

A Study on the Surrogate Duality Theory

Surrogate 쌍대이론에 관한 연구

오 세 호 *

요 약

본 연구에서 고찰한 surrogate relaxation은 Lagrangian relaxation 방법과는 달리 제약식들을 선형 조합으로 묶어 문제를 푼다. 수리계획문제가 convexity를 만족하지 못하는 경우에는 Lagrangian의 경우와 마찬가지로 surrogate gap이 발생한다. Lagrangian 쌍대이론을 토대로 surrogate optimality condition을 알아보고 수리계획법의 특별 형태인 정수선형계획법에 적용해 보았다. 일반적으로 surrogate gap은 Lagrangian gap보다 작기 때문에 좀더 근사하게 된 문제의 최적 목적함수값에 접근할 수 있다. 따라서 branch and bound 알고리즘을 개발할 때 중요한 정보를 제공하는 것이다.

1. Introduction

A surrogate constraint is suggested to get useful information not being captured from the original constraints individually.

Egon Balas²⁾ and Arthur Geoffrion⁴⁾ extended and modified the definition of surrogate constraint originally suggested by F. Glover.

Fred Glover³⁾ showed further extension by reference to those definitions that he had proposed and suggested, which are the stronger surrogate constraints accommodating additional restrictions.

F. Glover¹⁾ has given necessary and sufficient conditions for optimality both with and without the imposition of complementary slackness comparable with those for Lagrangian duality.

H. J. Greenberg and W. P. Pierskalla¹⁴⁾ studied on the general surrogate problem theoretically and suggested some results useful for the construction of the strongest surrogate constraints.

One of the important merits of duals in integer programming is to bracket the primal objective function value. In fact, efficiency of branch-and-bound procedure depends on good bracketing.

M. H. Karwan¹⁶⁾ made his attention to surrogate duality to develop the efficient branch-and-bound procedure.

M. H. Karwan and R. L. Rardin¹⁵⁾ elaborated on relationships between bounds obtained from Lagrangian duals and those derived from the powerful surrogate duals.

Section 2 deals with the basic surrogate duality theory in general mathematical programming Comparing with that of the Lagrangian.

Section 3 considers the linear integer programming as the special case of the general mathematical programming.

2. Basic Theory of Surrogate Duality

Consider the primal problem of mathematical programming given by :

$$P : \min f(\mathbf{x}), \text{ subject to } g(\mathbf{x}) \leq 0, \text{ and } \mathbf{x} \in X,$$

* 청주대학교 산업공학과 전임강사.

(iii) If $SP(S)$ has the integrality property, then

$$v^*(P(\bar{S})) = v^*(L(\bar{S})) = v^*(LD(\bar{S})) = v^*(SD(\bar{S})).$$

where $f(\mathbf{x})$ and each component $g_i(\mathbf{x})$ of the vector $\mathbf{g}(\mathbf{x})$ are real-valued functions defined on X .

A surrogate constraint for P is a linear combination of the component constraints of $\mathbf{g}(\mathbf{x}) \leq 0$ associating a multiplier u_i with each $g_i(\mathbf{x})$.

Then we define the surrogate problem :

$$SP(\mathbf{u}) : \min f(\mathbf{x}), \text{ subject to } \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{x} \in X.$$

where $\mathbf{u} = (u_i)$.

Without loss of generality, we consider $\mathbf{u} \in U = \{\mathbf{u} : \mathbf{u} \geq 0, \sum u_i = 1\}$

Further, we define the optimal objective function value for $SP(\mathbf{u})$:

$$v(\mathbf{u}) = \inf f(\mathbf{x}), \mathbf{x} \in X(\mathbf{u}) = \{\mathbf{x} \in X : \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) \leq 0\},$$

over the set

$$U = \{\mathbf{u} \in U, \inf \{f(\mathbf{x}) : \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{x} \in X\} > -\infty\}.$$

Then clearly, $v(\mathbf{u})$ provides a lower bound on the optimal objective function value for P . Thus choices of the vector \mathbf{u} that provide the greatest values of $v(\mathbf{u})$ yield strongest surrogate constraints in a natural sense, and motivate the definition of the surrogate dual :

$$SD : \max v(\mathbf{u}), \mathbf{u} \in U.$$

The Lagrangian relaxation is

$$L(\mathbf{u}) = \inf \{f(\mathbf{x}) + \mathbf{u} \cdot \mathbf{g}(\mathbf{x}), \mathbf{x} \in X\}$$

The corresponding Lagrangian dual is defined by :

$$LD : \max L(\mathbf{u}), \mathbf{u} \in U.$$

Modifying the definition of $L(\mathbf{u})$ by replacing X with $X(\mathbf{u})$ will result in $L(\mathbf{u}) \leq v(\mathbf{u})$ even if there is possible increasing in $L(\mathbf{u})$ because of the restriction $\mathbf{u} \cdot \mathbf{g}(\mathbf{x}) \leq 0$; that is, $L(\mathbf{u})$ may be regarded as an underestimating function for both the surrogate problem and the primal problem.

The following lemma shows that \mathbf{u} such that \mathbf{x} solves P exists for $SP(\mathbf{u})$ and provides a solution \mathbf{x} to P whenever $L(\mathbf{u})$ has no gap, which is the case with a convex program.

Lemma 1.

If there exists $\mathbf{u}^* \in U$ such that \mathbf{x}_L^* solves P , then \mathbf{x}_L^* solves $SP(\mathbf{u})$ for $\mathbf{u}^* \in U$.

Proof. See 14).

If P is Lagrange regular, then

$$\begin{aligned} & \min \{f(\mathbf{x}), \text{ subject to } \mathbf{u}^* \cdot \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{x} \in X\} \\ & = \min \{f(\mathbf{x}), \text{ subject to } \mathbf{u}^* \cdot \mathbf{g}(\mathbf{x}) = 0, \mathbf{x} \in X\}, \end{aligned}$$

since Lagrange regularity implies the complementary slackness condition. However for the surrogate problem in general,

$$\begin{aligned} & \min \{f(\mathbf{x}), \text{ subject to } \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{x} \in X\} \\ & \leq \min \{f(\mathbf{x}), \text{ subject to } \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) = 0, \mathbf{x} \in X\}. \end{aligned}$$

Now consider an example in which $L(\mathbf{u})$ has a gap, $SP(\mathbf{u})$ does not have a gap.

$$P : \min -x_1 - x_2, \text{ subject to } x_1 + 2x_2 \leq 4, 2x_1 + x_2 \leq 3; x_1, x_2 \geq 0 \text{ and integer.}$$

We can set up $L(\mathbf{u})$ problem as followings ;

$$L(\mathbf{u}) = \min -x_1 - x_2 + \mathbf{u} \cdot (2x_1 + x_2), \text{ subject to } 2x_1 + x_2 \leq 3; x_1, x_2 \geq 0 \text{ and integer.}$$

For $\mathbf{u} = 1/3$, two solutions are $\mathbf{x}^1 = (1, 1)$, $\mathbf{x}^2 = (0, 3)$. The values of $x_1 + 2x_2$ are 3 and 6 respectively. Thus the complementary condition is not satisfied, i.e. a gap occurs.

Next $SP(\mathbf{u})$ problem is given by :

$$SP(\mathbf{u}) : \min -x_1 - x_2, \text{ subject to } (1+u)x_1 + (2-u)x_2 \leq 4-u, 0 \leq u \leq 1, x_1, x_2 \geq 0 \text{ and integer.}$$

Then $SP(\mathbf{u})$ solves P for $u = 1/2$. Thus complementary slackness in the surrogate model is not required.

The next statements show how to detect a surrogate gap.

Assume that $SP(\mathbf{u})$ has a collection $S^*(\mathbf{u})$ of multiple solutions for a given \mathbf{u} . Define $I(\mathbf{u})$ as the index set of the elements of $S^*(\mathbf{u})$. Denote the convex hull of the points $\mathbf{g}(\mathbf{x}_k^*)$, $k \in I(\mathbf{u})$ by $\bar{C}(\mathbf{u})$; that is, $\mathbf{0} \in \bar{C}(\mathbf{u})$ iff $\mathbf{0}$ is a convex combination of the vectors $\mathbf{g}(\mathbf{x}_k^*)$ that are said to span $\bar{C}(\mathbf{u})$.

Theorem 2.1 Suppose there exists $u \in U$ such that $0 \in C(u)$. Then, either there exists $k \in I(u)$ such that x_k^* solves P , or there exists no $u \in U$ such that a solution to $SP(u)$ is a solution to P .

This theorem gives a sufficient condition for gap to occur. With a few additional assumptions, it is possible to provide sufficient conditions for the nonoccurrence of surrogate duality gap.

This gap can be eliminated if the perturbation function is a closed quasiconvex function.

For next theorem define problems related to the P problem and $SP(u)$ problem.:

$$P(y) : \min f(x), \text{ subject to } g(x) \leq y, x \in X.$$

$$P(y, u) : \min f(x) - u \cdot g(x), \text{ subject to } g(x) \leq y, x \in X(u).$$

$$v(y) = \inf f(x), \text{ subject to } g(x) \leq y, x \in X.$$

$$v(y, u) = \inf f(x) - u \cdot g(x), \text{ subject to } g(x) \leq y, x \in X(u).$$

$\bar{v}(y)$ is called the standard perturbation function and $\bar{v}(y, u)$ is named the parametric perturbation function.

For any solution $x_k^* \in S^*(u)$, $K \in I(u)$, let $B_k(u) = \{y : y \cdot g(x_k^*) \text{ and } u \cdot y \leq 0\}$, where 0 is origin.

Theorem 2.2 If $0 \in C(u^*)$, where u^* maximizes $v(u)$, if $f^*(y)$ is a closed quasiconcave function, and if $B_k(u)$ is a closed convex set for all $u \in U$, then $0 \in B_k(u^*)$.

Proof. See 14).

This theorem implies the sufficient condition for nonoccurrence of a gap.

Now discuss the surrogate optimality conditions.

In the above we define perturbation functions.

It should be noted that the objective function $\bar{v}(y, u)$ for $P(y, u)$ may be thought of as an overestimating function in the same sense that the Lagrangian is an underestimating function.

Using the above definition, F. Glover¹⁰⁾ formalized the surrogate optimality conditions as follows :

(i) $u \geq 0$.

(ii) x is optimal for $SP(u)$.

(iii) $g(x) \leq 0$.

Theorem 2.3 The surrogate optimality conditions imply x is optimal for P , u is optimal for SD , and their optimal objective function values are equal. Moreover, if there is a nonnegative u such that $v(0) = \bar{v}(u)$, then the set of all optimal solutions to P is precisely the set of x that together with u satisfy the surrogate optimality conditions.

By analogy with the Lagrangian, the strong optimality conditions are defined as follows :

(iv) $u \cdot g(x) = 0$.

Though the strong optimality conditions are the same as the Lagrangian optimality conditions upon replacing (ii) by the statement that x is optimal for the Lagrangian, those accommodate a somewhat broader range of possibilities than the Lagrangian optimality conditions.

Corollary. The strong optimality conditions imply x is optimal for both P and $P(0, u)$, and $\bar{v}(0, u) = \bar{v}(0)$. Moreover, if any pair u, x satisfies the strong optimality conditions, then these conditions characterize all x that are optimal for $P(0, u)$ and all u such that $\bar{v}(0, u) = \bar{v}(0)$.

For further step, define new concept extended from Lagrangian duality theory.

We will call $\delta(\alpha)$ a parametric subgradient of the function $F(y; \alpha)$ at y^* if $F(y; \alpha) \geq F(y^*; \alpha) + \delta(y - y^*)$ for all y . The parametric subgradient is an improving direction for the function $F(y; \alpha)$ For some given value of the parameter α .

With this definition we can identify necessary and sufficient conditions for the strong optimality conditions to hold. Extending the notion of the parametric subgradient, we define $\delta(\alpha)$ to be a relative subgradient of a function $F(y; \alpha)$ with respect to a function $G(y; \alpha)$ at the point y^* if $F(y; \alpha) \geq G(y^*; \alpha) + \delta(\alpha)(y - y^*)$ for all y . The relative subgradient points in a direction in which the function

Theorem 2.5 Assume $\mathbf{u} \geq 0$. Then the optimum for the overestimating surrogate is greater than or equal to the optimum for P if and only if $\bar{v}(\mathbf{y}, 2\mathbf{u}) \geq \bar{v}(0) - \mathbf{u} \cdot \mathbf{y}$ for all \mathbf{y} .

Since $f(\mathbf{x}) - \mathbf{u} \cdot \mathbf{g}(\mathbf{x})$ always equals or exceeds $f(\mathbf{x})$ for $\mathbf{x} \in X(\mathbf{u})$, the relation $\inf \{f(\mathbf{x}) - \mathbf{u} \cdot \mathbf{g}(\mathbf{x}) : \mathbf{x} \in X(\mathbf{u})\} \geq v(0) \geq \inf \{f(\mathbf{x}) : \mathbf{x} \in X(\mathbf{u})\}$ can be used to provide good lower and upper bound for the optimal objective function value for P. This results are useful in developing efficient branch-and-bound procedures

3. Surrogate Duality in Integer Programming

It is thought to be convenient to use the notations similar to those used in section 2 in this section. Consider the general integer linear programming problem :

$$P : \min \mathbf{c} \cdot \mathbf{x}, \text{ subject to } \mathbf{A} \cdot \mathbf{x} \leq \mathbf{b}, \mathbf{x} \in S$$

where $S = \{\mathbf{x} \geq 0 : \mathbf{G} \cdot \mathbf{x} \leq \mathbf{h}, \mathbf{x} \text{ satisfies some discrete constraints}\}$.

Here A and G are $m \times n$ and $k \times n$ matrices respectively.

We define the surrogate relaxation of P as followings :

$$SP(\mathbf{u}) : \min \mathbf{c} \cdot \mathbf{x}, \text{ subject to } \mathbf{u} \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}) \leq 0, \mathbf{u} \geq 0, \mathbf{x} \in S.$$

$v^*(\cdot)$ denotes the value of an optimal solution to problem (\cdot) .

Then the surrogate dual is :

$$SD : \max v^*(SP(\mathbf{u})), \mathbf{u} \geq 0.$$

And the Lagrangian relaxation is :

$$L(\mathbf{q}) : \min \mathbf{c} \cdot \mathbf{x} + \mathbf{q} \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}), \text{ subject to } \mathbf{x} \in S.$$

The corresponding Lagrangian dual is :

$$LD : \max v^*(L(\mathbf{q})), \mathbf{q} \geq 0.$$

Lagrangian duality in integer programming is based on the concept of an integrality property in the Lagrangian relaxation. Integrality property is relevant when taken in the context of a branch-and-bound algorithm in which a dual is used to bound the restricted version of P obtained when a number of variables have been fixed at specific values.

For the next theorem, define the integrality property.

Definition. The problem $L(\mathbf{q})$ has the integrality property if $v^*(L(\bar{S})) = v^*(L(S))$ for all $\mathbf{q} \geq 0$, i.e., if the Lagrangian relaxation can be solved as a linear program for all $\mathbf{q} \geq 0$.

A bar over the name of any constraint set represents the same set with all integrality requirements relaxed.

Theorem 3.1

- (i) $v^*(P(\bar{S})) = v^*(LD(\bar{S})) \leq v^*(L(\bar{S})) \leq v^*(LD(\bar{S})) = v^*(LD(\bar{C}(\bar{S}))) = v^*(P(\bar{C}(\bar{S})))$.
- (ii) If the problem $L(\mathbf{q})$ has the integrality property, then $v^*(P(\bar{S})) = v^*(L(\bar{S})) = v^*(LD(\bar{S}))$, where $\bar{C}(\bar{S})$ is a convex hull of \bar{S} .

This theorem gives implications for developing Lagrangian strategies for integer programming. (i) implies that the value of the Lagrangian dual may be obtained by solving the linear program formed when S is replaced in P by $\bar{C}(S)$. (ii) indicates that successful applications of Lagrangian duality do occur even when the integrality property is present.

We can also define an integrality property in the surrogate case ; $SP(\bar{S})$ has the surrogate integrality property if $v^*(SP(\bar{S})) = v^*(SP(\bar{S}))$ for all $\mathbf{u} \geq 0$. In the surrogate case the integrality property relates $v^*(LD(\bar{S}))$ to $v^*(SD(\bar{S}))$.

Theorem 3.2

- (i) $v^*(LD(\bar{S})) = v^*(SD(\bar{C}(\bar{S})))$.
- (i) $v^*(P(\bar{S})) = v^*(LD(\bar{S})) = v^*(SD(\bar{S}))$.

$F(\mathbf{y}; \alpha)$ improves relative to $G(\mathbf{y}^*; \alpha)$ for a given α .

Theorem 2.4 Assume $\mathbf{u} \geq \mathbf{0}$. Then the surrogate optimality conditions are met and the conclusions of Theorem 2.3 apply if and only if $-\mathbf{u}$ is a relative subgradient of $\bar{v}(\mathbf{y}, \mathbf{u})$ with respect to $\bar{v}(\mathbf{y})$ at the origin.

Comparison Theorem 2.3 with Theorem 2.4 leads to constructive conclusion which implies that the Lagrangian optimality conditions are met if and only if $-\mathbf{u}$ is a subgradient of $\bar{v}(\mathbf{y})$ at the origin; i.e., $\bar{v}(\mathbf{y}) \geq \bar{v}(0) - \mathbf{u} \cdot \mathbf{y}$ for all \mathbf{y} .

Finally let us mention an optimal surrogate constraint for the overestimating function.

The above theorem gives us some implications. If $SP(\S)$ has the integrality property, neither $LD(\S)$ nor $SD(\S)$ can improve on the bound from $P(\S)$ though they might provide a more efficient way to calculate $v^*(P(\S))$. And if $SP(\S)$ is relaxed and solved as a linear program when it does not have the integrality property, $v(P(\S))$ is gotten as a dual value.

Finally let us mention a composite relaxation which results from the combination of Lagrangian and surrogate relaxations. It is formed as follows;

$P_{\mathbf{u}} : \min \mathbf{c} \cdot \mathbf{x} + \mathbf{q} \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b})$, subject to $\mathbf{u} \cdot (\mathbf{A} \cdot \mathbf{x} - \mathbf{b}) \leq 0$, $\mathbf{x} \in \S$,
with corresponding dual ;

$D : \max v^*(P_{\mathbf{u}})$, $\mathbf{u}, \mathbf{q} \geq \mathbf{0}$.

The characteristics of these functions may determine whether there exist good search procedures to find optimal or near optimal dual multipliers. In section 2, $v^*(SP(\mathbf{u}))$ may be proved a quasiconcave function of \mathbf{u} , which leads to the development of search procedures for optimal surrogate multipliers. But the composite surrogate dual lacks quasiconcavity for $\mathbf{u} \cdot \mathbf{q} \geq 0$. However, M.H. Karwan and R. L. Rardin suggested a simple heuristic approach which alternately optimizes over Lagrangian and surrogate multipliers¹⁶⁾.

4. Conclusion

This paper discusses a surrogate duality theory inviting direct comparison with Lagrangian duality theory. This two duality theories are usually used for solving the problems lack of convexity.

The surrogate approach is more promising than the Lagrangian approach in case that the duality gap is an overriding consideration. In other words, surrogate duality gap is at least as small and often smaller than the Lagrangian duality gap. And surrogate dual solution values may yield excellent bounds in a branch-and-bound procedure. But the bounds would probably be more difficult to obtain because of the combinatoric nature of the surrogate dual problem.

In recent, many researchers focus their efforts on developing the algorithms for optimal surrogate constraints and extending the original dual concepts to reduce the gap in nonconvex problem.

5. Reference

- 1) E. Balas, "Discrete Programming by the Filter Method with Extension to Mixed-Integer Programming and Application to Machine Sequencing," ICC Report N 66/10, International Computation Center, Centre International de Calcul (Nov., 1966).
- 2) _____, "Discrete Programming by the Filter method," *Opns. Res.* 19, 915-917 (1967).
- 3) M. Bellmore, H. J. Greenberg, and J. J. Jarvis, "Generalized Penalty Function Concepts in Mathematical Optimization," *Opns. Res.* 18, 229-252 (1970).
- 4) M. E. Dyer, "Calculating Surrogate Constraints," *Mathematical Programming* 19 (1980) 255-278.
- 5) M. L. Fisher, "Optimal Solution of Scheduling," *Opns. Res.* 21, 1114-1127 (1973).
- 6) A. M. Geoffrion, "Implicit Enumeration Using an Imbedded Linear Program," No. 120, Western Mgt. Sci. Inst. (May, 1967).
- 7) _____, "Lagrangian Relaxation and Its Uses in Integer Programming," *Math. Program. Stud.* 2, 82-114 (1974).
- 8) F. Glover, "Multiphase-dual Algorithm for the Zero-One Integer Programming Problems," *Opns.*

Res. 13, 879 - 919 (1965).

- 9) _____, "Surrogate Constraints," Opns. Res. 16, 741 - 749 (1968).
- 10) _____, "Convexity Cuts," University of Texas, Austin, Dec., 1969.
- 11) _____, "Surrogate Constraint Duality in Mathematical Programming," Opns. Res. 23, 434 - 451 (1975).
- 12) H.J. Greenberg, "The Generalized Penalty Function Surrogate Model," Opns. Res. 21, 162 - 178 (1973).
- 13) _____, "Bounding nonconvex Programs with conjugates," Opns. Res. 21, 346 - 347 (1973).
- 14) _____, and W.P. Pierskalla, "Surrogate Mathematical Programs," Opns. Res. 18, 924 - 939 (1970).
- 15) M. H. Karwan and R. L. Rardin, "Some Relationships between Lagrangian and Surrogate Duality in Integer Linear Programming," Math. Program. 17 (1979) 324 - 334.
- 16) _____, "Searchability of the Composite and Multiple Surrogate Dual Functions," Opns. Res. 28, 1251 - 1257 (1980).
- 17) G.L. Neumhauser and W.B. Widhelm, "A modified linear program for columnar methods in mathematical programming," Opns. Res. 23, 372 - 382 (1975).