x)$ is convex and Fréchet-differentiable and $g_1(x), \ldots, g_m(x)$ are concave and Fréchet-differentiable.*

[1] If A is a bounded linear operator mapping X into Y_j the range of A is denoted by $\mathscr{R}(A)$.

1. *If x_0 is a feasible solution of the PP and $(x_1, T_{11}, \ldots, T_{m1})$ is a feasible solution of the DP then*

$$F(x_0) \geqq F(x_1) - \sum_{j=1}^{m} T_{j1} \circ g_j(x_1). \tag{$*$}$$

2. *If equality holds in $(*)$ then x_0 is an optimal solution of the PP and $(x_1, T_{11}, \ldots, T_{m1})$ is an optimal solution of the DP.*

Proof. 1. Since $F(x)$ is convex

$$F(x_1) + \lambda F(x_0) - \lambda F(x_1) \geqq F[x_1 + \lambda(x_0 - x_1)]$$

for any real number λ such that $0 \leqq \lambda \leqq 1$. Hence, for $0 < \lambda \leqq 1$

$$F(x_0) - F(x_1) \geqq \frac{1}{\lambda} \{F[x_1 + \lambda(x_0 - x_1)] - F(x_1)\}.$$

Since $F(x)$ is differentiable and K_0 is closed this implies

$$F(x_0) - F(x_1) \geqq F'(x_1)(x_0 - x_1).$$

Moreover, since $(x_1, T_{11}, \ldots, T_{m1})$ is a feasible solution of the DP we have

$$F'(x_1) = \sum_{j=1}^{m} T_{j1} \circ g_j'(x_1)$$

and therefore

$$F(x_0) - F(x_1) \geqq \sum_{j=1}^{m} T_{j1} \circ g_j'(x_1)(x_0 - x_1). \tag{1}$$

Since $g_j(x)$ is concave, we have

$$g_j'(x_1)(x_0 - x_1) \geqq g_j(x_0) - g_j(x_1). \tag{2}$$

Thus, (1), (2), $g_j(x_0) \geqq 0$ and $T_{j1} \geqq 0$ imply

$$F(x_0) - F(x_1) \geqq -\sum_{j=1}^{m} T_{j1} \circ g_j(x_1).$$

2. The second part of the theorem follows from the observation that for any feasible solution x of the PP, we have

$$F(x) \geqq F(x_1) - \sum_{j=1}^{m} T_{j1} \circ g_j(x_1) = F(x_0),$$

while for any feasible solution $(x, T_1, \ldots, T_m)$ of the DP, we have

$$F(x_1) - \sum_{j=1}^{m} T_{j1} \circ g_j(x_1) = F(x_0) \geqq F(x) - \sum_{j=1}^{m} T_j \circ g_j(x).$$

The next theorem is an extension of the Duality Theorem.

Theorem (2.2). *Suppose $F(x)$ is convex and continuously differentiable and $g_1(x), \ldots, g_m(x)$ are concave and continuously differentiable. If x_0 is an optimal solution of the PP, which satisfies the constraint qualification and has the property that for $j = 1, \ldots, m$ either*

$$g_j(x_0) > 0 \quad or \quad \mathscr{R}(g_j'(x_0)) = Y_j,$$

then there are $T_j \in L_{j0}$ *such that*

$$(x_0, T_1, \ldots, T_m)$$

is an optimal solution of the DP, and

$$F(x_0) = F(x_0) - \sum_{j=1}^{m} T_j \circ g_j(x_0).$$

Proof. By Theorem (1.1), there are $T_j \in L_{j0}$ such that

$$F'(x_0) = \sum_{j=1}^{m} T_j \circ g_j'(x_0)$$

and

$$T_j y \geqq 0 \qquad \text{whenever } y \in (K_j - g_j(x_0)), \quad j = 1, \ldots, m.$$

Since $g_j(x_0) \geqq 0$, $K_j \subset (K_j - g_j(x_0))$. Hence, $T_j \geqq 0$ and $(x_0, T_1, \ldots, T_m)$ is a feasible solution of the DP.

Furthermore, since $\pm g_j(x_0) \in (K_j - g_j(x_0))$, we have

$$T_j \circ g_j(x_0) \in K_0 \quad \text{and} \quad -T_j \circ g_j(x_0) \in K_0, \quad j = 1, \ldots, m.$$

By the assumption on K_0 this implies

$$T_j \circ g_j(x_0) = 0 \qquad \text{for } j = 1, \ldots, m.$$

Hence the assertion of the theorem follows from Theorem (2.1).

By means of simple counter-examples [9] it can be shown that the above theorem is not true if we only assume that $F(x)$ is pseudo-convex and/or $g_1(x), \ldots, g_m(x)$ are quasi concave. However, for the next theorem, the so called Converse Duality Theorem, these weaker assumptions are sufficient.

Theorem (2.3). *Suppose $F(x)$ is twice continuously differentiable and pseudo-convex and $g_1(x), \ldots, g_m(x)$ are twice continuously differentiable and quasi concave. Let $(x_1, T_{11}, \ldots, T_{m1})$ be an optimal solution of the DP which satisfies the constraint qualification. Then x_1 is an optimal solution of the PP, and*

$$F(x_1) = F(x_1) - \sum_{j=1}^{m} T_{1j} \circ g_j(x_1).$$

Proof. Let $i \in \{1, \ldots, m\}$ be arbitrary but fixed and define

$$\phi(T_i) = h(T_{11}, \ldots, T_{i-1,1}, T_i, T_{i+1,1}, \ldots, T_{m1})$$

$$\psi(T_i) = F(\phi(T_i)) - \sum_{\substack{j=1 \\ j \neq i}}^{m} T_{j1} \circ g_j(\phi(T_i)) - T_i \circ g_i(\phi(T_i)).$$

Since we have

$$\begin{aligned} &T_i \circ g_i(\phi(T_i)) - T_{i1} \circ g_i(\phi(T_{i1})) - T_{i1} \circ g_i'(\phi(T_{i1})) \circ \phi'(T_{i1})(T_i - T_{i1}) - (T_i - T_{i1}) \circ g_i(\phi(T_{i1})) \\ &\quad = T_{i1} \circ (g_i(\phi(T_i)) - g_i(\phi(T_{i1})) - g_i'(\phi(T_{i1})) \circ \phi'(T_{i1})(T_i - T_{i1})) \\ &\qquad + (T_i - T_{i1}) \circ (g_i(\phi(T_i)) - g_i(\phi(T_{i1}))) \\ &\quad = o(\| T_i - T_{i1} \|), \end{aligned}$$

it follows from the chain rule [2] that

$$\begin{aligned}\psi(T_i)-\psi(T_{i1}) &= \psi'(T_{i1})(T_i-T_{i1})+o(\|T_i-T_{i1}\|)\\ &= F'(\phi(T_{i1}))\circ\phi'(T_{i1})(T_i-T_{i1})-\sum_{j=1}^{m} T_{j1}\circ g_j'(\phi(T_{i1}))\\ &\quad \circ\phi'(T_{i1})(T_i-T_{i1})-(T_i-T_{i1})\circ g_i(\phi(T_{i1}))+o(\|T_i-T_{i1}\|)\\ &= \left(F'(x_1)-\sum_{j=1}^{m} T_{j1}\circ g_j'(x_1)\right)\circ\phi'(T_{i1})(T_i-T_{i1})\\ &\quad -(T_i-T_{i1})\circ g_i(x_1)+o(\|T_1-T_{i1}\|).\end{aligned}$$

Since $(x_1, T_{11}, \ldots, T_{m1})$ is an optimal solution of the DP this implies

$$\psi(T_i)-\psi(T_{i1})=-(T_i-T_{i1})\circ g_i(x_1)+o(\|T_i-T_{i1}\|)$$

and

$$(T_i-T_{i1})\circ g_i(x_1)\geqq 0 \tag{1}$$

for $T_i\geqq 0$ and $\|T_i-T_{i1}\|$ sufficiently small.

Suppose $T_{i1}\circ g_i(x_1)\neq 0$. If we choose $T_i=\lambda T_{i1}$, $\lambda\geqq 0$, then

$$(T_i-T_{i1})\circ g_i(x_1)=(\lambda-1)\, T_{i1}\circ g_i(x_1).$$

Since $T_i\geqq 0$ and, for any $\varepsilon>0$, there is $\lambda_0\in\{1-\varepsilon, 1+\varepsilon\}$ such that

$$(\lambda_0-1)\, T_{i1}\circ g_i(x_1)\notin K_0$$

we have a contradiction to (1). Hence,

$$T_{i1}\circ g_i(x_1)=0. \tag{2}$$

Next suppose $g_i(x_1)\notin K_i$. Since K_i is convex and closed there is a continuous linear functional f on Y_i such that

$$f(g_i(x_1))<0 \quad \text{and} \quad f(K_i)\geqq 0$$

[3, Theorem 10, p. 417].

Let $y\in K_0$ and $y\neq 0$. Choose $T_i=\lambda y f+T_{i1}$, $\lambda\geqq 0$. Then $T_i\geqq 0$ and

$$(T_i-T_{i1})\circ g_i(x_1)=\lambda y f(g_i(x_1))\notin K_0 \qquad \text{for } \lambda>0.$$

This is a contradiction to (1) and it follows that

$$g_i(x_1)\geqq 0. \tag{3}$$

Since (3) holds for every $i\in\{1, \ldots, m\}$, x_1 is a feasible solution of the PP. Furthermore, $T_{i1}\geqq 0$ and (2) imply for $i\in\{1, \ldots, m\}$ that

$$T_{i1}\, y\geqq 0 \qquad \text{for } y\in(K_i-g_i(x_1)).$$

Therefore, it follows from Theorem (1.1) that x_1 is an optimal solution of the PP.

References

1. Dieter, U.: Optimierungsaufgaben in topologischen Vektorräumen I: Dualitätstheorie. Z. Wahrscheinlichkeitstheorie verw. Geb. **5**, 89 – 117 (1966).
2. Dieudonné, J.: Foundations of modern analysis. New York: Academic Press 1960.
3. Dunford, N., and J. T. Schwartz: Linear operators, Part I. New York: Interscience Publishers 1966.
4. Hanson, M. A.: A duality theorem in nonlinear programming with nonlinear constraints. Austral. J. Statist. **3**, 64 – 72 (1961).
5. – Infinite nonlinear programming. J. Austral. math. Soc. **3**, 294 – 300 (1963).
6. Huard, P.: Dual programs. IBM J. Res. Develop. **6**, 137 – 139 (1962).
7. Kelly, R. J., and W. A. Thompson Jr.: Quadratic programming in real Hilbert spaces. SIAM J. appl. Math. **11**, 1063 – 1070 (1963).
8. Mangasarian, O. L.: Duality in nonlinear programming. Quart. appl. Math. **20**, 300 – 302 (1962).
9. – Pseudo-convex functions. SIAM J. Control **3**, 281 – 290 (1965).
10. Ritter, K.: Duality for nonlinear programming in a Banach space. SIAM J. appl. Math. **15**, 294 – 302 (1967).
11. – Optimization theory in linear spaces, Part III: Mathematical programming in partially ordered Banach spaces. Math. Ann. **184**, 133 – 154 (1970).
12. Van Slyke, R. M., and R. J. Wets: A duality theory for abstract mathematical programs with applications to optimal control theory. Boeing Scientific Research Laboratories Document D1-82-0671 (1967).
13. Wolfe, P.: A duality theorem for nonlinear programming. Quart. appl. Math. **19**, 239 – 244 (1961).

Privatdozent Dr. Klaus Ritter
Math. Research Center
University of Wisconsin
Madison, Wisc. 53706, USA

(Received September 2, 1968)