MTA SZTAKI Közlemények 35/1986 p. 89-104

DUALITY IN DYNAMIC PROGRAMMING

DINH THE LUC

Computer and Automation Institute
Hungarian Academy of Sciences

1. INTRODUCTION

To give an idea about duality in mathematical programming Hadley (1962) cited the following line from "The Rubayiat" of E. Fitzgeral:

"All of this of Pot and Potter-Tell me
Who is the Potter, pray, and who is the Pot?"

The highlight of duality lies in the reciprocal relations lingking "Pot" and "Potter". To see it let us remember the case of linear programming. Given a linear (minimization) problem, called primal, there is a second, corresponding (maximization) problem, called dual such that

(i) The dual of the dual is the primal
(ii) Any feasible solution of the primal has a value greater or at least equal to the value of any feasible solution of the dual.
(iii) If an optimal solution exists for the primal then one exists for the dual and their optimal values are equal.
(iv) If the primal has an unbounded solution, the dual has no feasible solution and vice versa.

These features have proved to be very important in the theoretical study of optimization problems and their computational practice. For the case of nonlinear programming similar results can be obtained, but of course, suitable conditions have to be imposed. What about case of dynamic programming? To our knowledge very little attention has been directed to this topic (some efforts towards it have been made by using Lagrangian multipliers in solving dynamic problems, (see Bellman (1957)) and a duality theory has been constructed for infinite horizon optimization of linear and convex models (see Weitzman (1973), Evers (1983) and Haneveld (1985)).

It is not the aim of the paper to develop the complete duality theory for dynamic programming. Our purpose is merely to study some duality aspects of a finite-stage dynamic system and in doing so we call attention of researchers to a useful technique for solving dynamic problems. The paper is structured as follows: in the second section we describe a dynamic problem to be considered. The third section is devoted to the Lagrangian functions associated with our problem and their properties. In section 4, another approach to duality theory, a conjugate function approach is presented and finally we give a brief discussion about the possibility of applying duality results in solving the problem.

2. DESCRIPTION OF THE MODEL

Suppose that we have a dynamic system which consists of N stages numbered from 1 to N. At stage $k \in \{1,\dots,N\}$ the system is characterized by a nonempty set of states X_k and a nonempty set of actions $U(x_k)$ corresponding to every $x_k \in X_k$. In the case $k \leq N-1$, under action

$$u_k \in U(x_k) \tag{1}$$

state

$$x_k \in X_k \tag{2}$$

will be transferred into state $x_{k+1} \in X_{k+1}$ of the next stage by the law:

$$x_k = g_k(x_k, u_k). \tag{3}$$

Let us start with x_k and apply action u_k to get x_{k+1}. Keeping in mind requirements (1), (2) and (3) and carrying on the above procedure we get a sequence of states $(x_k,\dots,x_N)$ and a sequence of actions $(u_k,\dots,u_{N-1})$. Such sequences of states and actions are called (N-k)-process and (N-k)-policy, respectively. It is obvious that when x_k and u_k are chosen, x_{k+1} is completely defined.

Further, for any (N-k)-process and (N-k)-policy, let $R_k(x_k,\dots,x_N,u_k,\dots,u_{N-1})$ be the cost of getting $(x_k,\dots,x_N)$ by using policy $(u_k,\dots,u_{N-1})$.

The problem that we are going to deal with is to find an N-policy for a given initial state x_1 such that it minimizes

$$R_1(x_1,\dots,x_N,u_1,\dots,u_{N-1})$$

over all possible N-processes starting with x_1 and all possible N-policies.

For the sake of simplicity we shall assume that the additivity property holds for $R_1,\dots,R_N$, i.e.

$$R_k(x_k,\dots,x_N,u_k,\dots,u_{N-1}) = \sum_{i=1}^{N-1} f_i(x_i,u_i) + f_N(x_N)$$

where $k = 1,\dots,N-1$ and $R_N(x_N) = f_N(x_N)$.

It is well known (see for example Bellman (1957)) that under the additivity assumption Bellman's equations will be satisfied:

$$B_k(x_k) = \min\{f_k(x_k,u_k) + B_{k\ 1}(g_k(x_k,u_k)) : u_k \in U(x_k)\}$$

where $k = 1,\dots,N-1$ and $B_N(x_N) = f_N(x_N)$.

Finally, it will be assumed in the model that constraints $u_k \in U_k$ are explicitly expressed by relations:

$$h_k(x_k, u_k) \leq 0, \quad k = 1, \ldots, N-1 \tag{4}$$

where u_k is taken from an arbitrary space U. For the sake of convenience we will assume that $h_1, \ldots, h_{N-1}$ are scalar--valued, although all the results to be proved are valid for the vector-valued case too.

3. LAGRANGIAN FUNCTIONS

For every $k \in \{1, \ldots, N\}$ and $x_k \in X_k$ let us consider the following problem denoted by $P(k)$:

$$\min \left[\sum_{i=k}^{N-1} f_i(x_i, u_i) + f_N(x_N) \right]$$

$$\text{s.t.} \qquad x_{k+1} = g_i(x_i, u_i)$$

$$h_i(x_i, u_i) \leq 0, \quad i = k, \ldots, N-1.$$

Whenever k is indicated, we shall write x and u instead of $(x_{k\ 1}, \ldots, x_N)$ and $(u_k, \ldots, u_{N-1})$, respectively and

$$\lambda = (\lambda_k, \ldots, \lambda_{N-1}), \quad \mu = (\mu_k, \ldots, \mu_{N-1})$$

are $(N-k-1)$-vectors of real numbers. In what follows, we will not speak of the dimensions of the variables if it is clear from the context.

Associated with $P(k)$ the lagrangian function will be:

$$L_k(x_k,x,u,\lambda,\mu) = \sum_{i=k}^{N-1} [f_i(x_i,u_i) + \lambda_i h_i(x_i,u_i) +$$

$$+ \mu_i(x_{i+1}-g_i(x_i,u_i))] + f_N(x_N),$$

if $\lambda \geq 0$ and

$L_k(x_k,x,u,\lambda,\mu) = -\infty$ otherwise.

Definition 3.1 For a fixed $x_k \in X_k$, $(x^*,u^*,\lambda^*,\mu^*)$ is said to be a saddlepoint for L_k if

$L_k(x_k,x,u,\lambda^*,\mu^*) \geq L_k(x_k,x^*,u^*,\lambda^*,\mu^*) \geq L_k(x_k,x^*,u^*,\lambda,\mu)$

for every x,u,λ and μ.

Lemma 3.1 $(x^*,u^*,\lambda^*,\mu^*)$ is a saddlepoint for L_k if and only if

(i) $L_k(x_k,x^*,u^*,\lambda^*,\mu^*) = \min\{L_k(x_k,x,u,\lambda^*,\mu^*): x,u\}$

(ii) $x^*_{i+1} = g_i(x^*_i,u^*_i)$, $i = k,\ldots,N$,

where $x^*_k = x_k$,

(iii) $\lambda^*_i h_i(x^*_i,u^*_i) = 0$,

$h_i(x^*_i,u^*_i) \leq 0$, $i = k,\ldots,N-1$.

Proof If $(x^*,u^*,\lambda^*,\mu^*)$ is a safflepoint of L_k, then (i) is obvious. Moreover, $L_k(x_k,x^*,\lambda^*,\mu^*) =$

$$= \max\{L_k(x_k,x^*,u^*,\lambda,\mu):\lambda,\mu\}.$$

If one of the conditions in (ii) and (iii) does not hold, then by varying λ and μ suitably, $L_k(x_k, x^*, u^*, \lambda, \mu)$ may increase infinitely, which is impossible. The inverse assertion is trivial. #

Define $p_k(x_k, x, u) = \max \{L_k(x_k, x, u, \lambda, \mu) : \lambda, \mu\}$,

$$d_k(x_k, \ , \ ) = \min \{L_k(x_k, x, u, \lambda, \mu) : x, u\}.$$

It is easy to see that

$$p_k(x_k, x, u) = \sum_{i=k}^{N-1} f_i(x_i, u_i) + f_N(x_N)$$

if (3) and (4) hold and

$$p_k(x_k, x, u) = +\infty \quad \text{otherwise.}$$

Consequently, the problem $\min \{p_k(x_k, x, u) : x, u\}$ will be the same as $P(k)$. We call it the k-primal problem and denote its optimal value by $B_k(x_k)$.

The k-dual problem will be

$$\max \{d_k(x_k, \lambda, \mu) : \lambda, \mu\} \qquad D(k)$$

and itsoptimal value will be denoted by $C_k(x_k)$.

Proposition 3.1 (Weak duality) For every (x, u, λ, μ) we have $d_k(x_k, \lambda, \mu) \leq p_k(x_k, x, u)$.

Proof. This follows immediately from the definitions. #

Proposition 3.2 (x, u, λ, μ) is a saddlepoint for L_k if and only if (x, u) solves the k-primal problem, (λ, μ) solves the k-dual problem and their optimal values are equal.

Proof. Apply Lemma 3.1 to get this proposition. #

Proposition 3.3 Functions $d_k,\dots,d_N$ yield the following recurrence relations:

$$d_i(x_i,\lambda,\mu) = \min\{ f_i(x_i,u_i) + \\ + \lambda_i h_i(x_i,u_i) + \\ + \mu_i(x_{i+1}-g_i(x_i,u_i)) \\ + d_{i+1}(x_{i+1},\lambda',\mu'): x_{i+1},u_i\}$$

if $\lambda_i \geq 0$ and $d_i(x_i,\lambda,\mu) = -\infty$ otherwise;

$i = k,\dots,N-1$ and $d_N(x_N) = f_N(x_N)$,

where the sign "'" denotes a vector for which the first component is omitted.

Proof. If $\lambda \not\geq 0$, then by definition $d_i(x_i,\lambda,\mu) = -\infty$ and there is nothing to prove. Now assume $\lambda \geq 0$. Consider the following dynamic model consisting of stages numbered from k to N. The sets of states are the same as in the model described in Section 2. For every state $x_i \in X_i$ $(i \geq k)$, the set of actions $U(x_i) = (X_{i+1},U)$. Action (x_{i+1},u_i) transfers state x_i into a state of the next stage by the relation:

$$G_i(x_i,(x_{i+1},u_i)) = x_{i+1}.$$

The cost of getting $(x_k,\dots,x_N)$ by using policy $((x_{i+1},u_i),\dots,(x_N,u_{N-1}))$ will be

$$\sum_{i=k}^{N-1} F_i(x_i,(x_{i+1},u_i)) + F_N(x_N),$$

where $F_i(x_i,(x_{i+1},u_i)) = f_i(x_i,u_i) +$

$$+ \lambda_i h_i(x_i,u_i)$$

$$+ \mu_i(x_{i+1}-g_i(x_i,u_i)),$$

$$i = k,\dots,N-1 \quad \text{and} \quad F_N(x_N) = f_N(x_N).$$

Apply Bellman's equations to this model to get the following recurrence relations:

$$B_i(x_i) = \min\{F_i(x_i,(x_{i+1},u_i)) +$$

$$+ B_{i+1}(G_i(x_i,(x_{i+1},u_i))): (x_{i+1},u_i)\}$$

$$i = k,\dots,N-1 \quad \text{and} \quad B_N(x_N) = F_N(x_N).$$

This gives the relations of the proposition. The proof is complete. #

Corollary 3.1 The following relation holds

$$d_k(x_k,\lambda,\mu) = \min\{f_k(x_k,u_k) +$$

$$+ \lambda_k h_k(x_k,u_k) -$$

$$- \mu_k g_k(x_k,u_k)): u_k\} +$$

$$+ \min\{\mu_k x_{k+1}+d_{k+1}(x_{k+1},\lambda',\mu'): x_{k+1}\}.$$

Proof. This follows immediately from Proposition 3.3. #

Proposition 3.4 If $(\bar{x},\bar{u},\bar{\lambda},\bar{\mu})$ is a saddlepoint for L_k with initial point $\bar{x}_k$, then $(\bar{x}',\bar{u}',\bar{\lambda}',\bar{\mu}')$ is a saddlepoint for L_{k+1} with initial point $\bar{x}_{k+1} = g_k(\bar{x}_k,\bar{u}_k)$.

Proof. By virtue of Lemma 3.1 it suffices to show that

$$L_{k+1}(\bar{x}_{k+1},\bar{x}',\bar{u}',\bar{\lambda}',\bar{\mu}') = \min\{L_{k+1}(\bar{x}_{k+1},x',u',\bar{\lambda}',\bar{\mu}'):x',u'\}. \tag{5}$$

Indeed, again by Lemma 3.1, $\bar{\lambda}_k \geq 0$ and $\bar{x}_{k+1} = g_k(\bar{x}_k,\bar{u}_k)$.

It follows from Proposition 3.3 that

$$\begin{aligned} d_k(\bar{x}_k,\bar{\lambda},\bar{\mu}) = f_k(\bar{x}_k,\bar{u}_k) + \bar{\lambda}_k h_k(\bar{x}_k,\bar{u}_k) &+ \\ + \bar{\mu}_k(\bar{x}_{k+1}-g_k(\bar{x}_k,\bar{u}_k)) &+ \\ + d_{k+1}(\bar{x}_{k+1},\bar{\lambda}',\bar{\mu}'). & \end{aligned} \tag{6}$$

Remembering the definition of d_k we have

$$\begin{aligned} d_k(\bar{x}_k,\bar{\lambda},\bar{\mu}) &= L_k(\bar{x}_k,\bar{x},\bar{u},\bar{\lambda},\bar{\mu}) = \\ &= f_k(\bar{x}_k,\bar{u}_k) + \bar{\lambda}_k h_k(\bar{x}_k,\bar{u}_k) + \\ &\quad + \bar{\ }_k(\bar{x}_{k+1}-g_k(\bar{x}_k,\bar{u}_k)) + \\ &\quad + L_{k+1}(\bar{x}_{k+1},\bar{x}',\bar{u}',\bar{\lambda}',\bar{\mu}'). \end{aligned}$$

Compare the latter relation and (6) to get (5). The proof is complete. #

Corollary 3.2 If $B_k(x_k) = C_k(x_k)$, then $B_i(x_i) = C_i(x_i)$ for all $i \geq k$, where $x_{i+1} = g_i(x_i,u_i)$

u_i solves $B_i(x_i)$.

Proof. If $B_k(x_k) = C_k(x_k)$, then any optimal solutions of P(k) and D(k) form a saddlepoint for L_k. Now the assertion implied by Proposition 3.4. #

In the rest of the section, we are concerned with the question of the existence of λ and μ for a given optimal solution (x,u) of P(k) so that (x,u,λ,μ) forms a saddlepoint for L_k. As it is known from the theory of mathematical programming, the answer to this question is not always positive. For λ and μ to exist some more assumptions about the model are needed. As before, let x_k be fixed from X_k.

Proposition 3.5 Assume that the following conditions hold:

(i) There are $u_k,\ldots,u_{N-1}$ such that $h_i(x_i,u_i) < 0$ with $x_{i+1} = g_i(x_i,u_i)$, $i = k,\ldots,N-1$

(ii) $g_i(x_i,u_i) = V_i)x_i) + W_i(u_i) + c_i$,

where V_i's and W_i's are linear maps and c_i's are vectors

(iii) $f_k,\ldots,f_N$ and $h_k,\ldots,h_{N-1}$ are convex.

Then for every optimal solution (x,u) of P(k) there are λ and μ such that (x,u,λ,μ) is a safflepoint for L_k.

Proof. Consider the following problem

$$\begin{aligned} &\min f(y) \\ &\text{s.t. } My = b, \\ &\qquad h(y) \leq 0, \end{aligned}$$

where $y = (x_{k+1},\dots,x_N,u_k,u_{N-1})^T$,

$$h(y) = (h_k(x_k,u_k),\dots,h_{N-1}(x_{N-1},u_{N-1}))$$

$$f(y) = R_k(x_k,\dots,u_{N-1}),$$

$$b \quad = (c_k+V_k(x_k),c_{k+1},\dots,c_{N-1})^T,$$

(T denotes the transposition),

$$M = \begin{pmatrix} E_{k+1} & O & \dots & & O & -W_k & O & \dots & O \\ -V_{k+1} & E_{k+2} & \dots & & O & O & -W_{k+1} & \dots & O \\ & & \dots & & & & \dots & & \\ O & O & \dots & -V_{N-1} & E_N & O & O & \dots & -W_{N-1} \end{pmatrix}$$

(E_i denotes the unit matrix of demension of X_i).

Observe that the Lagrangian function associated with this problem is the same as L_k. Using this fact and taking conditions (i),(ii) and (iii) into account we are now able to apply the Lagrange multipler theorem of convex programming (see Luenberg (1969), p.217) to the problem and this makes the proof complete. #

4. CONJUGATE FUNCTIONS

In this section we maintain all the assumptions made in Proposition 3.5. Following the method of conjugate functions, first we define

$$\rho_k(x_k,x,u,w,v) = R_k(x_k,x,u)$$

if $$x_{i+1} = g_i(x_i,u_i) + v_i,$$

and $$h_i(x_i,u_i) \leq w_i, \quad i = k,\dots,N-1;$$

$$\rho_k(x_k,x,u,w,v) = +\infty \qquad \text{otherwise,}$$

and its conjugate function

$$\rho_k(x_k,t,s,\lambda,\mu) = \sup\{<t,x>+<s,u>+<w,\lambda>+<v,\mu> -$$

$$- \rho_k(x_k,x,u,w,v) : x,u,w,v\}.$$

The primal and dual problems will be

$$\min\{\rho_k(x_k,x,u,0,0) : x,u\} \tag{P}$$

$$\max\{-\rho_k^*(x_k,0,0,\lambda,\mu) : \lambda,\mu\}. \tag{D}$$

It is easy to see that (P) is the same as P(k), and in the case solutions for ρ_k^* exist, (D) is the same as D(k). We have similar results:

i) $-\rho_k^*(x_k,0,0,\lambda,\mu) \leq \rho_k(x_k,x,u.0,0)$ for each x,u,λ,μ

ii) If $-\rho_k^*(x_k,0,0,\lambda,\mu) = \rho_k(x_k,x,u,0,0)$

for some x,u,λ,μ, then (x,u) and (λ,μ) are optimal solutions of (P) and (D), respectively.

Now we define the primal and dual perturbation functions Φ_k, Ψ_k associated with ρ_k and ρ_k^* as follows:

(7)

$$\Phi_k(x_k,w,v) = \inf\{\rho_k(x_k,x,u,w,v):\ x,u\}$$
(7)
$$\Psi_k(x_k,t,s) = \inf\{\rho_k^*(x_k,t,s,\lambda,\mu):\ \lambda,\mu\}.$$

Proposition 4.1 Under the conditions of Proposition 3.5, Φ_k is a convex function of (w,v).

Proof. A direct verification will yield the proposition. #

Proposition 4.2 Assume additionally that (7) is solvable, then $\Phi_k,\dots,\Phi_{N-1}$ yield the following recurrence relations

$$\Phi_i(x_i,w,v) = \inf\{f_i(x_i,u_i) + \Phi_{i+1}(g_i(x_i,u_i)+v_i,w',v'):$$
$$x_i,u_i \text{ with } h_i(x_i,u_i) \le w_i\}$$

$$i = k,\dots,N-i \quad \text{and} \quad \Phi_N(x_N) = f_N(x_N).$$

Proof. As the proof of this proposition is similar to that of Proposition 3.3, we omit it. #

Remember that problem (P) is said to be stable if the subdifferential $\partial\Phi_k(x_k,0,0)$ is a nonempty set. Now we can apply Theorem 5.11 (Avriel (1976)) to get the following result:

Let $\Phi_k(x_k,0,0)$ be finite. Then problem (D) has an optimal solution $(\bar{\partial},\bar{\mu})$ and

$$\Phi_k(x_k,0,0) = \max\{-\rho_k(x_k,0,0,\lambda,\mu):\ \lambda,\partial\}$$
(8)
$$= -\rho_k^*(x_k,0,0,\bar{\lambda},\bar{\mu})$$

if and onyl if (P) is stable. Moreover, $(\bar{\lambda},\bar{\mu}) \in \partial\Phi_k(x_k,0,0)$

if and only if (8) holds.

The following results are immediate:

Corollary 4.1 Assume that (P) has an optimal solution (x,u). Then there exists (λ,μ) such that (x,u,λ,μ) is a saddlepoint for L_k if and only if (P) is stable. #

Corollary 4.2 Suppose that (7) is solvable. If $P(k)$ is stable, then so are $P(k+1),\dots,P(N)$ (these problems are determined by $x_{k+1},\dots,x_N$ where $(x_{k+1},\dots,x_N)$ together with some u is a solution of (7)). #

CONCLUSIONS

To conclude this paper we should emphasize that the results obtained are merely theoretical aspects of a duality approach for solving the dynamic problem described in Section 2. They provide us with a possibility of solving problem

$$\max\{d(x_i,\lambda,\mu):\ \lambda,\mu \text{ without constraints}\}$$

instead of

$$\min\{R_1(x_i,x,u):\ \text{under constraints } (2),(3),(4)\},$$

where $d(x_i,\lambda,\mu)$ may be calculated by recurrence relations given in Corollary 3.1.

REFERENCES

Avriel, M., 1976, Nonlinear Programming (London: Prentice-Hall, INC.). Bellman, R., 1957, Dynamic Programming (New Yersey: Princeton University Press).

Evers, J.J.M., 1983, A Duality Theory for Infinite-horizon Optimization of Concave Input/Output Processes, Mathematics of Operations Research 8, 479-497.

Hadley, G., 1962, Linear Programming (London: Addision-Wesley).

Haneveld, W.K.K., 1985, Duality in Stochastic linear and dynamic programming (Amsterdam: Centrum voor Wiskunde en Informatica),

Luenberg, D.G., 1969, Optimization by Vector Space Methods (New York: John Wiley and Sons).

Weitzman, M.L., 1973, Duality Theory for Infinite-horizon Convex Models, Management Science 19, 783-789.

Dualitás a dinamikus programozásban

Dinh The Luc

Összefoglaló

A szerző véges-állapotu /"finite stage"/ dinamikus rendszerekre vonatkozó dualitás-elméletet dolgoz ki, felhasználva a matematikai programozásban használatos Lagrange és konjugált függvény módszert. Néhány dualitás tételt és nyereg-pont tételt is nyer, valamint rekurrencia-összefüggéseket, amelyek a rendszer dinamikus tulajdonságait jellemzik.

Двойственность в динамическом программировании

Дхин Тхе Лук

Р е з ю м е

В статье разработана теория двойственности для динамических систем конечного состояния /"finite-stage"/, используя метод двойственных функций Лагранжа в математическом программировании. Доказано несколько теорем двойственности и теорем о седловой точке, а также несколько рекуррентных соотношений характеризующих динамические свойства системы.