Arthur F. Veinott, Jr.

Guide to Dynamic Programming

Dynamic Programming is the study of the optimal control of systems over time. Dynamic programs can be classified in many ways including the following.

Time Frame

- **Time Parameter:** discrete or continuous.
- **Horizon**: finite, infinite or random.
- **Time Dependence:** stationary or nonstationary (periodic, arithmetic, geometric).

Basic Data and Representation on Graph

- **States** $s \in \mathcal{S}$.
  - Number: finite, countable, uncountable.
  - Order: ordered, lattice, partially ordered, unordered.
  - Convexity.
- **Actions** $a \in A_s$.
  - Number: finite, countable, uncountable.
  - Order: ordered, lattice, partially ordered, unordered.
  - Convexity.
- **Transition Law** $p(t|s,a)$.
  - Deterministic or stochastic.
  - Individual or population/branching.
  - Linear or nonlinear.
  - Stochastic, substochastic, strictly substochastic, semi-Markov.
  - Circuitless.
  - Transient.
  - Bounded, recurrent.
- **Rewards** $r(s,a)$.
  - Nonlinearity: Linear, quadratic, convex, concave, general.
  - Composition: Additive, multiplicative, minimum, maximum, composed.
  - Risk posture and utility.
- **Immigration Streams.**
- **Stochastic Constraints.**

State Information

- **Perfect.**
- **Imperfect.**

Policies

- **Individuals.**
  - Nonanticipative.
  - Markov.
  - Deterministic or randomized.
  - Stationary or nonstationary (periodic).
  - Stopping, transient, halting.
  - Myopic.
  - Linear or nonlinear.
- **Homogeneous Populations:** cohort.

Multiperson

- **Teams:** common objective, but different information.
- **Games:** differing objectives.

Preference Relations

Maximum Value/Utility.

Expected value: N-period, circuitless, transient, positive, negative.
Expected additive utility.
Expected multiplicative utility (constant risk posture).
Expected instantaneous rate of return.
Instantaneous rate of expected return.
Expected present value.
Expected future value.

Immigration-Induced Optimality: physical and value interpretations, unified preferences.

Maximum Present Value for Small Interest Rates.

Strong maximum present value.
Limiting present value for binomial immigration streams.
Moment Optimality.

Cesàro Overtaking.

Strong Cesàro overtaking.
Cesàro overtaking for binomial immigration streams: reward rate, total reward, float.
Future value: negative interest rates and inflation.

Cesàro Geometric Overtaking.

Expected multiplicative utility (constant risk posture).

Equivalence: Limiting present value and Cesàro overtaking preference relations.

Maximin.

Issues

Existence of Optimal Policies.

Characterization of Optimal Policies and their Values.

Computation of Optimal Policies and their Values.

- Methods.
  - Value iteration: successive approximations, Gauss Seidel.
  - Policy improvement.
    - Simplex method (linear programming).
    - Newton's method.
  - Index rules
  - Certainty equivalents.
- Approximation.
- Complexity.

Sensitivity Analysis.

Interest rate.
Horizon length.
Immigration stream.

Structure.

Accessibility/controllability.
Qualitative behavior.
State-action tradeoff.
Size of state-action description.

Tools Used in Proofs.

Comparison Lemma.
Policy improvement, successive approximations.
Expansions: Neumann, Laurent, polynomial.
Choice of rewards and final value.

Applications

Reliability/Quality Control/Maintenance/Replacement: 100M$ Arizona road maintenance.
Cutting Stock: 100M$ Weyerhaeuser log cutting.
Finance: 2B$ AT&T bond refunding, portfolio management, procurement/sale timing.
Water/Power: reservoir and hydropower management.
Claims Processing: EDS using for a score of years.
Insurance: segmenting customers by risk categories.
Credit Management.
Disaster Relief Strategies: Floods, hurricanes, etc.
Guidance/Control: vehicles in space.
Transportation: airline seat booking/overbooking, vehicle routing, air-traffic control.
Production Planning/Inventory Control: production/procurement scale and timing.
Purchasing/Selling: assets, commodities, products.
Pipeline Operations: stepping up pipeline pressure.
Capacity Planning: acquisition/sale, scale and timing.
Optimal Algorithms: order of multiplying matrices, search for maxima/zeros of functions, sorting, etc.
Combinatorial/Integer Programming: knapsack, concave-cost network flows, duality.
Text Formatting: paragraphing TEX.
Language: speech recognition.
Statistics: sequential analysis, sequential design of experiments.
Scheduling: PERT.
Telecommunications: routing calls.
Controlled Queues: arrival/service rates, routing.
Manpower Management: recruitment/promotion/retention policies.
Harvesting/Culling: fish, animals, wildlife, trees, crops.
Health Care: cancer screening, dosing strategies, hospital admissions.
Emergency-Service Provisioning: police, fire, ambulance.
Economics: economic growth, managing strategic oil reserve.
Psychology: order of teaching concepts.
Agriculture: integrated pest management, soil management, frost protection, irrigation level.
Marketing: scale and timing of advertising.

Early History (1875-1955)

Secretary Problem: Arthur Cayley (1875), *London Times*.
Wald: sequential analysis in 1940s.
Masse: reservoir control in 1940s.
Shapley: introduced finite Markov decision chains and two-person stochastic games in 1953.
Bellman: coined term *dynamic programming* and demonstrated the breadth of application in 1950s.
Pontryagin: introduced optimal control in 1950s.

Comparison of Function-Space and State-Space Dynamic Mathematical Programming Formulations

Function Space	State Space
Linear programs	Discrete-time-parameter finite state and action
Countable linear programs	Discrete-time-parameter countable state and action
Abstract linear programs	General time-parameter, state and action
Unconstrained convex quadratic	Convex quadratic (certainty equivalents)
Calculus of variations	Continuous-time-parameter optimal deterministic control