IDA

INSTITUTE FOR DEFENSE ANALYSES

Structural Dynamic Programming Models Applications for DOD Research

DATAWorks 2022

Mikhail Smirnov

March 2022
Approved for public release;
distribution is unlimited.
IDA Document NS D-33019
Log: H 22-000105

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22301

About This Publication
The work was conducted by the Institute for Defense Analyses (IDA) under CRP C6608.

For More Information:
Dr. John W. Dennis III, Project Leader
jdennis@ida.org, 703-845-2166
ADM John C. Harvey, Jr., USN (ret) Director, SFRD
jharvey@ida.org, 703-575-4530

Executive Summary

Dynamic programming (DP) is a method for finding optimal decision rules for dynamic sequential decision problems under uncertainty. This is a very large and general class of problems, and DP models are a key component of most game theoretic and applied macro- and micro-economic models, including analyses of many applications that are relevant to the Department of Defense (DOD). The structure of a DP model includes intertemporal optimization over a finite or infinite horizon, transition functions for state variables, and a payoff function. This structure allows DP models to be used for causal inference and analysis of valid counterfactual or hypothetical policies, which is not always possible with other machine learning (ML) models.

The main drawback to using DP models has been the computational complexity that comes from the curse of dimensionality. DP models that are easy enough to solve have generally been too simplistic for applied work, which has left DP as a mostly academic pursuit. Recent developments in approximation methods and econometric algorithms, combined with greater availability of raw computational power, make solving complex and realistic DP models possible. Neural networks and other ML algorithms can be used to approximate solutions to DP models, and conditional choice probability methods can be used to estimate DP models that incorporate endogeneity due to persistent unobserved heterogeneity.

Many DOD applications can benefit from recent analytical advancements utilizing complex, realistic DP models. We can improve military readiness by analyzing maintainers' decisions to repair or replace critical parts, investment/divestment decisions for specific platforms, and the optimal way to structure a phased readiness cycle. Personnel policy can benefit from designing incentives that are necessary to retain the right service members at the right time in their careers and develop education and training plans that have the best payoffs. Additionally, we can examine strategic competition through the lens of DP game theoretic models to help the DOD make optimal choices in uncertain environments.

This page is intentionally blank.

Structural Dynamic Programming Models

Applications for DOD Research

DATAWorks 2022

Mikhail Smirnov

April 28, 2022

Institute for Defense Analyses
730 E. Glebe Rd. • Alexandria, VA 22305

What is dynamic programming (DP)?

Finding optimal decision rules for dynamic sequential decision problems under uncertainty

The foundation of most game theory and applied macro- and microeconomics

Examples of dynamic sequential decision problems:

- Maintainer: when to replace an engine on a vehicle?
- Planner: what is the right level of inventory to keep on hand?
- Service member: re-enlist or leave the service?

1

Structure of a DP model

$$V(s_0) = \max_{\{d_0,\ldots,d_T\}} E\{\sum_{t=0}^{T} \beta^t u(s_t, d_t) | s_0\}$$

t indexes time

s are state variables

$d \in D$ are decisions

β is a discount factor

u is a payoff function

p is a transition function

Commonly referred to as the "structure" of the model

$$V(s) = \max_{d \in D}[u(s,d) + \beta \int V(s')p(s'|s,d)]$$

2

Why structural DP?

Huge class of extremely important problems

Structure allows for causal inference

$u(x,s)$ gives information about payoffs and the decision process

Easy to test counterfactual policies or scenarios

Contrast with machine learning (ML) models

Expert input required to properly specify the models

Static models miss the intertemporal nature of the problem

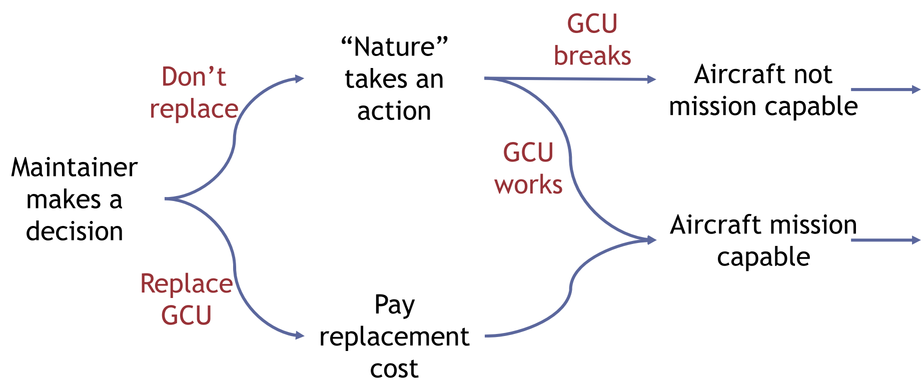
3

IDA

Aircraft GCU replacement

Every period a maintainer has to decide to replace an aircraft generator control unit (GCU) or not

- Replacing a GCU incurs a cost
- GCUs with more hours on them have a higher failure rate



IDA

4

Using the GCU model

We observe the decisions of the maintainer and want to learn the cost and their risk tolerance

$u(hours, replace) = cost$

$u(hours, keep) = risk * hours$

Is the cost implied by the maintainer's behavior equal to monetary cost?

What would happen if the reliability improved 10%?

IDA

5

Main challenge to using DP is tractability

Curse of dimensionality

Estimating DP models involves solving the DP many times

Realism vs tractability

- All models are wrong, some models are useful
- Specifying the right model and objective
- Much more popular in academia than in applied settings

It is now easier than ever to use DP

Complex and realistic DP models are now possible to use

- Solving DP models that are good representations of reality
- Estimating models where some relevant information is unobserved

Computational advances

- Chess and Go are DP problems that AI algorithms have "solved"
- Raw computational power is cheap and abundant

Approximation methods

Econometric algorithms

IDA

Approximation methods

Old idea, now possible

Approximate the value function using a parametric form

$V \sim V_\omega$, where ω is a finite set of parameters

For example, neural networks or polynomial expansions

ω can be estimated using nonlinear least squares or similar

Sophisticated ML approximations can come arbitrarily close to "solving" the problem

IDA

Econometric algorithms

Key problem for econometric inference – unobserved variables aka "endogeneity"

$u(d, s, z)$ includes a variable z that we do not observe

Biased estimator if z is ignored

Solutions to the endogeneity problem

All require significant computational power

Specify a distribution $G(z)$ and integrate it out

Approximation methods make this tractable

IDA

GPU replacement, again

More detailed model

- Include additional information about the aircraft
- Information about the missions/readiness phase
- Payoff structure that depends on the type/timing of failure
- Most aircraft have more than one GCU

Including unobserved variables

- Risk tolerance depends on mission
- Some GCUs may simply be "duds"
- Previous service history

10

Other military applications

Investments in readiness

- What is the right number of a specific platform to procure?
- What is the best way to structure a phased readiness cycle?

Personnel policy

- What incentives are necessary to retain the right service members?
- What education and training investments have the best payoff?

Strategic competition

11

Mikhail Smirnov
msmirnov@ida.org

This page is intentionally blank.