Clearing the Jungle of Stochastic Optimization

PAPER BY: WARREN B. POWELL

PRESENTED BY: MEHDI DADFARNIA

Outline

Decisions

State of sequential, stochastic decision-making in literature

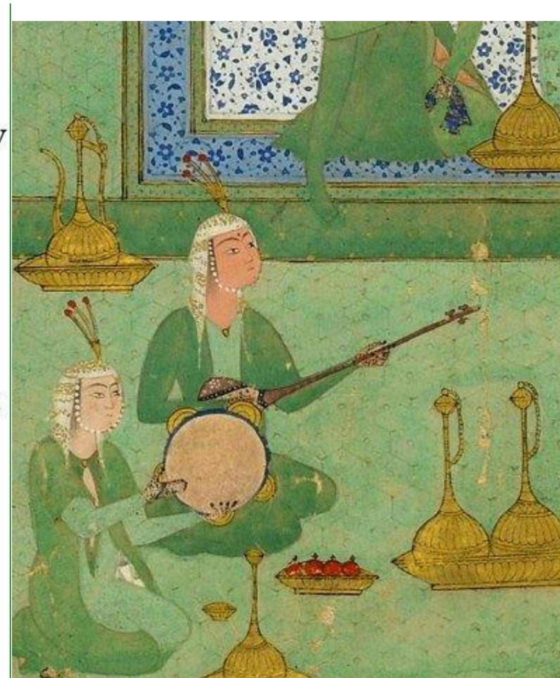
Stochastic optimization models

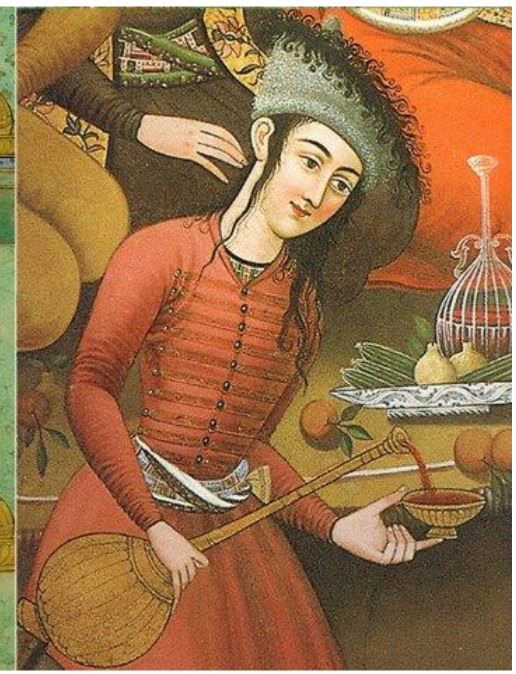
Stochastic optimization solution strategies: Policies!

- Brief Example

Decisions?

If an important decision is to be made, they [the Persians] discuss the question when they are drunk, and the following day the master of the house where the discussion was held submits their decision for reconsideration when they are sober. If they still approve it, it is adopted; if not, it is abandoned. Conversely, any decision they make when they are sober, is reconsidered afterwards when they are drunk.





Modeling Decision Problems

Deterministic Modeling

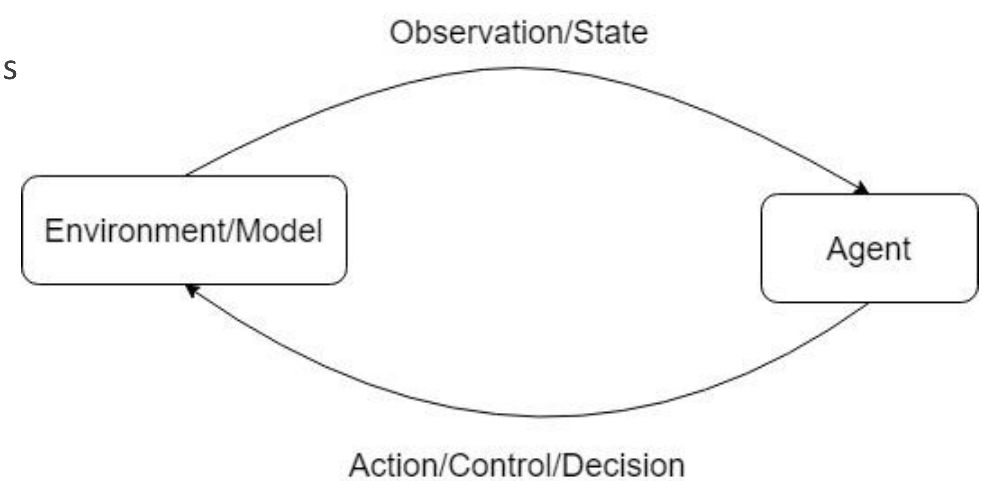
Stochastic Modeling
- Model Uncertainty (dynamics are unknown, agent learns through state/observation, e.g. RL)
- State Uncertainty (noise in agent's observation)

Sequential

Single Agent

Multi-Agent

Game Theory Side Note:
- At least 2 agents
- Usually no probabilistic model of behavior of other agents/environment, but models of their utilities exists



Deterministic Optimization Modeling

Canonical form (for optimization or canonical modeling)

- Minimize costs, due to some constraints, etc

$$
\begin{aligned}
&\min cx \\
&Ax = b \\
&x \geq 0
\end{aligned}
$$

Or its sequential (over time) form:

$$
\begin{aligned}
&\min \sum_{t=0}^{T} c_t x_t \\
&A_t x_t - B_{t-1} x_{t-1} = b_t \\
&D_t x_t \leq u_t \\
&x_t \geq 0
\end{aligned}
$$

Stochastic Optimization Modeling

Modeling sequential decision problems under uncertainty

- Parameters unknown at time of decision
- Probability distribution of parameter estimated; imply expected values of objective functions
- "Jungle of books" across dozens of academic fields – but why???

- **Can there be a canonical form for stochastic optimization modeling?**

Key Differences in Solution Strategies

Deterministic optimization

- Solution is an **optimal decision** (i.e. a decision vector)

Multistage stochastic optimization

- Solution is a (hopefully) **optimal function** (known as policy)
- Four classes of policies (or functions) exist;
- Four classes can be applied to any multistage stochastic opt. problem
- Quite rare to actually find the optimal policy (trade-offs)

“Models” in the “Jungle of Books”

Note: Algorithm, not an actual model.

Also note: models without algorithms are useless.

Dynamic Programming: Bellman’s Equation

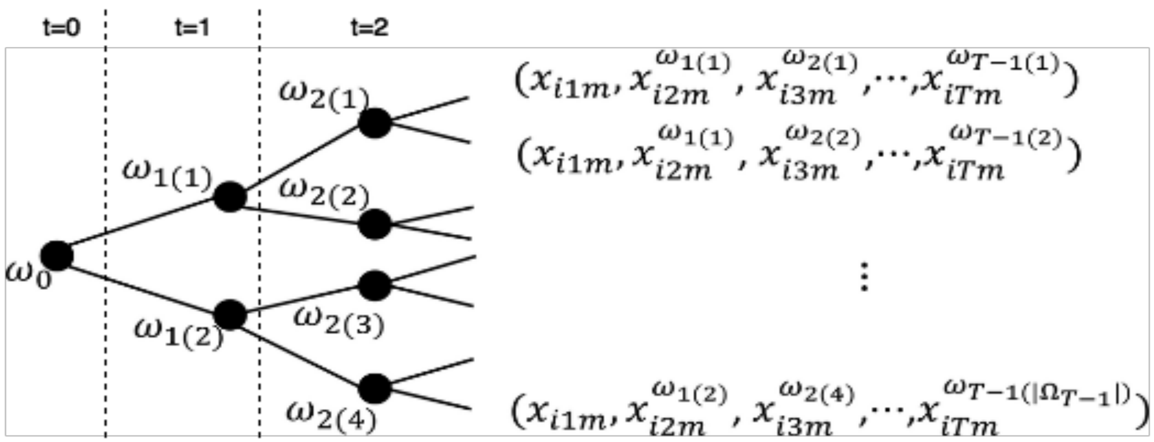
$$V(s) = \min_a \left(C(s,a) + \gamma \sum_{s'} p(s' \mid s,a) V(s') \right)$$

$$\begin{aligned} s &= \text{"State variable"} \\ a &= \text{Discrete action} \\ p(s' \mid s,a) &= \text{"Model" (transition matrix, transition kernel)} \\ V(s) &= \text{Value of being in state } s \\ \gamma &= \text{Discount factor} \end{aligned}$$

"Models" in the "Jungle of Books"

Another algorithm: Stochastic Programming

$$\min c_t x_t + \sum_{\omega_t \in \Omega_t} p(\omega_t) \sum_{t'=t+1}^{t+H} c_{tt'}(\omega) x_{tt'}(\omega)$$



Each node ω_t is a scenario in period *t* and has a probability $p^{\omega t}$;

Decision x_t made for time *t*

5 Elements to Model a Stochastic, Sequential Decision

Powell’s proposal to capture problem’s characteristics in a model:

- State variables
- Decision variables
- Exogenous information (~environment, external)
- Transition function
- Objective function

Other characteristics:

- Time scales of decision-making (seconds vs weeks vs decades)
- Modeling uncertainty from transition function, state observation, exogenous information

State Variables

Hard to define across communities

- Powell: Minimally dimensional function of history that is necessary+sufficient to compute the decision problem

Divided into:

- Physical/Resource state (energy in storage)
- Information state (prices of energy)
- Knowledge state, belief about unobservables (belief about equipment status)
- Information state may include resources state, and the knowledge state may include information state; division aids thought-process

$\boldsymbol{S_t=(R_t, I_t, K_t)}$

Decision Variables

OR community: Decision variables x

Comp Sci folks: Actions a

Control theory people: Control signals u

Don't specify a decision $x/a/u$, BUT:

- Define the function $X_t^{\Pi}(S_t)$ where Π is the policy (or function) that produces feasible actions/decisions
- Set of feasible actions may depend on the state

Exogenous Information

Environmental, external, e.g. prices, costs, forecasts

Random variables W_t are observed in sequence $W_1, W_2, \ldots$ so that states, actions, and exogenous information evolve as:

$$(S_0, x_0, W_1, S_1, x_1, W_2, \ldots, S_t, x_t, W_{t+1}, \ldots, S_T)$$

ω refers to the sample realization of the random variables $W_1, W_2, \ldots$ and Ω is the set of all possible realizations.

Transition Function

Synonyms across communities: model, plant model, transition law, etc

$$S_{t+1} = S^M\left(S_t, x_t, W_{t+1}(\omega)\right)$$

Describes evolution of system from t to t+1; depends on current state, the decision, and the exogenous information realized by t+1

When unknown, the exogenous information is unobservable and problem is known as “model-free”

Objective Function

Cost function we want to minimize, given the transition function $S_{t+1} = S^M\left(S_t, x_t, W_{t+1}(\omega)\right)$

$$\min_{\pi} E^{\pi}\left\{\sum_{t=0}^{T} \gamma^t C\left(S_t, X_t^{\pi}(S_t)\right)\right\}$$

Expectation over all random outcomes

Cost function

Decision function (policy)

Finding the best policy/function pi

Cost function may also be dependent on the new exogenous information W_{t+1}, or next state S_{t+1}

Other minimizations: minimizing risk measures or robustness

How to search over an abstract space of policies?

Powell’s 4 fundamental classes of policies/functions:

- Policy function approximations (PFAs)
- Optimizing a cost function approximation (CFAs)
- Policies that depend on a value function approximation (VFAs)
- Lookahead policies

Hybrids between these 4 fundamental classes are also possible

Search through their tunable parameters

Policy Function Approximations

Lookup tables
- Given a state, take a certain action

Parametric models
- Linear & Nonlinear graphs

Rules
- Inventory model: product inventory dips below x, bring it up to y (y>x)

Imbedded decisions
- Explicit programming of doing tasks

Tunable Parameters: regression parameters on parametric models, parameters defining the rules (x,y), lookup table and imbedded decision state-to-action mapping

$$X^{PFA}(S_t) = \theta_0 + \theta_1 S_t + \theta_2 S_t^2$$

Cost Function Approximations

Basically a cost function, not too creative

$$X^{CFA}(S_t \mid \theta) = \arg\min_{x_t \in X_t^{\pi}(\theta)} \bar{C}^{\pi}(S_t, x_t \mid \theta)$$

Minimizing the cost, usually with a parametric approximation

- Basis functions of the form $S_t * x_t$, $S_t * x_t^2$, x, x^2, *etc.*

Tunable Parameters: θ captures all tunable parameters based on how cost function is set up;

E.g. parameters may be bonus or penalties to encourage behaviors

Value Function Approximations

Dynamic programming

$$X_t^{VFA}(S_t) = \arg\min_{x_t} \left(C(S_t, x_t) + \gamma \bar{V}_t^x \left(S_t^x(S_t, x_t) \right) \right)$$

Value function:

- Depends on state immediately after decision has been made, before new data arrives
- Does not depend on computationally expensive expectations

Different than CFAs:

- Mechanisms for fitting regression are different,
- The future contributions (given the current state) are actually approximated with VFAs i.e. uses some idea of the future

Tunable Parameters: θ captures all tunable parameters based on how value function is set up

Lookahead Policies

In literature, look-ahead models referred to as "the model"

Approximation of base model, used in lookahead policies

- Limiting time horizon
- Dimensionality reduction
- Discretization

Base models require determining decisions AND policies for every time period in the future

Tunable Parameter: type of approximation, planning horizon, number of scenarios, etc

Premise: optimizing the model over a horizon of time

Deterministic lookahead/model-predictive control

$$X_t^{LA-D}(S_t) = \arg \min_{\tilde{x}_{tt},\ldots,\tilde{x}_{t,t+H}} C(\tilde{S}_{tt}, \tilde{x}_{tt}) + \sum_{t'=t+1}^{T} C(\tilde{S}_{tt'}, \tilde{x}_{tt'})$$

Stochastic lookahead/stochastic programming

$$X_t^{LA-S}(S_t) = \arg\min_{\tilde{x}_{tt},\tilde{x}_{t,t+1},\ldots,\tilde{x}_{t,t+T}} C(\tilde{S}_{tt}, \tilde{x}_{tt}) + \sum_{\tilde{\omega}\in\tilde{\Omega}_t} p(\tilde{\omega}) \sum_{t'=t+1}^{T} \gamma^{t'-t} C(\tilde{S}_{tt'}(\tilde{\omega}), \tilde{x}_{tt'}(\tilde{\omega}))$$

Other types

Policy Evaluations & Recommendations

Computing the objective is rare, Monte Carlo simulations used to estimate value of a policy.

PFA: best for low-dimensional problems where policy structure is apparent from the problem

CFA: better for high-dimensional function where the deterministic variation of the model works well; able to achieve desired behavior by manipulating the cost function

VFA: Useful when value of the future is easy to approximate; can deal with high-dimensional problems but has issues with nonseparable interactions

Lookahead: Great if forecast available; deterministic and stochastic lookahead policies should always be compared; should only be used when all else fails (more computationally expensive)

Example: Energy Storage

Major takeaway:

- The same problem with slightly different data each has a different optimal policy

Problem:	Problem description	PFA	CFA Error correction	VFA	Determ. Lookahead	CFA Lookahead
A	A stationary problem with heavy-tailed prices, relatively low noise, moderately accurate forecasts.	**0.959**	0.839	0.936	0.887	0.887
B	A time-dependent problem with daily load patterns, no seasonalities in energy and price, relatively low noise, less accurate forecasts.	0.714	**0.752**	0.712	0.746	0.746
C	A time-dependent problem with daily load, energy and price patterns, relatively high noise, forecast errors increase over horizon.	0.865	0.590	**0.914**	0.886	0.886
D	A time-dependent problem, relatively low noise, very accurate forecasts.	0.962	0.749	0.971	**0.997**	0.997
E	Same as (C), but the forecast errors are stationary over the planning horizon.	0.865	0.590	0.914	0.922	**0.934**

Joint research with Prof. Stephan Meisel, University of Munster, Germany.

- Deterministic
  - » Objective function

$$\min_{x_0,\dots,x_T} \sum_{t=0}^{T} c_t x_t$$

  - » Decision variables:

$$(x_0,\dots,x_T)$$

  - » Constraints:
    - at time t

$$\left.\begin{array}{r} A_t x_t = R_t \\ x_t \geq 0 \end{array}\right\} \mathcal{X}_t$$

    - Transition function

$$R_{t+1} = b_{t+1} + B_t x_t$$

- Stochastic
  - » Objective function

$$\min_\pi E^\pi \left\{ \sum_{t=0}^{T} \gamma^t C\left(S_t, X_t^\pi(S_t)\right) \mid S_0 \right\}$$

  - » Policy

$$X^\pi : S \mapsto \mathcal{X}$$

  - » Constraints at time t

$$x_t = X_t^\pi(S_t) \in \mathcal{X}_t$$

  - » Transition function

$$S_{t+1} = S^M\left(S_t, x_t, W_{t+1}\right)$$

  - » Exogenous information

$$(W_1, W_2, \dots, W_T)$$