Cooperative Dynamic Programming*

James Hess

Department of Economics
University of Southern California
University Park
Los Angeles, California 90007

Harriet Kagiwada and Robert Kalaba

HFS Associates
3117 Malcolm Avenue
Los Angeles, California 90034
and
University of Southern California

and

Chris Tsokos

Department of Mathematics
University of South Florida

ABSTRACT

A many-decision-maker dynamic-control problem under uncertainty is presented as an extension of the standard control problem. It is assumed that only one payoff function exists, identical for all decision makers. The outstanding problem is to find a sequence of several decision rules that optimally coordinate activities both across time and between decision makers. This cooperative dynamic programming problem requires the researcher to consider how to solve large numbers of systems of integral equations.

*Research sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant No. 77-3383.

 0096/3003/79/010069 + 6$01.75

:NTRODUCTION

he optimal control of dynamical systems requires that a decision maker dinate the actions he takes in several succeeding time periods. The ulus of variations, Pontryagin's theory of optimal processes and Bellman's amic programming provide alternative ways of describing the optimal ience of actions. Uncertainty and learning have been incorporated in the lels of stochastic and adaptive control. Only recently has the theory of mal control considered the problem of cooperative control of a system by roup of decision makers [1–5]. Cooperative decision making in a single iod model has been studied in the theory of teams [6] and the theory of entralization [7].

n this paper, a simple model of cooperative dynamic control will be sented in order to focus attention on a few of the outstanding analytic l computational problems. No attempt is made to produce the most ieral formulation, because that has been done elsewhere [2] and because iple examples are sufficient to indicate the degree of difficulty of the ny decision maker problem.

The mathematical problem discussed here might be referred to as a blem of dynamical team decision making, since there is a harmonious oup of decision makers with differing information about an uncertain and namic environment. However, there are already several conflicting defini- ons of what is meant by dynamic teams (Refs. 3 and 6), so we have used the rm "cooperative dynamic programming" to describe the models that are vestigated here. Further qualifications of the topic include the assumptions discrete, finite time and of perfect knowledge of all probability distribu- ons of the random variables.

. DYNAMIC, DECENTRALIZED PROFIT MAXIMIZATION

To illustrate the concepts of cooperative dynamic programming, imagine business firm with two decision makers, a president and a foreman. The resident is responsible for all capital expenditures, and the foreman must lecide on the amount of labor that is hired. If, in any period, the firm has K inits of capital and L units of labor services, the profit of the firm is

$$pf(K,L)-wL-rK, \tag{1}$$

where p is the competitive price of output, $f(K,L)$ is a production function relating inputs to outputs, w is the wage rate, and r is the interest rate on capital. The firm hires labor and buys capital at the beginning of each of T time periods, which are indexed by $t=1,2,\ldots,T$. The factor markets are

assumed to be competitive, so that the firm has no influence on w or r. However, random fluctuations in market supply and demand make the wage and interest uncertain from the viewpoint of the individual firm. The firm, however, knows that the frequency of any particular (w,r) pair is described by a probability density $g(w,r)$ (which for simplicity will be assumed to be independent of time).

The firm's decision makers specialize in observing the market conditions for the factors they must buy. To simplify the analysis, it will be assumed that this specialization implies that the president learns the market interest rate r_t before he makes his decision to buy more capital goods in period t, and that the foreman learns the market wage rate w_t before hiring labor in period t. There is no communication between president and foreman concerning their specialized knowledge. This assumption can be relaxed, but is meant to capture the realistic fact that decision makers seldom have identical knowledge about the environment of the firm.

It is assumed that the president and foreman make their decisions independently (authority is decentralized). The president must establish a rule for relating his knowledge of the interest rate to the changes in the capital stock. This rule is called the *investment policy* and is denoted

$$I_t=I_t(r_t). \tag{2}$$

The foreman must establish a rule for how he will relate his knowledge of the wage rate to the size of the labor force. This rule is called the *hiring policy* and is denoted

$$L_t=L_t(w_t). \tag{3}$$

The objective of this firm is to maximize the total expected profit over the T period time horizon by the selection of a sequence of investment and hiring policies. That is, the firm wishes to

$$\begin{aligned}\underset{(I_t(\cdot),L_t(\cdot))_{t=1}^{T}}{\text{maximize}}\ \sum_{t=1}^{T}\iint\big[\,&pf(K_t,L_t(w_t))-w_tL_t(w_t)-r_tK_t\\&-C(K_t-K_{t-1})\big]\,g(w_t,r_t)\,dw_t\,dr_t,\end{aligned} \tag{4}$$

(where C is the cost of capital adjustment) subject to

$$K_{t+1}=K_t+I_{t+1}(r_{t+1}), \tag{5}$$

$$K_0 \text{ given.} \tag{6}$$

functional in (4) gives the total expected profit. Equation (5) is the ital accumulation equation for the firm, and Eq. (6) specifies the firm's al capital stock. It should be remarked that although both authority and rmation are decentralized, both decision makers are attempting to maxi- e the same objective function: they constitute a *team*. There are no game oretic aspects to the firm's problem. In addition, cooperation between the decision makers is imperative because the objective function is not itively separable in K and L; the marginal productivity of capital (labor) nfluenced by the quantity of labor (capital) used.

PRINCIPLE OF OPTIMALITY

The cooperative dynamic programming problem of the relations (4)–(6) y be expressed in recursive form by applying Bellman's principle of imality [8, 9]. No matter what the initial capital stock and initial invest- nt and hiring policy, the remaining policies must constitute optimal licies with regard to the capital resulting from the first decision. Define a nction $\phi_t(K)$ to be the maximum total expected profit for the time periods $t+1,\ldots,T$ if the capital stock inherited from the previous period is K. is function $\phi_t(K)$ must satisfy the recurrence relation

$$(K) \equiv \max_{L_t(w,K) I_t(r,K)} \iint [\, pf(K+I_t(r,K), L_t(w,K)) - wL_t(w,K) - C(I_t(r,K)) - r\cdot(K+I_t(r,K)) + \phi_{t+1}(K+I_t(r,K))] \times g(w,r)\, dw\, dr, \quad (7)$$

$= 1,\ldots, T-1$, and

$$(K) \equiv \max_{L_T(w,K), I_T(r,K)} \iint [\, pf(K+I_T(r,K), L_T(w,K)) - wL_T(w,K) - C(I_T(r,K)) - r\cdot(K+I_T(r,K))]\, g(w,r)\, dw\, dr. \quad (8)$$

quation (8) expresses the fact that when only one time period remains in ne process, maximum total expected profit with inherited capital is achieved y maximizing $E\{pf(K+I_T, L_T) - wL_T - r(K+I_T)\}$ with respect to the olicies I_t, L_t. When we are at time t in the process, Eq. (7) requires hat L_t, I_t maximize the sum of expected profit earned during that time eriod and the maximum expected profit for the remainder of the process eginning with the capital stock $K+I_t$.

4. ANALYTIC AND COMPUTATIONAL ASPECTS

What difficulties are there in evaluating the optimal sequence of investment and hiring policies? At any stage in the process the recurrence relation (7) must be solved for the optimal functions $L_t(w,K)$ and $I_t(r,K)$. For each possible value of K these optimal policies must satisfy the following person by person optimality conditions (see Ref. [6]):

$$0 = \int [\, pf_1(K+I_t(r,K), L_t(w,K)) - r + \phi'_{t+1}(K+I_t(r,K)) - C^1(I_t(r,K))]\, g(w|r)\, dw \quad \text{for all } r, \quad (9)$$

$$0 = \int [\, pf_2(K+I_t(r,K), L_t(w,K)) - w]\, g(r|w)\, dr \quad \text{for all } w, \quad (10)$$

where ϕ'_{t+1} is the derivative of ϕ_{t+1}, $g(w|r)$ and $g(r|w)$ are posterior probabilities, and f_i is the partial derivative of f with respect to its ith argument. Equations (9) and (10) are a coupled system of nonlinear integral equations. This presents a severe analytic and computational problem. At each stage, not one but many systems of nonlinear integral equations must be solved (as K ranges from zero to infinity). Techniques of parameter imbedding in K (see Ref. [10]) may simultaneously solve the optimality conditions (9) and (10) for every K, but this may involve large amounts of computer time.

In addition, once the optimal decision rules are found, the maximum expected profit must be computed by evaluating the multiple integrals defined in the relations (7) and (8). This could involve large errors or large computational costs or both.

The problem developed above is the simplest possible example. Further difficulties could arise if the information of the decision makers consists of several variables instead of just one. In this case the integral equations (9) and (10) would involve multiple integrals. If investment or labor were constrained in any way (say, to be non-negative) then the optimization problem of (7) and (8) would be considerably more difficult. Finally, if the probability densities are not known, then adaptation to accumulating information would have to be modeled in some fashion.

While the difficulties of cooperative dynamic programming are indeed large, the situation is not hopeless. If the production function of the firm is linear-quadratic,

$$f(K,L) = a_0 + a_1K + a_2L + a_{11}K^2 + a_{12}KL + a_{22}L^2, \quad (11)$$

and the random variables are distributed Gaussian-normal, then it is well known [6] that the decision rules with one stage to go are linear in the information variables and K. It can then be shown that the decision rules at any stage are linear. This implies that instead of solving integral equations, the cooperative dynamic programming problem involves solving linear algebraic equations.

Even if the linear-quadratic and Gaussian assumptions fail to hold, it may be permissible to restrict the decision rules to a function space which is linear in a finite number of parameters (as is done in standard regression analysis). This restriction will also result in linear algebraic equations rather than the more complicated integral equations.

5. DISCUSSION

The many decision maker dynamic control problem has been presented as an extension of the standard control problem. The assumption of only one payoff function, identical for all decision makers, removes many game theoretic difficulties. The outstanding problem is to find a sequence of several decision rules that optimally coordinate activities both across time and between agents. The multi-agent nature of cooperative dynamic programming requires the researcher to consider how to solve large numbers of systems of integral equations.

REFERENCES

1 P. P. Varaiya, Trends in the theory of decision-making in large systems, *Annals of Economic and Social Measurement* 1 (4):493–500 (Oct. 1972).
2 H. S. Witsenhausen, Separation of estimation and control for discrete time systems, *Proc. IEEE* 59 (11):1557–1566 (Nov. 1971).
3 Y. C. Ho and K. C. Chu, Team decision theory and information structures in optimal control problems—Parts I, II, *IEEE Transactions on Automatic Control* AC-17 (1):15–28 (Feb. 1972).
4 M. Aoki, On decentralized linear stochastic control problems with quadratic cost, *IEEE Transactions on Automatic Control* AC-18 (3):243–250 (June 1973).
5 M. Athans, Theory and applications survey of decentralized control methods, *Annals of Economic and Social Measurement* 4 (2):345–355 (1975).
6 J. Marschak and R. Radner, *Economic Theory of Teams*, Yale U. P., New Haven, Conn., 1972.
7 L. Hurwicz, The design of mechanisms for resource allocation, *American Economic Review* 63:1–30 (May 1973).
8 R. Bellman and R. Kalaba, *Dynamic Programming and Modern Control Theory*, Academic, New York, 1966.
9 R. Bellman, *Dynamic Programming*, Princeton U. P., Princeton, N. J., 1957.
10 H. Kagiwada and R. Kalaba, *Integral Equations via Imbedding Methods*, Addison-Wesley, Reading, Mass., 1974.

Rational Approximations of Trigonometric Matrices with Application to Second-Order Systems of Differential Equations*

Steven M. Serbin

Department of Mathematics
University of Tennessee
Knoxville, Tennessee 37916

Transmitted by Robert Kalaba

ABSTRACT

We consider the direct treatment of the second-order system of equations $y''(t)+Ay(t)=f(t)$, such as might arise in finite-element or finite-difference semidiscretizations of the wave equation. We develop the exact solution and some three-term recurrences involving trigonometric matrices. We approximate these trigonometric matrices by rational approximants of Padé type and thus develop a two-parameter family of approximation schemes. We analyze the stability behavior and computational complexity of members of this family and isolate four schemes for numerical experimentation, the results of which we tabulate. We single out as particularly effective the classical Stormer-Numerov method and also a new sixth-order scheme.

1. INTRODUCTION

There is a vast amount of literature dealing with the application of rational approximation of the exponential function to the development of approximation schemes for the solution of the initial-value problem

$$y'(t)=Ay(t)+q(t),$$
$$y(t_0)=y_0, \qquad (1.1)$$

*This research was supported in part by National Science Foundation Grant #MCS 76 24453.