Multiperiod Optimization:
Dynamic Programming vs. Optimal Control
(Harry Mapp, Oklahoma State University, Presiding)

Dynamic Programming: Has Its Day Arrived?

Oscar R. Burt

A look at John Kennedy's recent review article on dynamic programming (DP) applications in agriculture, forestry, and fisheries leaves no doubt that DP is a useful analytical and numerical technique [Kennedy]. However, the method appears to be in more common use among Australian than American agricultural economists, particularly in the farm management/production economics area. When John Allison and I published "Farm Management Decisions with Dynamic Programming" [Burt and Allison] twenty years ago, I though DP would be as routinely used by now as linear programming (LP), but this is obviously not the case. After outlining the fundamental principles of DP, the apparent lack of its popularity among research agricultural economists will be addressed.

For simplicity in exposition, consider a dynamic process which can be described by a single state variable and a first order difference equation,

$$x_{t+1} = h(u_t, x_t, \varepsilon_t) \tag{1}$$

where u, x, and ε are the decision variable, state variable, and a random variable, respectively. Frequently, the economic criterion is a periodic cash flow, the expectation of which can be written as a function of u_t and x_t, i.e., $G(u_t, x_t)$. A one-period discount factor appropriately defined for the length of time period t is denoted β. Since the conditional expected periodic return function, $G(u_t, x_t)$, is a random variable for periods $j>t$ from the decision agent's point of view at time t, a decision function $D_t(x_t)$ is sought which maximizes the expected present value criterion

$$E\left[\sum_{t=1}^{\infty} \beta^t G(D_t(x_t), x_t)\right] \tag{2}$$

subject to (1) and an initial value x_o, where $E(\cdot)$ is the mathematical expectation operator.

The dynamic optimization model can be restated as a recursive equation by defining $v_n(x)$ as the expected present value of net returns from an n-period decision process when the *optimal* decision rule is followed and the initial state of the process is x. Then application of Bellman's "principle of optimality" [Bellman 1957a] gives

$$v_n(x) = \max_u [G(u,x) + \beta E\{v_{n-1}(h(u,x,\varepsilon))\}]. \tag{3}$$

The subscript n is a reverse ordering with respect to chronological time t, and is convenient for considering limits as the process is extended indefinitely. The number of periods duration of the process is called the number of stages; thus (3) refers to an n-stage process.

When $G(\cdot)$, $h(\cdot)$, and the distribution of ε are independent of stages, $v_n(\cdot)$ converges to a limiting function as $n \to \infty$ and the decision rule is invariant over stages. Therefore, the limiting case of (3) is a functional equation where $v(\cdot)$ replaces $v_n(\cdot)$ and $v_{n-1}(\cdot)$. Many applications involve a finite number of periods and the functions are changing over time (stages), but the simpler case is used here for discussion purposes. The basic recurrence equation in (3) generalizes in an obvious way to allow, u,x, and ε to be vectors with $h(\cdot)$ a vector function. The nonstochastic case occurs when the random variable ε is

Oscar R. Burt is a professor in the Department of Agricultural Economics and Economics at Montana State University.

deleted from $h(\cdot)$, i.e., the state variable changes from stage to stage as an exact function instead of randomly.

The most common method of solving (3) is by successive approximations to $v_{n-1}(\cdot)$. An arbitrary value is assigned to $v_o(\cdot)$, say identically zero; then (3) is solved for $v_1(\cdot)$ by solving the optimization problem on the right hand side for many discrete values of the state variable, then $v_2(\cdot)$ is obtained with the approximation for $v_1(\cdot)$, then $v_3(\cdot)$ and so on. Using discrete values of a continuous state variable as an approximation converts (3) into a finite Markovian decision process which was the type of DP problem analyzed in Ronald Howard's pioneering work [Howard].

Obstacles to Implementation

Although the computational and input/output problems with data are substantial in getting solutions to empirical formulations of (3), the greatest obstacle to applications seems to be conceptualization difficulties, i.e., understanding how to formulate an empirical situation as a DP model. The conceptual and computational difficulties compound one another because the most intuitive and direct way to structure the model is often infeasible computationally, or at least cumbersom and expensive in computer resources. Clifford Hildreth remarked 25 years ago in the context of this same type of decision theory model: "...although theoretical discussion is concerned almost exclusively with solving given decision problems, formulating appropriate problems is one of the most difficult tasks in successful application" [Hildreth].

It is important to recognize that the solution technique is relatively unimportant compared to formulation of the practical problem in such a way that it is amenable to analysis by the particular technique chosen. An outstanding example of imaginative use of stochastic simulation in conjunction with statistical response surface exploration methods is the early application of Zusman and Amiad to a stochastic DP problem which appeared too complex for a direct DP solution. The recent analysis of U.S. wheat storage policies by Taylor and Talpaz is another clever use of two techniques together — approximate certainty equivalence DP and stochastic simulation.

The primary objective in all modeling is to capture the essential aspects of the phenomenon under study and yet keep the model as simple as possible. The tendency for stochastic sequential decision problems to mushroom in complexity requires more imagination in the application of this basic principle of modeling. Modern computers and cheap rates at many research facilities have encouraged the substitution of computer simulation in a cavalier manner for well constructed economic models. The result has been more numerical output than can be assimilated by the analyst, let alone presented in an objective form for the professional literature. Many papers are being published which try to subjectively summarize voluminous output from simulation models which are a complete black box to the prospective audience.

Although I do not understand why, there seems to be more difficulty than with most other mathematical modeling techniques in making the transition from a rudimentary knowledge of the method of DP to original formulations of problems. Dreyfus and Law begin the preface of their text:

> It is our conviction, based on considerable experience teaching the subject, that the art of formulating and solving problems using dynamic programming can be learned only through active participation by the student. No amount of passive listening to lectures or of reading text material prepares the student to formulate and solve novel problems. The student must first discover, by experience, that proper formulation is not quite as trivial as it appears when reading a textbook solution. Then, by considerable practice with solving problems *on his own*, he will acquire the feel for the subject that ultimately renders proper formulation easy and natural.

The above statement makes an excellent premise on which to teach any technical course such as econometrics or operations research techniques, but in a relative sense, it is especially important in DP. My personal experi-

ences in teaching the subject support the view of Dreyfus and Law.

In my opinion, part of the lack of popularity of DP in agricultural economics research emanates from its cursory treatment in graduate courses which survey various operations research techniques. The many topics which compete for time in such courses, combined with the limited credit hours devoted to quantitative methods, discourage instructors from covering the subject in much depth. Another factor is the generally poor background in mathematics and statistics of our graduate students which discourages them in their study of abstract topics such as DP, particularly the stochastic models. A final obstacle to the students learning the subject is the poor quantitative training of the instructors themselves in many of these courses. When these quantitative courses are offered in specialized departments on campus with well qualified instruction, that is where the students should be enrolled. Technical competence in our profession is still suffering from the "relevancy binge" taken back in the late 1960's and early 70's.

The use of control theory in its popular deterministic, continuous time version, associated with the Pontryagin maximum principle, has received a great deal of exposure in the general economics literature during the past decade. However, numerical solution of applications is a rarity because these papers have been almost entirely of a theoretical nature where general mathematical structure of the solution is the objective. Explicit solution of control theory problems is seldom possible except for trivial exercises. Most agricultural economists, particularly the younger ones, have had some exposure to this type of dynamic economic analysis. Control theory applications in the agricultural economics literature tend to mimic the theoretical research in general economics by stating the necessary conditions for an optimum, á la the Pontryagin maximum principle, and then give an intuitive economic interpretation, but theorems on the structure of the solution are rare.

The discrete time characteristic of DP is actually more realistic in most economic applications than continuous time modeling. One may wonder why the discrete time analogues of Pontryagin's maximum principle are not more widely used in general economics. The reason is apparently the more awkward mathematical analysis and less opportunity to exploit results already proved in the physical and mathematical sciences. The cost of using continuous approximations to discrete time phenomena is that certain tacit assumptions are made about the optimal solution to the actual discrete time model. For example, certain types of oscillatory behavior of the state variables are ruled out in the continuous time version which could affect the solution. In a sense, the necessity of making tacit assumptions to apply continuous time models to discrete phenomena is the dual of the assumption required to derive the Pontryagin maximum principle from Bellman's principle of optimality. Mathematically, the relaxed assumptions needed to prove the maximum principle was the main contribution of Pontryagin over Bellman's results (see Intrilligator p. 329, 356, and footnote 5 for the nature of the assumption required to use Bellman's approach).

When the dynamic model is nonstochastic and the objective of the analysis is to set out the necessary conditions for an optimal solution with a little intuitive economic interpretation, it is largely irrelevant whether we think of the necessary conditions as having been derived through the calculus of variations, control theory, or DP; the result is the same. If further mathematical analysis to delineate the structure of the solution is attempted, it appears that mathematical economists usually prefer classic variational methods over DP in deterministic models.

Bellman introduced the term "curse of dimensionality" to describe the extremely large computer storage requirements resulting when a direct computational procedure is applied to the recurrence relation of (3). Unfortunately, the term has tended to be associated with DP and not recognized as an

inherent characteristic of dynamic optimization problems in general when the number of state variables is large (more than 3 or 4). Many methods which are used with the idea of avoiding the curse, such as simulation, tacitly pretend that it does not exist. The result of ignoring the true complexity of the optimization problem is a partial, approximate, or at worst, no solution to the problem.

Most solution methods associated with the control theory or calculus of variations formulation reduce the dimensionality problem by obtaining the solution to a very specific statement of the optimization, *viz.*, a specific initial state of the process, instead of getting the solution to an entire family of problems associated with all possible initial states as is obtained with DP. The latter method also yields the entire family of solutions for all possible lengths of planning horizons in the context of many applications. In empirical agricultural economics research, the objective is to learn much more about a dynamic process than merely *an* optimal solution path starting from some initial state, in contrast to many applications such as in the space program or individual firm investment decisions.

DP has a certain generality and simplicity for numerical solution of low dimension problems in that concavity/convexity assumptions are not required and inequality and integer constraints actually simplify the computations. But good numerical algorithms have been developed in recent years for some classes of continuous state and decision variable models, and these algorithms provide powerful alternatives when appropriate for the problem at hand. When one can obtain the optimal time path of the state vector for an initial state with one of these algorithms at a relatively low cost, the structure of the optimal decision rule could be approximated by using the principles of experimental design to strategically choose a relatively small number of initial states. The data thus generated would provide an empirical basis for fitting a continuous vector function of the state variables as an approximation to the decision rule, i.e. $u_t = D_t(x_t)$, where u and x are vectors.

The above discussion has been with reference to deterministic models; when the state variable difference (differential) equation is stochastic as in (1), DP appears to be the only viable method for numerical solution. The stochastic version of the maximum principle does not appear to be very useful in suggesting solution algorithms [Kushner and Schweppe]. The method used recently by Dixon and Howitt (attributed to Athans) has serious logical problems in that a deterministic solution path is used as a reference locus for minimizing expected squared deviations. Various approximation methods applied directly to (3) would appear to give better results; some of these are discussed below. The basic difficulty in stochastic control models is that the decision must be made after the random outcomes of the state variables are known for periods $j \leq t$ when the decision is made at period t, sometimes referred to as a "wait and see" solution, otherwise valuable information is thrown away. In the parlance of control theory, the decision rule needs to be in "closed loop" form where the decision variable is expressed as a function of the state variables. This is the natural way in which the decision rule is obtained from a DP solution procedure, see (1) through (3).

Making DP More Operational

The primary consideration in formulating problems in a DP model is the choice of state variables which must jointly describe the entire history of the process. The first order difference equation in (1) requires that the state and decision vectors on the right hand side summarize all pertinent information contributing to optimization of the dynamic process under study. Deficiencies in state variables can directly discard information since a decision rule makes the decision variables functions of the state variables at each stage, or the loss of information can be indirect by causing $h(\cdot)$ in (1) to be inaccurate and create errors in the time paths of the included state variables. Frequently, informa-

tion is lost in both ways when a relevant state variable is omitted from the model.

The expected present value function, $v_n(x)$, ultimately determines the importance of a state variable; x is now defined as a vector. If $\partial v_n(\cdot)/\partial x_j = 0$, then x_j is a redundant state variable which could be deleted. Going from this extreme case, the importance of x_j in the decision process is closely related to the magnitude of $\partial v_n(\cdot)/\partial x_j$, particularly in a relative sense to the other state variables. We can avoid the scale of measurement problem by converting $\partial v_n(\cdot)/\partial x_j$ to an elasticity, denoted η_j. Then $|\eta_j|$ relative to $|\eta_k|$ tells us something about the importance of x_j compared to x_k.

Another important consideration is the amount of independent variation in x_j taken jointly with the other state variables. The ideas here are much like those in multivariate statistical models, factor analysis in particular. If all the independent variation of the $\{x_j\}$ takes place in $q<m$ dimensions and x is an m-component vector, then there exists a transformation from m to q dimensions such that the new state variables $y_1, \ldots y_q$ describe the process equally well. The meaning of independence is linear and stochastic for deterministic and stochastic processes, respectively, at least in an operational sense. Most likely, the transformation would be linear in either case which implies the normal distribution for a stochastic process, possibly after a previous transformation of the state variables such as logarithmic.

The above considerations about the state variables would not be very useful if economic modeling were an exact science. Approximations are always necessary and these ideas can be used to choose among various ways of simplifying the model. This is particularly true in conjunction with a method of deriving proximate decision rules given by Arrow for DP formulations. Observe that (3) would be a static optimization problem if the function $v_{n-1}(x)$ were known. The basic idea of Arrow's is to use an approximation to $v_{n-1}(\cdot)$ which provides a proximate decision rule when the one-stage maximization problem in (3) is solved.

Reductions in the State Vector Dimension

There is no set way to arrive at an approximation to $v_{n-1}(x)$. In some types of applications, such as firm growth models, dynamic linear programming coupled with the standard parametric programming options of the algorithms could be used to estimate a present value function starting from various states. Then $v_{n-1}(x)$ could be approximated by fitting a continuous function, such as a polynomial, to the points generated by parametric linear programming. The form of (3) from which the final proximate decision rule is derived could be either deterministic or stochastic.

In a recent wheat storage study [Burt, Koo, and Dudley], a simplified model containing only two state variables was used to estimate $v_{n-1}(x)$ with the other state variables suppressed. Then the special nature of the additional state variables, *viz.*, higher order lags in econometric equations, permitted their gradual phasing into the expected present value function over six additional stages. Each additional iteration contained an increment to the number of state variables and provided an improved approximation to $v_{n-1}(x)$ in (3). The model was stochastic and the final iteration implicitly contained 14 state variables. This example illustrates the importance of exploiting any special structure in the application.

The method used in [Burt, *et al.*] to suppress all of the state variables except two was a very informal application of the principles discussed above on independent variation among the state variables which were economic time series varibles and highly collinear. Subjective estimates of the relative magnitudes of the $\{\eta_j\}$ were also used in deciding which state variables to retain for the two state variable model. Another consideration in this application was the role of a commodity market equilibrium where the lagged variables in the econometric equations converge to equilibrium values; the two state variable model could be interpreted as an "equilibrium neighborhood" model.

Robert Taylor just completed a study of wild oats control in the Northern Great Plains which solved a stochastic DP model containing five state variables. The proximate solution procedure of Arrow was used by first solving a three state variable problem which kept the most important variables based on a priori reasoning. Then the two remaining state variables were introduced in sequence with two more iterations. The model contained four stochastic and one deterministic state variables. Here too, the key to successful solution with relative ease on a modest size computer was exploitation of the special structure of the application.

Taylor's approach to simplifying the problem was an informal, subjective use of relative elasticities, $|\eta_j|$; two state variables were suppressed in getting the estimate of $v_{n-1}(\cdot)$. Some limited validation of the a priori assumption about relative elasticities can be performed after the final model when results for all five state variables have been estimated, but care must be taken in recognizing the influence of the approximating procedure itself on empirical estimates of the $\{\eta_j\}$. Also, the size of $|\eta_j|$ must be considered jointly with the amount of variation in x_j during the process; if x_j is constant over the process it is a redundant variable.

Giaver's early study of dairy cow replacement was a formal application of multivariate normal theory to reduce the number of state variables by a linear transformation to lower dimension space [Giaver]. An important set of state variables in dairy cow replacement decisions is lactation history; if a cow is kept for as many as six lactations, then six state variables are involved in describing her production history. Giaver assumed a multivariate normal distribution on production by lactation and used a single state variable to replace the entire history. This variable was a linear combination of individual lactations obtained from the linear regression coefficients of a predictor equation for the next lactation beyond the current history. Since the conditional distribution of a normal variate, given a linear constraint on a subset of variates with which it is jointly distributed, is also normally distributed, the distribution for stochastic state transitions was easily derived from the multivariate distribution over all lactations. The particular linear combination chosen had the intuitive appeal of being the conditional mean of the lactation for the next stage of the decision process, given the cow's lactation history.

Although Giaver's method has a certain amount of intuitive appeal, a better choice for a single linear combination of the state variables would be the eigenvector associated with the largest principal component of the estimated covariance matrix. Point estimation is not the objective, but instead, a linear relation which captures the most information with respect to the conditional distribution, given the linear restriction on the subset of multivariate normal variates. In the normal distribution, the most information is synonomous with minimum variance in the conditionally distributed random variable. The principal component method also generalizes directly to two or more linear combinations of the state variables to get a smaller set of state variables which captures the most information. In distributions which are not multivariate normal, the principal component method, which minimizes conditional variance, would seem to be a good approximation to an optimal transformation, although it cannot be justified as well in a formal statistical model.

The above discussion of transforming multivariate normal state variables into a smaller set, and thus throwing away information, illustrates a general method of reducing the dimension of the state vector in stochastic DP which is much easier to defend philosophically than its counterpart in deterministic models. A subset of the original state variables, or a reduced set obtained by transformations, is used as the state variables in the DP model and the probability distribution associated with the state transition from stage to stage has greater variance. The reduced model uses a conditional distribution based on the smaller set of state variables

which connotes less information in the decision process. Since we are dealing with probability distributions, there is no "error" or "bias" in the modeling process, just less information relative to the ideal.

In deterministic models, we can only speak of approximations and orders of magnitude on the errors created by a reduction in the dimension of the state vector. In the rare instance that an exact linear combination exists among the state variables, the dimension can be reduced without error, but any approximate linear combination will introduce some error if used to reduce the number of state variables. As a practical matter, the situation is much like that in stochastic models, but nothing as precise as a less informative conditional probability distribution can be used.

Taylor's Series Approximations to $v_{n-1}(\cdot)$

When continuous state variables are approximated by discrete segments and the transitions for that variable are deterministic, some form of interpolation between the discrete points is necessary. Usually linear interpolation is adequate if the intervals are not too wide. Let the random variable ε be suppressed in (3) and suppose the one continuous state variable x is approximated by a set of discrete values, $x^1, x^2, \ldots x^p$. Then the function $v_{n-1}(\cdot)$ is approximated by a table of values $v_{n-1}(x^i)$, $i=1, 2, \ldots p$. In general $h(u,x)$ in the right hand side of (3) will fall between two values, say $h(u,x^j)$ and $h(u,x^{j+1})$, which requires some sort of interpolation scheme between the two table values. A method of interpolation for deterministic state variables which comprise a subset of the total in a stochastic model is also necessary in most cases.

Bellman has suggested that the curse of dimensionality can be partially overcome by approximating $v_n(x)$ in (3) as a polynomial on the relevant interval [Bellman 1961, p. 244]. It is easily shown that the table of values used to approximate $v_n(x)$ is replaced by the coefficients of a polynomial. It might take 50 discrete values with linear interpolation to give the same accuracy as a 5th degree polynomial with only 6 parameters. The method readily extends to several state variables and some orthogonal form of polynomial such as Legendre polynomials could be used. These ideas could be used in fitting spline functions instead of polynomials if empirical evidence suggested a better approximation with the same number of parameters.

The polynomial approximation of $v_n(x)$ would appear relatively more powerful in stochastic DP where the state variables (or a subset) are continuous, although I have not seen it suggested. For a continuous single state variable, let $h(u,x,\varepsilon)=x+\phi(u,x,\varepsilon)$ without any loss in generality. Then $Ev_{n-1}(h(u,x,\varepsilon))$ is expanded in a Taylor's series around the point x,

$$(4) \qquad Ev_{n-1}(x+\phi(u,w,\varepsilon)) =$$

$$v_{n-1}(x)+E\phi(u,x,\varepsilon)v'_{n-1}(x)$$

$$+E\phi(u,x,\varepsilon)^2v''_{n-1}(x)/2!+$$

$$E\phi(u,x,\varepsilon)^3v'''_{n-1}(x)/3!+\ldots$$

If a polynomial approximation is used on $v_{n-1}(\cdot)$, the derivatives evaluated at x are easily calculated and only the moment functions $E\phi(u,x,\varepsilon)^j$ are required to an arbitrary order of approximation desired for the maximization operation in (3). The main attraction of this method in the stochastic model is avoidance of discrete approximations by means of a finite state Markov chain model [Howard]; only a few moments of $\phi(u,x,\varepsilon)$ are required which contain u and x as arguments. But these moment functions are calculated once and for all, and are just part of the data for the model.

The use of (4) in the right hand side of (3) does not depend in any essential way on $v_{n-1}(\cdot)$ being globally approximated by a polynomial. The method is operational if the first few derivatives of $v_{n-1}(\cdot)$ can be calculated at an arbitrary point x. This could be done by an interpolation formula using a

sufficient number of points and degree polynomial applied to a tabular approximation of $v_{n-1}(\cdot)$ as is traditional in deterministic DP. The interpolation would get rather complex with several state variables which would make the global polynomial approximation of $v_{n-1}(\cdot)$ look relatively more attractive. The transition from deterministic to stochastic DP might not be as onerous as most of us have thought if this method were as powerful as it appears.

An analytical approach to getting approximately optimal decision rules for certain types of DP applications has been quite successful [Burt 1964a, 1976, 1981; Burt and Cummings]. The ideas are similar to those just discussed in relation to substituting (4) into the right hand side of (3) except that approximations to the derivatives are also derived. In nonstochastic models the approximations become exact in a neighborhood of state space centered on the equilibrium state of the process. For decision processes where the state variables change slowly over time in conjunction with discounting, the approximation seems to give good results at any interior point in state space. In applications to ground water [Burt 1964a] and soil conservation [Burt 1981], the approximations were excellent except when the state variables were close to a nonnegativity boundary, which is not surprising in that the derivation of the approximations assumed interior solutions in the optimization.

The method can be used in either deterministic or stochastic applications. Application of the approximation requires solution of a system of equations for a given state of the process; a subset of the variables in the solution constitutes the approximately optimal values of the decision variables for the given state. As an example, the one decision and state variable model requires two and four simultaneous equations for the first and second order approximations, respectively, although one of the variables can be simply eliminated in the first order case. In the first order case the variables are the decision variable, u, and $v'(x)$, while in the second order approximation they are u, $v'(x)$, $v''(x)$, and $\partial u/\partial x$. These additional variables besides the decision variable are themselves of economic interest; $v'(x)$ is marginal expected present value of the state variable and $v''(x)$ is its rate of change. The $\partial u/\partial x$ is a measure of the sensitivity of the decision variable to changes in the state variable.

The most general treatment of this method is found in [Burt 1976] where the model is stochastic with multiple decision and state variables. Most of the analysis is for a first order approximation, but the general second order case with one decision and state variable is analyzed in the appendix. Although the paper is oriented around multiaquifer ground water management, the mathematical results are in general terms.

When the assumptions are met for certainty equivalence DP, [Simon] and [Theil], the second order approximation simplifies to the standard results; this was demonstrated in [Burt 1967] with a certainty equivalance model for intertemporal allocation of ground water. The second order approximation uses a locally quadratic approximation to the functional equation $v(x)$, and when the assumptions are met for the standard certainty equivalance theorem, $v(x)$ is apparently globally quadratic.

Algorithms for Discrete DP Markov Processes

For purposes of numerical solution, the recursive DP equation in (3) is usually approximated as a finite state Markov decision process,

$$(5) \qquad v_n(i) = \max_k \left[q_i^k + \beta \sum_{j=1}^{m} p_{ij}^k v_{n-1}(j) \right], \qquad \begin{matrix} i = 1, 2, \ldots M \\ k = 1, 2, \ldots K \end{matrix}$$

in the following notation:

i = one of M possible states

k = one of K possible decisions

q_i^k = discrete valued approximation to $G(u,x)$

p_{ij}^k = conditional probability of going to state j under decision k when the current state is i

$v_n(i)$ = discrete valued approximation to $v_n(x)$.

Efficient solution methods to obtain the asymptotic decision rule in (5) as $n \to \infty$ are discussed below.

Several modifications of Bellman's standard successive approximation algorithm [Bellman 1957b] have been published, but their usefulness is difficult to assess. The following discussion draws heavily on an unpublished paper [Hendrikx, Van Nunen, and Wessels] which gives an excellent summary of the many modified algorithms and some comparative computation times on four applications. Hendrikx, *et al.* emphasize the importance of choosing a combination of variants in algorithms which exploits the special structure in an application, but some variants performed uniformly well across all four of their applications. Emphasis here will be placed on these more promising algorithms.

The method of successive approximation applied to (5) yields monotonic improvement in the decision rule as well as the value function $v_n(\cdot)$. An unequivocal convergence criterion comparable to that for Howard's policy improvement algorithm is not available for the decision rule. The relevant criterion in value iteration is to find a decision rule such that $v_n(i)$ is sufficiently close to v(i), the maximum expected present value possible. The basis for such a convergence criterion was provided by MacQueen who derived formulae for upper and lower bounds on v(i) that can be calculated at stage n. The iterations on (5) are terminated when the difference between the upper and lower bounds is less than some small value δ times $|v_n(i)|$ for all i.

Algorithms should always include this type of convergence criterion although the best strategy may be not to apply it at each iteration because of the extra computation. There is a tradeoff between calculating extra iterations and the computations required to check for convergence. When the number of alterative decisions at each stage is large relative to the number of states, convergence should be checked at each iteration, but as the ratio K/M get sufficiently small, every iteration is too frequent. A stochastic strategy could be used to advantage by checking only a subset of the M states for convergence at each iteration, with the subset chosen by random sampling. The best size sample would depend directly on K/M.

Most modified algorithms require extra computations at each iteration and would have more advantage when the maximization operation in (5) is relatively costly. However, some policy improvement variants, such as the Gauss-Seidel, can give impressive increases in convergence with essentially the same amount of computation at each iteration [Porteous]. The Gauss-Seidel variant exploits a near upper or lower triangular form in the matrix of transition probabilities under a given decision rule (common in replacement applications).

The bisection variant [Bartman] requires considerable extra computation at each iteration, but can greatly accelerate convergence when the policy improvement uniformly closes in one direction the gap between the upper and lower bounds on v(i) for all x at a given stage n. When the application has the right structure to exploit the Gauss-Seidel variant, the computation times reported in [Hendrikx, *et al.*] suggest that these two variants of the successive approximation algorithm complement one another. In one of these applications, the computation time was reduced by a factor of 25 and in another by 10.

A simple device which is advantageous when the number of decision alternatives is large relative to the number of states is to iterate (5) several times with the same decision rule [Van Nunen]. This procedure is somewhat of a combination between Howard's policy and Bellman's value iteration algorithms. Since the former tends to require fewer iterations but more difficult computations at each iteration than the latter, Van

Nunen's variant would seem to be a good strategy on some types of problems. A serious limitation of Howard's policy iteration algorithm is the large dimension of the system of linear equations that must be solved at each iteration, which is equal to the number states.

The following remarks on computational efficiency are applicable to generalizations of (5) where the q_i^k and p_{ij}^k are functions of the stage n, and thus, the decision rule is dependent on the stage and the maximum number of stages is finite. A large proportion of the elements in transition probability matrices are zero in most applications, and usually the positive elements are clustered together in adjacent columns for a given row. Both computer storage and computational requirements can be greatly reduced by structuring the algorithm to avoid the zeros. Taking i and k as given, let S(i,k) be the smallest integer for j where $p_{ij}^k>0$ and let L(i,k) be the largest. Then the lower and upper limits on the summation in (5) to take expected values can be replaced by S(i,k) and L(i,k), respectively. In many problems, the computations are reduced to a small fraction of what they would have been with (5). The cost of this technique is separate calculation and storage of the transition probabilities with S(i,k) and L(i,k) as leading entries in the array of probabilities for each i and k combination, but the storage space for S(·) and L(·) is usually trivial compared to the zero elements of the $\{p_{ij}^k\}$ for which storage space is eliminated.

In order to keep the number of states M to a manageable number, lengths of the intervals into which continuous state variables are divided sometimes get rather large. The usual procedure in calculating the $\{p_{ij}^k\}$ is to take the initial state i as the midpoint of the interval in a one state variable case, or the center of the multidimension cell if i is associated with a combination of several state variables. The following discussion uses a single state variable but the generalization is obvious.

As the interval gets larger and larger, using the midpoint of the interval as a reference point tends to bias the $\{p_{ij}^k\}$ in many decision processes. The bias occurs because the interval is too large relative to the mean and variance of the random variable which underlies the process. The problem can be largely overcome by specifying a uniform probability distribution for the position within the ith interval instead of using the midpoint of the interval as a deterministic reference point. Calculation of the $\{p_{ij}^k\}$ is more onerous, but this is a once and for all calculation. Suppose the interval is divided into ten subintervals and each subinterval has an equal conditional probability of occurring, given that the large interval is the current state i. Then the probability p_{ij}^k is calculated ten times, once for each of the midpoints of the ten subintervals taken as the given state i; the final probability is the simple average of these ten values. The large intervals will still yield an inferior decision rule compared to smaller intervals because less precise information is available on which to base decisions, but the compounding of probabilities will remove most of the systematic bias from using midpoints of the large intervals.

Another device that can be used in some applications to allow wider intervals in discrete approximations to state variables is lengthening the time period which specifies the stage of the process. This method worked well in an intertemporal ground water application [Burt 1964b] because changes in stocks (the state variable) are a slow, evolutionary process under an optimal decision rule. The stage was taken as a five-year period instead of one year which prevents the decision variable from changing more frequently than every five years. Intertemporal allocation problems in many natural resources would appear amenable to this type of modeling device, and in some respects, the restriction on frequency of changes in decision variables makes economic sense. In the case of water, the restriction forces a certain stability on the supplies of water available in the production process.

Computation of Ancillary Results

Although the optimal decision rule based

on an expected value criterion is the primary objective in most applications of stochastic DP, the finite Markovian form in (5) can be used to calculate much additional information about the problem. Various decision rules, whatever the source, can be compared with the optimal rule. Variance, or higher moments, of the return function can be calculated for any decision rule. The measure of returns can be present value or periodic (by stage), the latter in association with the equilibrium probability vector corresponding to the decision rule.

Starting from any initial state, the state probability vector for any future period (stage) can be calculated in a straightforward manner. Then the same kind of statistical measures on returns as for the equilibrium state probability vector can be calculated for the future period. This auxillary analysis could be done quite easily at a higher level of aggregation than that used to derive the optimal decision rule, i.e., the interval lengths in the state variables used to get (5) could be increased.

This type of computations for Markov processes is often overlooked, but could satisfy the apparent preference of many agricultural economists for the flexibility of simulation. Its advantage over simulation is the orderly structure of the modeling effort, and the availability of an optimal decision rule based on maximum expected value as a benchmark. Even without an optimal decision rule as a guide the Markov process computations would appear to be more efficient than stochastic simulation if very many alternative decision rules are evaluated.

A Neglected Area of Application

Although most agricultural economists, even those well indoctrinated to DP, are pessimistic about the usefulness of DP models applied to farm-firm growth and farm planning in general [Kennedy], I am optimistic about the opportunities for application of stochastic DP in this area. Clearly, such applications will not be routine formulations and will require a major research effort, but DP is still the most promising analytical model for the conceptual problems involved. The trick for success is an astute choice of a few state variables at an aggregated level of abstraction. What would appear to be serious omissions in state variables must be taken into account by nesting the omitted state variables into the model implicitly by various conditional optimizations. Various aspects of the firm growth process must be classified into purely static and short, intermediate, and long run dynamic components. All but the long run dynamic aspects of the process would be dealt with by various suboptimizations and approximations.

My limited experience with a simple firm growth model about ten years ago [Larson, Stauber, Burt] convinced me that stochastic DP is the most promising method for empirical analysis of the firm growth process. This model had only two state variables, (1) cash reserves (debt) divided by value of capacity, and (2) capacity measured in acres of cropland (one state was reserved for bankruptcy). The decision variable was amount of land bought or sold. The primary criterion was maximum expected value of net worth at the end of 25 years, but some experimentation with a concave utility function of terminal net worth was also done. Crop prices were taken as fixed and the main source of risk was from the variation in dryland wheat yields.

In spite of the simplicity and apparent naieveté of the model, it provided much information on the risks faced by operators in various financial positions, and the nature of their best strategies for survival and growth. Some contrasting results with and without government subsidy programs illustrated the dynamic reduction in risk from the government program. A necessary generalization under current world grain price instability is another state variable associated with an autoregressive model of wheat prices. Also, more sophistication in the way decisions are made to acquire and dispose of land are needed.

A particularly attractive safety first criterion which is easily used in stochastic DP is to

minimize the probability of net worth falling below a specified level at the end of a finite planning horizon. This criterion minimizes the sum of a subset of the probabilities of being in various states at the termination of the process, and is a simple modification of Bellman's original application of DP to Markov processes [Bellman 1957b]. The most conservative version of the criterion would minimize the probability of bankruptcy.

References

Arrow, Kenneth J. "Statistics and Economic Policy." *Econometrica*, 25(1957):523-31.

Athans, M. "The Discrete Time Linear-Quadratic-Gaussian Stochastic Control Problem." *Annals of Economic and Social Measurement*, 1(1972):449-91.

Bartman, D. "A Method of Bisection for Discounted Markov Decision Problems." *Zeitschrift für Operations Research*, 23(1979):275-87.

Bellman, Richard. *Dynamic Programming*. Princeton: Princeton University Press, 1957a.

———. "A Markovian Decision Process." *Journal of Mathematics and Mechanics*, 6(1957b):679-84.

———. *Adaptive Control Processes: A Guided Tour*. Princeton: Princeton University Press, 1961.

Burt, Oscar R. "Optimal Resource Use Over Time With an Application to Ground Water." *Management Science*, 11(1964a):80-93.

———. "The Economics of Conjunctive Use of Ground and Surface Water." *Hilgardia* 36(1964b):31-111.

———. "Groundwater Management Under Quadratic Criterion Functions." *Water Resources Research*, 3(1967):673-82.

———. "Ground Water Management and Surface Water Development for Irrigation." in *Economic Modeling for Water Policy Evaluation*, edited by R. M. Thrall, *et al.*, New York: North-Holland, 1976.

———. "Farm Level Economics of Soil Conservation in the Palouse Area of the Northwest." *American Journal of Agricultural Economics*, 63(1981):83-92.

Burt, Oscar R. and John R. Allison. "Farm Management Decisions with Dynamic Programming." *Journal of Farm Economics*, 45(1963):121-36.

Burt, Oscar R. and Ronald G. Cummings. "Natural Resources Management, the Steady State, and Approximately Optimal Decision Rules." *Land Economics*, 53(1977):1-22.

Burt, Oscar R., Won W. Koo, and Norman J. Dudley. "Optimal Stochastic Control of U.S. Wheat Stocks and Exports." *American Journal of Agricultural Economics*, 62(1980):173-87.

Dixon, Bruce L. and Richard E. Howitt. "Resource Production Under Uncertainty: A Stochastic Control Approach to Timber Harvest Scheduling." *American Journal of Agricultural Economics*, 62(1980):499-507.

Dreyfus, Stuart E. and Averill M. Law. *The Art and Theory of Dynamic Programming*. New York: Academic Press, 1977.

Giaver, H. *Optimal Dairy Cow Replacement Policies*. Unpublished Ph.D. thesis, University of California, Berkeley, 1966.

Hendrikx, Marcel, Jo Van Nunen, and Jaap Wessels. "Some Notes on Iterative Optimization of Structured Markov Decision Processes With Discounted Rewards." COSOR Memorandum 80-20, Eindhoven University of Technology, Eindhoven, The Netherlands, 1980.

Hildreth, Clifford. "Problems of Uncertainty in Farm Planning." *Journal of Farm Economics*, 39(1957):1430-1441.

Howard, Ronald A. *Dynamic Programming and Markov Processes*. New York: John Wiley and MIT Press, 1960.

Intrilligator, Michael D. *Mathematical Optimization and Economic Theory*. Englewood Cliffs: Prentice-Hall, 1971.

Kennedy, John O. S. "Applications of Dynamic Programming to Agriculture, Forestry, and Fisheries: Review and Prognosis." *Review of Marketing and Agricultural Economics*, 49 (Dec. 1981).

Kushner, Harold J. and Fred C. Schweppe. "A Maximum Principle for Stochastic Control Systems." *Journal of Mathematical Analysis and Applications*, 8(1964):287-302.

Larson, Donald K., M. S. Stauber, and O. R. Burt. *Economic Analysis of Farm Firm Growth in North-central Montana*. Research Report 62, Montana Agricultural Experiment Station, Bozeman, Montana, 1974.

MacQueen, J. "A Modified Dynamic Method for Marko-

vian Decision Problems." *Journal of Mathematical Analysis and Applications*, 14(1966):38-43.

Van Nunen, J. A. E. E. "A Set of Successive Approximation Methods for Discounted Markovian Decision Problems." *Zeitshcrift für Operations Research*, 20(1976):203-08.

Porteus, E. L. "Bounds and Transformations for Discounted Finite Markov Decision Chains." *Operations Research*, 23(1975):761-84.

Simon, Herbert A. "Dynamic Programming Under Uncertainty With a Quadratic Criterion Function." *Econometrica*, 24(1956):74-81.

Taylor, C. Robert, and Hovav Talpaz. "Approximately Optimal Carryover Levels for Wheat in the United States." *American Journal of Agricultural Economics*, 61(1979):32-40.

Theil, H. "A Note on Certainty Equivalence in Dynamic Programming." *Econometrica*, 25(1957):346-49.

Zusman, Pinhas and Amotz Amiad. "Simulation: A Tool for Farm Planning Under Conditions of Weather Uncertainty." *Journal of Farm Economics*, 47(1965):574-94.