Optimal Sequential Decisions

Extended from Chapter 5 of
Sustainable Management of Natural Resources. Mathematical Models and Methods
by Luc DOYEN and Michel DE LARA

Michel De Lara
Cermics, École des Ponts ParisTech
Université Paris-Est
France

École des Ponts ParisTech

January 4, 2016

Outline of the presentation

1. A bird's eye view of trade-offs in intertemporal optimization
2. Dressing an intertemporal optimization problem in formal clothes
3. Pontryaguin's "maximum" principle and Hotelling rule
4. Dynamic programming for the additive payoff case
5. Examples in natural resources optimal management
6. Non additive criteria

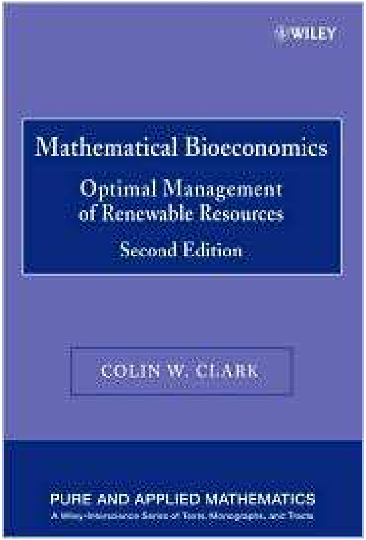
Outline of the presentation

Colin W. Clark's *Mathematical Bio-economics* (1976)

"Perhaps the most important initial realization for the question of sustainable development is that the overwhelming environmental and resource problems now facing humanity are the result of economically rational individual decisions made every day by each and every one of us"

CLARK C.W., (1976), *Mathematical Bio-economics. The Optimal Management of Renewable Resources*, J. Wiley & Sons, New York

What is the deep reason why whales have been depleted?



It is economically sensible to deplete the stock of whales

- Imagine that you manage a stock of 1,000 whales.
  How can you exploit them?
- Any "good" whales manager should be a "good" manager, whales or not, and should consider the following data on yearly growth rates

	yearly growth rate
whales	2 — 5%
money	5%

- Any "good" manager should consider at least two management strategies
  - harvest every year the "whales surplus", say 3% of the stocks, and sell it; the population is stationary, and the process can go forever
  - deplete the stock of whales, sell it and invest the money at 5%
- A whales manager would be economically sensible to sell its stocks and to invest this money at a higher rate

Here are intertemporal trade-offs

- Getting more fish catches now is at the expense of having less stock in one year
- In a hydropower dam, turbinating water today reduces the stock, hence precluding to turbinate much in the future when prices may be higher

Many economic analysis and insights are based upon (dynamic) optimization models

- Optimality stands as a basic tool in economics
  - utility maximization, cost minimization
  - cost-benefit analysis
  - cost-effectiveness analysis
- Many economic analysis are built upon optimization
  - Ramsey growth model on optimal allocation of levels of consumption/investment over successive generations
  - Solow long-run economic growth model
  - Hotelling rule on optimal extraction path of an exhaustible resource
  - discounted utility criterion for climate change policy evaluation

The use of optimization is said to have been triggered by the Second World War

The Map Room of W. Churchill's Cabinet War Rooms

- In May 1940, the Royal Air Force (RAF) sustained massive losses of 500 operational fighters
- and remained with 620 operational fighters, well below the 1,200 fighters thought to be the minimum number needed to win an air battle over the U. K.
- W. Churchill brought in experts to determine a program that best met objectives without exceeding existing limited resources

What is "optimization"?

Optimizing is obtaining the best compromise between needs and resources
Marcel Boiteux (président d'honneur d'Électricité de France)

- Needs: multiple targets
- Resources: multiple limits and multiple possible allocations
- Best compromise: value, trade-offs, balancing the costs against the benefits

Optimization in the wild

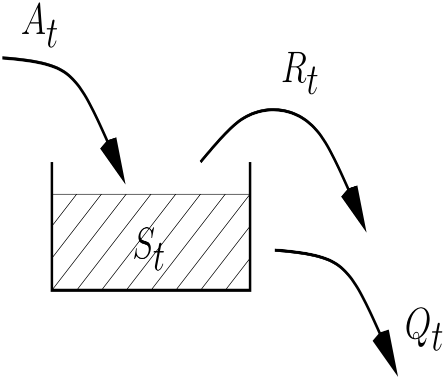
- "In laying bricks, the motions used in laying a single brick were reduced from 18 to 5, with an increase in output from 120 bricks an hour to 350 an hour", Frank B. Gilbreth, 1915
- "The motions of a girl putting paper on boxes of shoe polish were studied. Her methods were-changed only slightly, and where she had been doing 24 boxes in 40 seconds, she did 24 in 20 seconds, with less effort", Frank B. Gilbreth, 1915
- Improvement in military actions:
  "1000 percent increase in bombs on targets",
  "optimal size of a merchant convoy", etc.

in S. I. Gass and A. A. Assad,
An Annotated Timeline of Operations Research: An Informal History

The Diet Problem, or the birth of Linear Programming by George Dantzig

- Jerry Cornfield's diet problem: the Army wanted a low cost diet that would meet the nutritional needs of a GI soldier
- A footnote at the very end of the article *The Cost of Subsistence* by George Stigler reveals that bureaucrats had recommended low cost diets that cost twice as much as Stigler's (suboptimal) solution
- The optimal solution comprises spinach ;-)
- George Dantzig: "The trouble with a diet is that one's always hungry. What I need to do is maximize the feeling of feeling full."
- George Dantzig: "I placed an upper bound of three on the number of bouillon cubes consumed per day. That was how upper bounds on variables in linear programming first began."

Outline of the presentation


A_t
R_t
S_t
Q_t

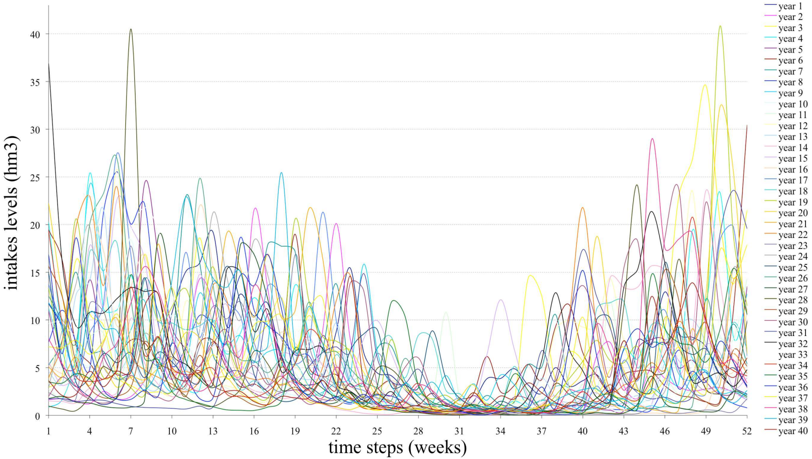
We pay attention to time periods when we lay out descriptive variables

- time $t \in \overline{\mathbb{T}} = \{t_0, t_0 + 1, \dots, T - 1, T\}$ is discrete,
  and t denotes the beginning of the period $[t, t+1[$ (day or 8 hours)
- stock $S(t)$, volume of water at the beginning of period $[t, t+1[$
- turbined outflow volume $q(t)$ during $[t, t+1[$
  - decided at the beginning of period $[t, t+1[$
  - $q(t)$ is a control variable in the hands of the decision-maker
    who can pick up its value in a given range

As for data, we distinguish between parameters and scenarios: parameters do not depend on time

- $S^\sharp$ is the dam capacity (maximum dam volume)
- $q^\sharp$ is the outflow capacity,
  the maximum which can be turbined by time unit (and produce electricity)
- time $t \in \overline{\mathbb{T}} = \{t_0, t_0 + 1, \ldots, T - 1, T\}$ is discrete,
  - t_0 is the initial time
  - T is the horizon

Water inflows historical scenarios



As for data, we distinguish between parameters and scenarios: scenarios are time sequences

A scenario is a time sequence

that runs from initial time t_0 to horizon T or to $T-1$

- inflow water volume $a(t)$ (rain, upper dam outflow) during $[t, t+1[$ gives rise to a scenario of water inflows

$$a(\cdot) = \big(a(t_0), \dots, a(T-1)\big)$$

- price $p(t)$ of water turbined during $[t, t+1[$ gives rise to a scenario of prices

$$p(\cdot) = \big(p(t_0), \dots, p(T-1)\big)$$

In the deterministic setting, a single scenario is part of the data of the problem, hence is supposed to be known in advance

Static constraints relate variables at the same time t

- Stocks are volumes of water

$$0 \leq S(t) \leq \underbrace{S^{\sharp}}_{\text{dam capacity}}$$

- Turbined outflows are volumes of water taken from the stock

$$0 \leq q(t) \leq S(t)$$

- Turbined outflows are limited by turbined capacity

$$0 \leq q(t) \leq \underbrace{q^{\sharp}}_{\text{turbined capacity}}$$

A dynamical equation relates variables between times t and $t+1$

More precisely, a so-called "state variable" at time $t+1$
is expressed as a function of variables at time t

$$\underbrace{S(t+1)}_{\text{future volume}} = \min\{S^{\sharp}, \underbrace{S(t)}_{\text{volume}} - \underbrace{q(t)}_{\text{turbined}} + \underbrace{a(t)}_{\text{inflow volume}}\}, \quad t = t_0, t_0+1, \ldots, T-1$$

- $S(t)$ volume (stock) of water at the beginning of period $[t, t+1[$
- $a(t)$ inflow water volume during $[t, t+1[$
- $q(t)$ turbined outflow volume during $[t, t+1[$
  - decided at the beginning of period $[t, t+1[$
  - chosen such that

$$0 \leq q(t) \leq S(t) \text{ and } 0 \leq q(t) \leq q^{\sharp}$$

The so-called "history space" can have a large size

Single dam histories

$$(S(\cdot), q(\cdot)) = \Big(\overbrace{S(t_0), \ldots, S(T)}^{\text{stocks}}, \overbrace{q(t_0), \ldots, q(T-1)}^{\text{turbined}}\Big)$$

- For a single dam managed over a year with one turbinating decision a day, $\big(S(\cdot), q(\cdot)\big) \in \mathbb{R}^{366+365} = \mathbb{R}^{731}$
- For a single dam managed over a year with three turbinating decisions a day, $\big(S(\cdot), q(\cdot)\big) \in \mathbb{R}^{3\times 731} = \mathbb{R}^{2193}$
- For five dams managed over a year with three turbinating decisions a day, $\big(S(\cdot), q(\cdot)\big) \in \mathbb{R}^{5\times 2193} = \mathbb{R}^{10965}$

State and control constraints reduce the set of admissible trajectories to account for feasibility issues

Admissible trajectories for a single dam dynamical model

$$\left\{ \big(S(\cdot), q(\cdot)\big) \;\middle|\; \begin{array}{rll} S(t_0) = & S_0 \; , & \\ S(t+1) = & \min\{S^\sharp, S(t) - q(t) + a(t)\} \; , & t \in \mathbb{T} \\ q(t) \in & [0, \min\{q^\sharp, S(t)\}] & t \in \mathbb{T} \end{array} \right\}$$

When each history has a value, optimization is possible

- "In economics, there is the concept of value"
- Optimization requires measuring value

Intertemporal payoff for a single dam

$$\sum_{t=t_0}^{T-1} \overbrace{\underbrace{p(t)}_{\text{price } \times} \underbrace{q(t)}_{\text{quantity}}}^{\text{turbined water profit}} + \overbrace{K\big(S(T)\big)}^{\text{final stock utility}}$$

We are now ready to dress an intertemporal optimization problem in formal clothes

- Domain: the product set of histories with generic element $\big(S(\cdot), q(\cdot)\big)$
- Criterion/objective function:

$$\big(S(\cdot), q(\cdot)\big) \mapsto \sum_{t=t_0}^{T-1} \underbrace{p(t)q(t)}_{\text{instantaneous payoff}} + \underbrace{\mathtt{K}\big(S(T)\big)}_{\text{final payoff}}$$

- Constraints: induced by state, control and dynamics constraints

$$\mathcal{T}^{\mathrm{ad}}(t_0, S_0) = \left\{ \big(S(\cdot), q(\cdot)\big) \;\middle|\; \begin{array}{rll} S(t_0) = & S_0 \;, & \\ S(t+1) = & \min\{S^\sharp, S(t) - q(t) + a(t)\} \;, & t \in \mathbb{T} \\ q(t) \in & [0, \min\{q^\sharp, S(t)\}] & t \in \mathbb{T} \end{array} \right\}$$

$$\max_{\big(S(\cdot), q(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}(t_0, S_0)} \sum_{t=t_0}^{T-1} p(t)q(t) + \mathtt{K}\big(S(T)\big)$$

We know illustrate the two well-known approaches
to tackle intertemporal optimization problems

- à la Pontryaguin
- à la Bellman

In the Pontryaguin approach, the dynamics equations are treated as constraints that inherit a multiplier, the adjoint state

- We can handle the resolution of the intertemporal optimization problem by mathematical programming, that is, the maximization of a function over $\mathbb{R}^n$ under equality and inequality constraints
- In one approach, one writes the KKT conditions for the equality constraints induced by the dynamics equations
- In another approach, one annihilates a gradient
  - the volumes $S(t)$ are intermediary variables, completely determined by the choice of the controls $q(t)$
  - however, such intermediary variables may be used to compute gradients by means of an adjoint state

In the Bellman approach, a value is attached to a stock at any time

Question: What happens to the final stock if we maximize $\sum_{t=t_0}^{T-1} p(t)q(t) + \mathtt{K}\big(S(T)\big)$ with final value of water $\mathtt{K}(S) = 0$?

Answer:

In economic terms, the final value of water $\mathtt{K}(S)$
is the best payoff that can be obtained over an infinite horizon
starting from the stock S at time t_0

To beat myopic strategies, an intuition of Bellman equation

- A myopic decision-maker maximizes, at each time t, the instantaneous payoff:

$$\max_{0 \leq q(t) \leq \min\{q^\sharp, S(t)\}} p(t)q(t)$$

- Denote by $V(t,S) = \max_{\big(S(\cdot),q(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}(t,S)} \sum_{s=t}^{T-1} p(s)q(s) + \mathtt{K}\big(S(T)\big)$ the best payoff that can be achieved over the time range $\{t, t+1, \ldots, T-1, T\}$ starting from the stock S at time t
- A non myopic decision-maker solves, at each time t

$$\max_{0 \leq q(t) \leq \min\{q^\sharp, S\}} \underbrace{p(t)q(t)}_{\text{instantaneous payoff}} + \underbrace{V(t+1, \min\{S^\sharp, S - q(t) + a(t)\})}_{\text{future payoff}}$$

which reveals a trade-off between immediate and future reward

- The Bellman equation follows by definition of the left-hand side

$$V(t,S) = \max_{0 \leq q \leq \min\{q^\sharp, S\}} p(t)q + V(t+1, \min\{S^\sharp, S - q + a(t)\})$$

Outline of the presentation

Discrete-time nonlinear state-control systems are special input-output systems

A specific output is distinguished, and is labeled state,
when the system may be written as

$$x(t+1) = \texttt{Dyn}(t, x(t), u(t)), \quad t \in \mathbb{T} = \{t_0, t_0+1, \dots, T-1\}$$

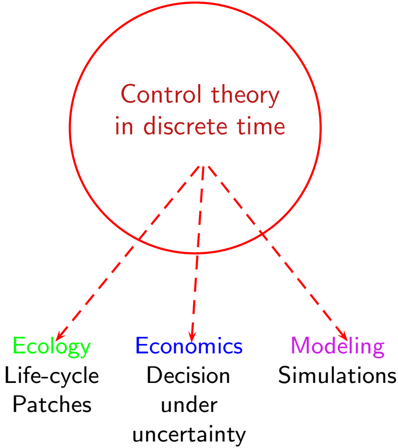
- the time $t \in \overline{\mathbb{T}} = \{t_0, t_0+1, \dots, T-1, T\} \subset \mathbb{N}$ is discrete with initial time t_0 and horizon T ($T < +\infty$ or $T = +\infty$) (*the time period* $[t, t+1[$ *may be a year, a month, etc.*)
- the state variable $x(t)$ belongs to the finite dimensional *state space* $\mathbb{X} = \mathbb{R}^{n_{\mathbb{X}}}$; (*stocks, biomasses, abundances, capital, etc.*)
- the control variable $u(t)$ is an element of the *control space* $\mathbb{U} = \mathbb{R}^{n_{\mathbb{U}}}$ (*outflows, catches, harvesting effort, investment, etc.*)
- the dynamics $\texttt{Dyn}$ maps $\mathbb{T} \times \mathbb{X} \times \mathbb{U}$ into $\mathbb{X}$ (*storage, age-class model, population dynamics, economic model, etc.*)

A historical snapshot on the distinction between states and controls

The Maximum Principle of optimal control: A history of ingenious ideas and missed opportunities, by Hans Josef Pesch and Michael Plail,
Control and Cybernetics, vol. 38 (2009) No. 4A

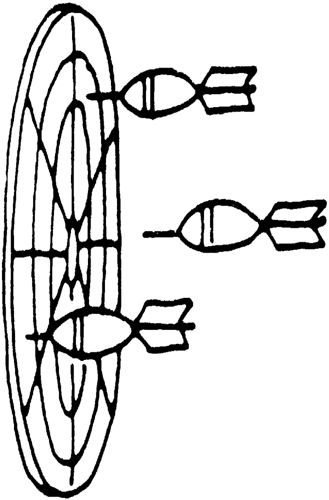
- Lawrence M. Graves (1932, 1933) distinguished the state variables and the degrees of freedom by different letters
- Buried in RAND reports (1949, 1950) Magnus R. Hestenes has definitely introduced different notations for the state and the control variables
- RAND (Research ANd Development) corporation:
  Magnus R. Hestenes, Rufus P. Isaacs, Richard E. Bellman
- Later, Rudolf E. Kálmán as well introduced
  the concept of state and control variables;
- The letter u stands for the Russian word for control: *upravlenie*
- Russian school: Pontryagin, Gamkrelidze, Boltyanskii

We dress natural resources management issues in the formal clothes of control theory in discrete time



- Problems are framed as
  - find controls/decisions driving a dynamical system
  - to achieve various goals
- Three main ingredients are
  - controlled dynamics
  - constraints
  - criterion to optimize

We mathematically express the objectives pursued as control and state constraints



- For a state-control system,
  we cloth objectives as constraints
- and we distinguish
  control constraints (rather easy)
  state constraints (rather difficult)
- Viability theory deals with state constraints

Constraints may be explicit on the control variable

and are rather easily handled by reducing the decision set

Examples of control constraints

- Irreversibility constraints, physical bounds 
  $0 \leq a(t) \leq 1$, $0 \leq h(t) \leq B(t)$
- Tolerable costs $c\big(a(t), Q(t)\big) \leq c^{\sharp}$

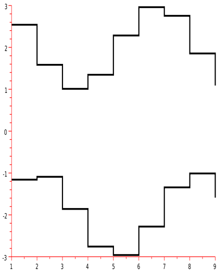
Control constraints / admissible decisions

$$\underbrace{u(t)}_{\text{control}} \in \underbrace{\mathbb{B}\big(t, x(t)\big)}_{\text{admissible set}} , \quad t = t_0, \ldots, T-1$$

Easy because control variables $u(t)$ are precisely those variables whose values the decision-maker can fix at any time within given bounds

Meeting constraints bearing on the state variable is delicate

due to the dynamics pipeline between controls and state



State constraints / admissible states

$$\underbrace{x(t)}_{\text{state}} \in \underbrace{\mathbb{A}(t)}_{\text{admissible set}} \quad , \quad t = t_0, \ldots, T$$

Examples ("tipping points")

- CO_2 concentration $M(t) \leq M^\sharp$
- biomass $B^\flat \leq B(t) \leq B^\sharp$

State constraints are mathematically difficult because of "inertia"

$$x(t) = \underbrace{\text{function}}_{\text{iterated dynamics}} \Big(\underbrace{u(t-1), \ldots, u(t_0)}_{\text{past controls}}, x(t_0) \Big)$$

Target and asymptotic state constraints are special cases

- Final state achieves some target

$$\underbrace{x(T)}_{\text{final state}} \in \underbrace{\mathbb{A}(T)}_{\text{target set}}$$

Example: CO_2 concentration

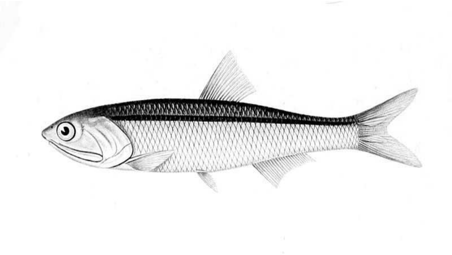
- State converges toward a target

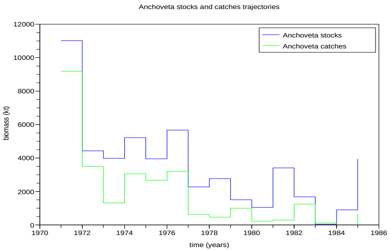
$$\underbrace{\lim_{t\to+\infty} x(t)}_{\text{asymptotic state}} \in \underbrace{\mathbb{A}(\infty)}_{\text{target set}}$$

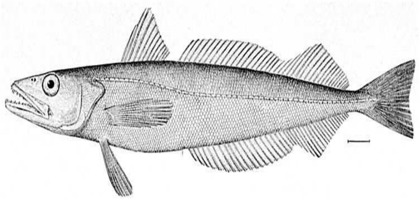
Example: convergence towards an endemic state in epidemiology

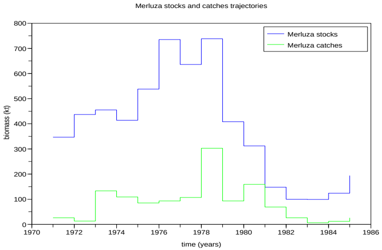
Outline of the presentation

Anchoveta and merluza stock and catches trajectories, in Perú from 1971 to 1985









Trajectories are time sequences (of states and controls), also called paths

- Control trajectory

$$u(\cdot) = \underbrace{\big(u(t_0), u(t_0+1), \ldots, u(T-1)\big)}_{\text{control path}}$$

- State trajectory

$$x(\cdot) = \underbrace{\big(x(t_0), x(t_0+1), \ldots, x(T-1), x(T)\big)}_{\text{state path}}$$

- State-control trajectory

$$\big(x(\cdot), u(\cdot)\big) = \underbrace{\big(x(t_0), \ldots, x(T), u(t_0), \ldots, u(T-1)\big)}_{\text{state-control path}}$$

IMARPE data from 1971 to 1985 in thousands of tonnes (10^3 tons)

- anchoveta stocks [11019 4432 3982 5220 3954 5667 2272 2770 1506 1044 3407 1678 40 900 3944]
- merluza stocks [347 437 455 414 538 735 636 738 408 312 148 100 99 124 194]
- anchoveta captures [9184 3493 1313 3053 2673 3211 626 464 1000 223 288 1240 118 2 648]
- merluza captures [26 13 133 109 85 93 107 303 93 159 69 26 6 12 26]

A history is a whole path of states and controls, and the history set is the natural domain for an intertemporal optimization problem

- A state-control trajectory is called a history

$$\big(x(\cdot), u(\cdot)\big) = \underbrace{\big(x(t_0), \ldots, x(T), u(t_0), \ldots, u(T-1)\big)}_{\text{history}}$$

- The set of state and control trajectories is the so-called history set

$$\underbrace{\big(x(\cdot), u(\cdot)\big)}_{\text{history}} \in \underbrace{\mathbb{X}^{T+1-t_0} \times \mathbb{U}^{T-t_0}}_{\text{history set}}$$

Single dam histories

$$\big(S(\cdot), q(\cdot)\big) = \Big(\overbrace{S(t_0), \ldots, S(T)}^{\text{stocks}}, \overbrace{q(t_0), \ldots, q(T-1)}^{\text{turbined}}\Big)$$

State and control constraints reduce the set of admissible trajectories to account for feasibility issues

Admissible trajectories $\big(x(\cdot), u(\cdot)\big)$ in $\mathcal{T}^{\mathrm{ad}}(t_0, x_0)$ satisfy

- dynamics $x(t+1) = \mathtt{Dyn}\big(t, x(t), u(t)\big)$
- control constraints $u(t) \in \mathbb{B}\big(t, x(t)\big)$
- state constraints $x(t) \in \mathbb{A}(t)$

$$\mathcal{T}^{\mathrm{ad}}(t_0, x_0) = \left\{ \big(x(\cdot), u(\cdot)\big) \;\middle|\; \begin{array}{rll} x(t_0) = & x_0 \ , & \\ x(t+1) = & \mathtt{Dyn}\big(t, x(t), u(t)\big) \ , & t \in \mathbb{T} \\ u(t) \in & \mathbb{B}\big(t, x(t)\big) \ , & t \in \mathbb{T} \\ x(t) \in & \mathbb{A}(t) \ , & t \in \overline{\mathbb{T}} \end{array} \right\}$$

Admissible trajectories for a single dam dynamical model

$$\mathcal{T}^{\mathrm{ad}}(t_0, S_0) = \left\{ \big(S(\cdot), q(\cdot)\big) \;\middle|\; \begin{array}{rll} S(t_0) = & S_0 \ , & \\ S(t+1) = & \min\{S^\sharp, S(t) - q(t) + a(t)\} \ , & t \in \mathbb{T} \\ q(t) \in & [0, \min\{q^\sharp, S(t)\}] & t \in \mathbb{T} \end{array} \right\}$$

An intertemporal criterion assigns a value to each history

Intertemporal criterion

An (intertemporal) criterion or (intertemporal) objective function

$$\texttt{Crit}\Big(x(t_0), x(t_0+1), \ldots, x(T-1), x(T), u(t_0), u(t_0+1), \ldots, u(T-1)\Big)$$

is a function defined over the set of histories

$$\begin{aligned} \texttt{Crit} :\quad & \mathbb{X}^{T+1-t_0} \times \mathbb{U}^{T-t_0} && \to \mathbb{R} \\ & \big(x(\cdot), u(\cdot)\big) && \mapsto \texttt{Crit}\big(x(\cdot), u(\cdot)\big) \end{aligned}$$

Intertemporal payoff for a single dam

$$\texttt{Crit}\big(S(\cdot), q(\cdot)\big) = \sum_{t=t_0}^{T-1} \overbrace{\underbrace{p(t)}_{\text{price}} \times \underbrace{q(t)}_{\text{quantity}}}^{\text{turbined water profit}} + \overbrace{\texttt{K}\big(S(T)\big)}^{\text{final stock utility}}$$

A criterion reflects the intertemporal preferences of the decision-maker (impatience, intergenerational equity, etc.)

- The additive and time-separable criterion

$$\mathtt{Crit}\big(x(\cdot),u(\cdot)\big) = \sum_{t=t_0}^{T-1} \underbrace{\mathtt{L}\big(t,x(t),u(t)\big)}_{\text{instantaneous gain}} + \underbrace{\mathtt{K}\big(x(T)\big)}_{\text{final gain}}$$

  is the most common and covers many well-known examples
  - Discounted present value (or net present value) $\sum_{t=t_0}^{T-1} \delta^{t-t_0} \mathtt{L}\big(x(t),u(t)\big)$
  - Green Golden $\mathtt{K}\big(T,x(T)\big)$
  - Chichilnisky $\theta \sum_{t=t_0}^{T-1} \delta^{t-t_0} \mathtt{L}\big(x(t),u(t)\big) + (1-\theta)\mathtt{K}\big(T,x(T)\big)$
- The Maximin or Rawls criterion

$$\mathtt{Crit}\big(x(\cdot),u(\cdot)\big) = \min_{t=t_0,\dots,T-1} \mathtt{L}\big(t,x(t),u(t)\big)$$

The most common additive and time-separable criterion allows for compensations between time periods

- The most usual criterion is additive and time-separable

$$\texttt{Crit}\big(x(\cdot), u(\cdot)\big) = \sum_{t=t_0}^{T-1} \texttt{L}\big(t, x(t), u(t)\big) + \texttt{K}\big(x(T)\big)$$

- Additive criteria allow for possible compensations between time periods (like the sums of times spent on a graph)
- Environmental economists sanction the present value

$$\texttt{Crit}\big(x(\cdot), u(\cdot)\big) = \overbrace{\sum_{t=t_0}^{+\infty} (\frac{1}{1+r_e})^{t-t_0} \texttt{L}\big(x(t), u(t)\big)}^{\text{discounted utility}}$$

as "dictatorship of the present" (because of discounting)

Discounting erases the future

The French public discount rate

En France, le rapport *Révision du taux d'actualisation des investissements publics* (Commissariat général du Plan, groupe d'experts présidé par Daniel Lebègue, janvier 2005) a conduit à diviser par deux (de 8% à 4%) le taux d'actualisation à retenir pour évaluer la rentabilité des choix d'investissements publics

$$\frac{1}{1+r_e} = \frac{1}{1+0.04} \approx 0.96$$

The future in one hundred years is valued, seen from today, 2%

$$(\frac{1}{1+0.04})^{10} \approx 0.68 \; , \quad (\frac{1}{1+0.04})^{50} \approx 0.14 \; , \quad (\frac{1}{1+0.04})^{100} \approx 0.02$$

The Maximin focuses on minimal utility over time

- Equity: a focus on the poorest generation / utility level of the least advantaged generation
- The maximin form in the finite horizon case

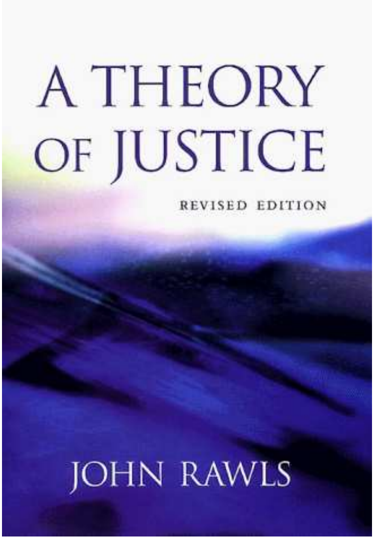
$$\texttt{Crit}\big(x(\cdot), u(\cdot)\big) = \underbrace{\min_{t=t_0,\ldots,T-1} \overbrace{\texttt{L}\big(t, x(t), u(t)\big)}^{\text{generation utility}}}_{\text{worse generation utility}}$$

- In the infinite horizon case

$$\texttt{Crit}\big(x(\cdot), u(\cdot)\big) = \min_{t=t_0,\ldots,+\infty} \texttt{L}\big(t, x(t), u(t)\big)$$

- There can be no compensations between time periods
- John Rawls, *A Theory of Justice*, 1971

John Bordley Rawls (1921–2002)



- John Bordley Rawls
  was an American philosopher and a leading figure in moral and political philosophy, famous for having written
  A Theory of Justice (1971)
- Two of John Rawls's younger brothers died as children —
  from illnesses they contracted from him
- Rawls believed he developed
  his life-long stutter as a result of guilt over his brothers' deaths

The Green Golden criterion is a "dictatorship of the future"

- In the finite horizon case

$$\mathtt{Crit}\big(x(\cdot), u(\cdot)\big) = \mathtt{K}\big(\underbrace{x(T)}_{\text{state}}\big)$$

- In the infinite horizon case

$$\mathtt{Crit}\big(x(\cdot), u(\cdot)\big) = \liminf_{T \to +\infty} \mathtt{K}\big(x(T)\big)$$

- The Green Golden criterion values only the final state and none of the controls (no consumption)

The Chichilnisky criterion is in-between

- The Chichilnisky form with ponderation parameter $\theta \in [0,1]$

$$\texttt{Crit}\big(x(\cdot),u(\cdot)\big) = \theta \underbrace{\sum_{t=t_0}^{T-1} \Big(\frac{1}{1+r_e}\Big)^{t-t_0} \texttt{L}\big(x(t),u(t)\big)}_{\text{dictatorship of the present}} + (1-\theta) \underbrace{\texttt{K}\big(x(T)\big)}_{\text{dictatorship of the future}}$$

- Sustainability: to reconcile $\begin{cases} \text{present} \\ \text{future} \end{cases}$
- In the infinite horizon case

$$\texttt{Crit}\big(x(\cdot),u(\cdot)\big) = \theta \sum_{t=t_0}^{+\infty} \texttt{L}\big(t,x(t),u(t)\big) + (1-\theta) \liminf_{T\to+\infty} \texttt{K}\big(x(T)\big)$$

Outline of the presentation

We are now ready to dress an intertemporal optimization problem in formal clothes

- Domain: the product set of histories

$$\mathbb{X}^{T+1-t_0} \times \mathbb{U}^{T-t_0}$$

- Criterion/objective function:

$$\begin{aligned} \texttt{Crit} : \quad \mathbb{X}^{T+1-t_0} \times \mathbb{U}^{T-t_0} &\to \mathbb{R} \\ \big(x(\cdot), u(\cdot)\big) &\mapsto \texttt{Crit}\big(x(\cdot), u(\cdot)\big) \end{aligned}$$

- Constraints: induced by state, control and dynamics constraints

$$\mathcal{T}^{\mathrm{ad}}(t_0, x_0) = \left\{ \big(x(\cdot), u(\cdot)\big) \;\middle|\; \begin{array}{rll} x(t_0) = & x_0 \,, & \\ x(t+1) = & \texttt{Dyn}\big(t, x(t), u(t)\big) \,, & t \in \mathbb{T} \\ u(t) \in & \mathbb{B}\big(t, x(t)\big) \,, & t \in \mathbb{T} \\ x(t) \in & \mathbb{A}(t) \,, & t \in \overline{\mathbb{T}} \end{array} \right\}$$

An optimal trajectory maximizes the criterion over all admissible trajectories

- The optimal value

$$\texttt{Crit}^\star(t_0, x_0) = \max_{\big(x(\cdot),u(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}(t_0,x_0)} \texttt{Crit}\big(x(\cdot), u(\cdot)\big)$$

is abusively denoted by $\texttt{Crit}^\star(t_0, x_0) = \max_{u(\cdot)} \texttt{Crit}\big(x(\cdot), u(\cdot)\big)$

Optimal trajectory

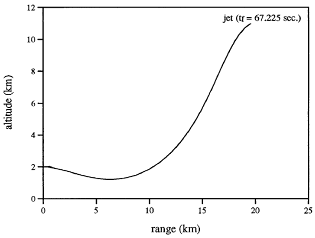
Any path $\big(x^\star(\cdot), u^\star(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}(t_0, x_0)$ such that

$$\max_{\big(x(\cdot),u(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}(t_0,x_0)} \texttt{Crit}\big(x(\cdot), u(\cdot)\big) = \texttt{Crit}\big(x^\star(\cdot), u^\star(\cdot)\big)$$

is a feasible optimal trajectory or an optimal path

Why *mathematical* optimization? (and not rules of thumb)

- Optimization can examine millions, or even an infinite number, of choices to find the best ones within the specified resource limits
- It can look at many more options than any human decision maker (using rules of thumb)



Optimum trajectory for a jet engine

- In 1962, Bryson and Denham calculated the optimal trajectory for a supersonic plane (F4H) to go from zero altitude to the altitude of 20 km in horizontal flight with Mach 1 speed
- They showed the counterintuitive result that the path of a supersonic aircraft should actually dive at one point
- The minimal time obtained by optimal control is, depending on sources, 10% (to 50%) less than the time needed by the best of pilots!
- The optimum is by no ways obvious, and requires mathematical techniques to be characterized

Summary

- Discrete-time nonlinear state-control systems
  are special input-output dynamical systems
  - control = input
  - state = specific output satisfying a dynamical equation
- Trajectories are time sequences (of states and controls), also called paths
- State and control constraints reduce the set of admissible trajectories
  to account for feasibility issues
- A history is a whole path of states and controls,
  and an intertemporal criterion assigns a value to each history
- A criterion reflects the intertemporal preferences of the decision-maker
  (impatience, intergenerational equity, etc.)
- An optimal trajectory maximizes the criterion over all admissible trajectories

Outline of the presentation

Outline of the presentation

In Pontryaguin's approach, dynamic equations are dualized, and a derivative is set to zero

Pontryaguin was incited to work on optimal control to beat the U-2 plane

- Treat dynamic equations as constraints bearing on trajectories
- Raise the constraints into the criterion by means of Lagrange multipliers and forge the Hamiltonian
- Set the derivative of the Hamiltonian to zero

In *De Methodis Serierum et Fluxionum*
(A Treatise on the Methods of Series and Fluxions) (1671), Newton stated

> *When a quantity is greatest or least, at that moment its flow neither increases nor decreases: for if it increases, that proves that it was less and will at once be greater than it now is, and conversely so if it decreases. Therefore seek its fluxion [by previously described methods] and set it equal to nothing.*

For an additive criterion, we define a Hamiltonian

- To the optimization problem

$$\max_{\left(x(\cdot),u(\cdot)\right)} \sum_{t=t_0}^{T-1} \underbrace{\mathtt{L}\big(t,x(t),u(t)\big)}_{\text{instantaneous gain}} + \underbrace{\mathtt{K}\big(x(T)\big)}_{\text{final gain}}$$

- with dynamics $x(t+1) = \mathtt{Dyn}\big(t,x(t),u(t)\big) \quad , \quad x(t) \in \mathbb{X} = \mathbb{R}^n$
- we attach the Hamiltonian function

$$\begin{aligned}\mathcal{H}(t,x,q,u) &= \sum_{i=1}^{n} q_i \mathtt{Dyn}_i(t,x,u) + \mathtt{L}(t,x,u) \\ &= \underbrace{\mathtt{Dyn}(t,x,u)'}_{\text{dynamic}} q + \underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}}\end{aligned}$$

where the new variable $q \in \mathbb{R}^n$ is called
adjoint state, adjoint variable, or multiplier

Pontryaguin's "maximum" principle

Assume that instantaneous utility L, final utility K and dynamic Dyn are continuously differentiable in the state and control variables (x, u)

- If the trajectory $\big(x^\star(\cdot), u^\star(\cdot)\big)$ is optimal,
- there exists a sequence $q^\star(\cdot) = (q^\star(t_0), \ldots, q^\star(T-1)) \in \mathbb{X}^{T-t_0}$ of adjoint states
- such that, for any $i = 1, \ldots, n$ and $j = 1, \ldots, p$, we have

Hamilton equations

$$
\begin{cases}
x_i^\star(t+1) = \dfrac{\partial \mathcal{H}}{\partial q_i}\big(t, x^\star(t), q^\star(t), u^\star(t)\big) , \quad t = t_0, \ldots, T-1 \\
q_i^\star(t-1) = \dfrac{\partial \mathcal{H}}{\partial x_i}\big(t, x^\star(t), q^\star(t), u^\star(t)\big) , \quad t = t_0+1, \ldots, T-1 \\
0 = \dfrac{\partial \mathcal{H}}{\partial u_j}\big(t, x^\star(t), q^\star(t), u^\star(t)\big) , \quad t = t_0, \ldots, T-1 \\
x_i^\star(t_0) = x_{i0} \\
q_i^\star(T-1) = \dfrac{\partial \mathrm{K}}{\partial x_i}\big(x^\star(T)\big)
\end{cases}
$$

Beware! The optimum may not be a maximum (in discrete-time)

Outline of the presentation

Harold Hotelling, 1895-1973

- Harold Hotelling
  was a mathematical statistician
  and an influential economic theorist
- His name is known to all statisticians
  because of Hotelling's T-square
  distribution and its use in statistical
  hypothesis testing
  and confidence regions
- He also introduced
  canonical correlation analysis,
  and is the eponym of
  Hotelling's law, Hotelling's lemma,
  and Hotelling's rule in economics

Wikipedia

Optimal extraction of an exhaustible resource

H. Hotelling. The economics of exhaustible resources.
Journal of Political Economy, 39:137–175, april 1931.

- For the dynamic

$$S(t+1) = \underbrace{S(t)}_{\text{stock}} - \underbrace{h(t)}_{\text{extraction}} \quad , \quad 0 \leq h(t) \leq S(t)$$

- and the optimal discounted utility problem

$$\max_{h(t_0),\ldots,h(T-1)} \left(\sum_{t=t_0}^{T-1} \delta^{t-t_0} \underbrace{\mathtt{L}(h(t))}_{\text{extraction utility}} + \delta^{T-t_0} \underbrace{\mathtt{L}(S(T))}_{\text{final stock utility}} \right)$$

- the Hamiltonian is

$$\mathcal{H}(t,S,q,h) = q \underbrace{(S-h)}_{\text{dynamics}} + \underbrace{\delta^{t-t_0}\mathtt{L}(h)}_{\text{instantaneous gain}}$$

The Hamilton equations yield a stationary multiplier

$$
\begin{aligned}
S^\star(t+1) &= S^\star(t) - h^\star(t) & \left(\frac{\partial \mathcal{H}}{\partial q} = S - h\right) \\
q^\star(t-1) &= q^\star(t) & \left(\frac{\partial \mathcal{H}}{\partial S} = q\right) \\
0 &= \delta^{t-t_0} \mathrm{L}'\big(h^\star(t)\big) - q^\star(t) & \left(\frac{\partial \mathcal{H}}{\partial h} = \delta^{t-t_0} \mathrm{L}'\big(h\big) - q\right) \\
S^\star(t_0) &= S_0 & \\
q^\star(T-1) &= \delta^{T-t_0} \mathrm{L}'\big(S^\star(T)\big) & \left(\frac{\partial \mathrm{K}}{\partial S} = \delta^{T-t_0} \mathrm{L}'\big(S\big)\right)
\end{aligned}
$$

The second equation shows that the multiplier $q^\star(t)$ is stationary and the third equation gives

$$
\mathrm{L}'\big(h^\star(t)\big) = q^\star(t_0)\delta^{t_0-t}
$$

The Hotelling rule states that the price of an exhaustible resource grows at the discount rate

- Economists interpret the marginal utility $\mathtt{L}'\big(h^{\star}(t)\big)$ as the price of the stock at time t, that we denote by

$$p(t) = \mathtt{L}'\big(h^{\star}(t)\big)$$

- Since the multiplier $q^{\star}(t)$ is stationary, we deduce that the price of the resource grows as

$$p(t+1) = \delta^{-1} p(t) = (1 + r_e) p(t)$$

- The so-called Hotelling rule states that the exhaustible resource price growth rate coincides with the discount rate r_e

$$\frac{p(t+1)}{p(t)} = 1 + \underbrace{r_e}_{\text{discount rate}}$$

The Hotelling rule has found an application to remedy the stubborn use of discounting with constant prices

M. Boiteux, À propos de la "critique de la théorie de l'actualisation telle qu'employée en France", Revue d'Économie Politique, 1976.

- "l'application obtuse de l'actualisation,
  à prix constants et sur les seules valeurs marchandes,
  trahit les réalités et les aspirations profondes de nos sociétés"
- "la procédure de l'actualisation nettoie, à terme, ce qui est accessoire parce que maîtrisable par le génie humain, pour mettre en relief l'essentiel : ce qui est intrinsinquement rare et non reproductible"

En France, la valeur tutélaire du carbone est fixée à l'aide de la règle de Hotelling

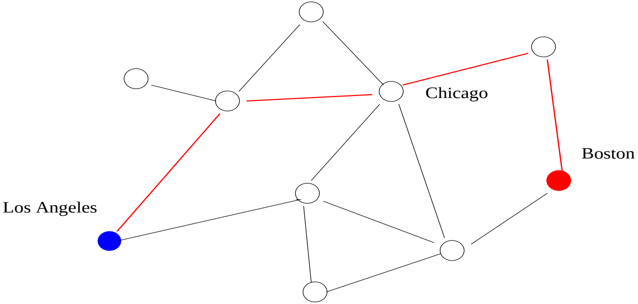
La valeur tutélaire du carbone, Commission du Centre d'analyse stratégique, présidée par Alain Quinet, 2009.

- Valeur fixée à 100 euros par tonne de CO2 à l'horizon 2030.
- Après 2030, cette valeur de 100 euros croît au rythme du taux d'actualisation public. Cette règle d'évolution, similaire à la règle de Hotelling pour l'exploitation optimale des ressources épuisables, est une règle de préservation de l'avenir. Elle garantit que le prix actualisé d'une ressource limitée reste constant au cours du temps et n'est pas "écrasé" par l'actualisation.
- Il est retenu un taux de croissance annuel de la valeur carbone de 4%. Avec ces hypothèses, la valeur du carbone croît de 100 euros la tonne de CO2 en 2030 à 200 euros en 2050.

Outline of the presentation

Outline of the presentation

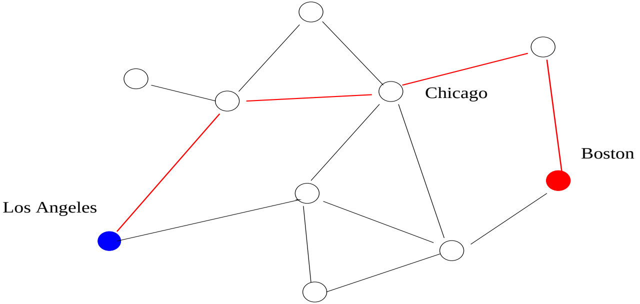
We consider a shortest distance travelling problem



Shortest time path to go from Los Angeles to Boston

- Time: day
- States: nodes of a graph (cities)
- Admissible controls: arcs starting from a node
- Criterion: sum of the arcs lengths
- Optimization problem: minimal distance to go from Los Angeles to Boston

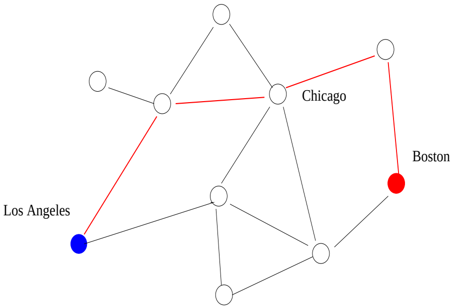
We illustrate two notions of solutions to the shortest time travelling problem



- Open loop: a sequence of cities
  Los Angeles → Las Vegas → Salt Lake City → ⋯ → Chicago → ⋯
  Cleveland → Boston
- Closed loop: a sequence of decision rules
  for *any* city (Los Angeles, San Diego, Dallas, Chicago, etc.),
  pinpoint the next city to visit

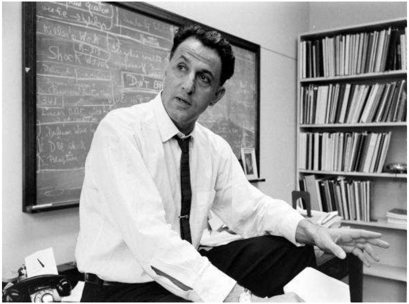
Shortest path to go to Boston

The shortest path on a graph illustrates Bellman's Principle of Optimality



For an auto travel analogy, suppose that the fastest route from Los Angeles to Boston passes through **Chicago**. *The principle of optimality translates to obvious fact that the Chicago to Boston portion of the route is also the fastest route for a trip that starts from Chicago and ends in Boston. (Dimitri P. Bertsekas)*

Bellman's Principle of Optimality



Richard Ernest Bellman
(August 26, 1920 – March 19, 1984)

> *An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision (Richard Bellman)*

A plan is time consistent if the passage of time alone gives no reason to change it (G. Heal)

Optimization of an intertemporal additive criterion

- We consider the optimization of an intertemporal additive criterion

$$\max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t_0,x_0)} \sum_{t=t_0}^{T-1} \underbrace{\mathtt{L}\big(t,x(t),u(t)\big)}_{\text{instantaneous gain}} + \underbrace{\mathtt{K}\big(x(T)\big)}_{\text{final gain}}$$

- The dynamic programming method breaks
  - an intertemporal optimization problem
  - into smaller static optimization subproblems
- and Richard Bellman's Principle of Optimality describes how to do this

The payoff-to-go / value function / Bellman function

The payoff-to-go from state x at time t is

$$V(t,x) = \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}\boxed{(t,x)}} \sum_{s=\boxed{t}}^{T-1} \mathtt{L}\big(s,x(s),u(s)\big) + \mathtt{K}\big(x(T)\big)$$

that is, the best payoff which can be obtained starting from state x at time t

- The function V is called the value function, or the Bellman function
- The original problem is $V(t_0,x_0)$, that is,
  a single problem among a myriad of others

Shortest path to go from Los Angeles to Boston

- Value function: minimal time to go from any city (node) to Boston
- The original problem is $V(0,$ Los Angeles $)$

The dynamic programming equation, or Bellman equation, is a backward equation satisfied by the value function

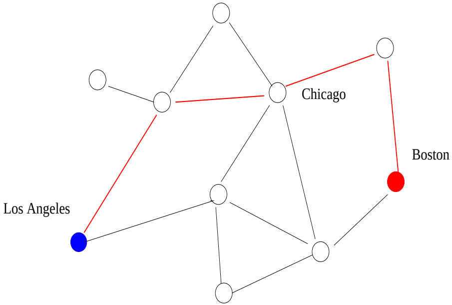
Proposition

In the case without state constraints ($\mathbb{A}(t) = \mathbb{X}$), the value function is the solution of the following backward dynamic programming equation (or Bellman equation) where t runs from $T-1$ down to t_0:

$$
\begin{aligned}
V(T,x) &= \mathtt{K}(x)\,, \quad \forall x \in \mathbb{X} \\
V(t,x) &= \max_{u \in \mathbb{B}(t,x)} \Big(\mathtt{L}(t,x,u) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big) \Big)\,, \quad \forall x \in \mathbb{X}
\end{aligned}
$$

Sketch of the proof

$$V(t,x) = \max_{u \in \mathbb{B}(t,x)} \Big(\underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}} + \overbrace{V\big(t+1, \underbrace{\mathtt{Dyn}(t,x,u)}_{\text{future state}}\big)}^{\text{optimal payoff}} \Big)$$



A decision u at time t in state x provides

- an instantaneous gain $\mathtt{L}(t,x,u)$
- and a future payoff for attaining the new state $\mathtt{Dyn}(t,x,u)$

Shortest path to go from LA to Boston

When you leave a node (city) by an arc

- it takes the travel time on the arc
- and it brings you to another node (city), from where the minimal time is supposed known

Sketch of the proof (2)

$$
\begin{aligned}
\mathcal{T}^{\mathrm{ad}}(t, x_t) &= \left\{ \big(x(\cdot), u(\cdot)\big) \;\middle|\; \begin{array}{rll} x(t) = & x_t \;, & \\ x(s+1) = & \mathtt{Dyn}\big(s, x(s), u(s)\big) \;, & s = t, \ldots, T-1 \\ u(s) \in & \mathbb{B}\big(s, x(s)\big) \;, & s = t, \ldots, T-1 \\ x(s) \in & \mathbb{A}(s) \;, & s = t, \ldots, T \end{array} \right\} \\
&= \left\{ \big(x(\cdot), u(\cdot)\big) \;\middle|\; \begin{array}{l} x(t) = x_t \in \mathbb{A}(t) \;, \\ u(t) \in \mathbb{B}(t, x_t) \;, \\ \big(x(\cdot), u(\cdot)\big) \in \mathcal{T}^{\mathrm{ad}}\Big(t+1, \mathtt{Dyn}\big(t, x_t, u(t)\big)\Big) \end{array} \right\}
\end{aligned}
$$

Sketch of the proof (3)

$$
\begin{aligned}
V(t,x) &= \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t,x)} \sum_{s=t}^{T-1} \mathtt{L}\big(s,x(s),u(s)\big) + \mathtt{K}\big(x(T)\big) \\
&= \max_{u(t)\in\mathbb{B}(t,x_t)} \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}\Big(t+1,\mathtt{Dyn}\big(t,x_t,u(t)\big)\Big)} \\
&\quad \sum_{s=t}^{T-1} \mathtt{L}\big(s,x(s),u(s)\big) + \mathtt{K}\big(x(T)\big) \\
&= \max_{u_t\in\mathbb{B}(t,x_t)} \mathtt{L}\big(t,x(t),u(t)\big) \\
&\quad + \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}\Big(t+1,\mathtt{Dyn}\big(t,x_t,u(t)\big)\Big)} \sum_{s=t+1}^{T-1} \mathtt{L}\big(s,x(s),u(s)\big) + \mathtt{K}\big(x(T)\big) \\
&= \max_{u\in\mathbb{B}(t,x)} \Big(\mathtt{L}(t,x,u) + V\big(t+1,\mathtt{Dyn}(t,x,u)\big)\Big)
\end{aligned}
$$

Building block for the proof: a maximum over a product set is a sequence of maxima

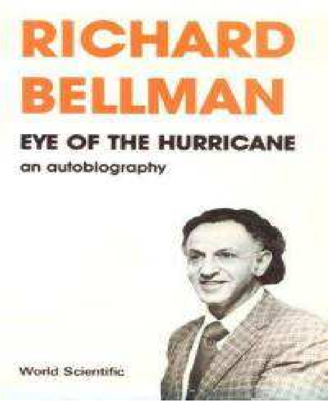
- A maximum over a product set is a sequence of maxima

$$\max_{a \in A, b \in B} f(a,b) = \max_{(a,b) \in A \times B} f(a,b) = \max_{a \in A} \Big(\max_{b \in B} f(a,b) \Big)$$

- Extension to the case when the domain B depends on the "first" decision: $B \to B(a)$

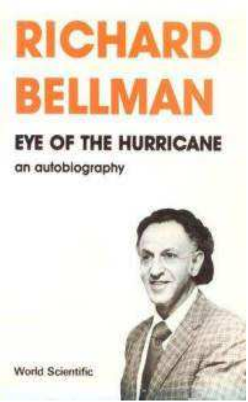
$$\max_{a \in A, b \in B(a)} f(a,b) = \max_{a \in A} \underbrace{\Big(\max_{b \in B(a)} f(a,b) \Big)}_{\hat{f}(a)}$$

“Where did the name, dynamic programming, come from?”



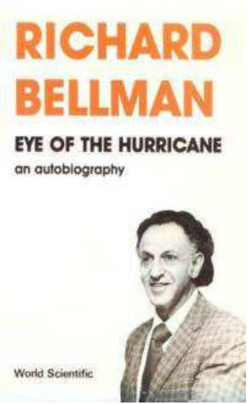
The 1950s were not good years for mathematical research. We had a very interesting gentleman in Washington named Wilson. He was Secretary of Defense, and he actually had a pathological fear and hatred of the word, research. I'm not using the term lightly; I'm using it precisely. His face would suffuse, he would turn red, and he would get violent if people used the term, research, in his presence. You can imagine how he felt, then, about the term, mathematical.

“Where did the name, dynamic programming, come from?”



What title, what name, could I choose? In the first place I was interested in planning, in decision making, in thinking. But planning, is not a good word for various reasons. I decided therefore to use the word, programming.

"Where did the name, dynamic programming, come from?"



I wanted to get across the idea that this was dynamic, this was multistage, this was time-varying. I thought, let's kill two birds with one stone. Let's take a word that has an absolutely precise meaning, namely dynamic, in the classical physical sense. It also has a very interesting property as an adjective, and that is it's impossible to use the word, dynamic, in a pejorative sense. Try thinking of some combination that will possibly give it a pejorative meaning. It's impossible. Thus, I thought dynamic programming was a good name.

Algorithm for the Bellman functions

initialization $V(T,x) = \mathtt{K}(x)$;

for $t = T, T-1, \dots, t_0$ **do**

  forall the $x \in \mathbb{X}$ **do**

    forall the $u \in \mathbb{B}(t,x)$ **do**

      $l\big(t,x,u\big) = \mathtt{L}\big(t,x,u\big) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big)$

    $V(t,x) = \max_{u \in \mathbb{B}(t,x)} l\big(t,x,u\big)$;

    $\mathbb{B}^\star(t,x) = \operatorname{argmax}_{u \in \mathbb{B}(t,x)} l\big(t,x,u\big)$

Case with state constraints

Proposition

$$V(T,x) = \mathtt{K}(x) , \quad \forall x \in \mathbb{V}\mathrm{iab}(T)$$

$$V(t,x) = \max_{u \in \mathbb{B}^{viab}(t,x)} \Big(\mathtt{L}(t,x,u) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big) \Big) , \quad \forall x \in \mathbb{V}\mathrm{iab}(t)$$

where $\mathbb{V}\mathrm{iab}(t)$ *is given by the backward induction*

$$\mathbb{V}\mathrm{iab}(T) = \mathbb{A}(T)$$

$$\mathbb{V}\mathrm{iab}(t) = \{x \in \mathbb{A}(t) \mid \exists u \in \mathbb{B}(t,x) , \quad \mathtt{Dyn}(t,x,u) \in \mathbb{V}\mathrm{iab}(t+1)\}$$

and where the supremum is over viable controls

$$\mathbb{B}^{viab}(t,x) = \{u \in \mathbb{B}(t,x) \mid \mathtt{Dyn}(t,x,u) \in \mathbb{V}\mathrm{iab}(t+1)\}$$

The dynamic programming equation yields a decision rule

- The dynamic programming equation

$$V(t,x) = \max_{u \in \mathbb{B}(t,x)} \Big(\underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}} + \overbrace{V\big(t+1, \underbrace{\mathtt{Dyn}(t,x,u)}_{\text{future state}}\big)}^{\text{optimal payoff}} \Big)$$

- yields a decision rule

$$\mathtt{Pol}^{\star}(t,x) \in \operatorname*{argmax}_{u \in \mathbb{B}(t,x)} \Big(\mathtt{L}(t,x,u) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big) \Big)$$

Again the intriguing thought: A solution is not merely a set of functions of time, or a set of numbers, but a rule telling the decisionmaker what to do; a policy (Richard Bellman)

The Bellman equation provides an optimal policy

Proposition

For any time t and state x, assume the existence of the policy

$$\mathtt{Pol}^{\star}(t,x) \in \operatorname*{argmax}_{u \in \mathbb{B}(t,x)} \Big(\mathtt{L}(t,x,u) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big) \Big)$$

The policy $\mathtt{Pol}^{\star} : (t,x) \mapsto \mathtt{Pol}^{\star}(t,x)$ *is an optimal strategy, in the sense that it yields an optimal trajectory as follows*

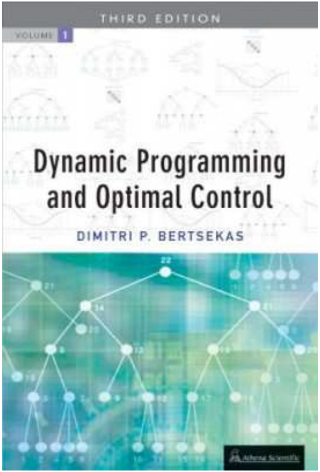
Optimal trajectories are calculated forward online

1. Initial state $x^\star(t_0) = x_0$
2. Plug the state $x^\star(t_0)$ into the "machine" $\texttt{Pol}^\star$ → initial decision $u^\star(t_0) = \texttt{Pol}^\star(t_0, x^\star(t_0))$
3. Run the dynamics → second state $x^\star(t_0+1) = \texttt{Dyn}\big(t_0, x^\star(t_0), u^\star(t_0)\big)$
4. Second decision $u^\star(t_0+1) = \texttt{Pol}^\star(t_0+1, x^\star(t_0+1))$
5. And so on $x^\star(t_0+2) = \texttt{Dyn}\big(t_0+1, x^\star(t_0+1), u^\star(t_0+1)\big)$ ...

Proposition

$$\begin{aligned} & \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t_0,x_0)} \texttt{Crit}\big(x(\cdot),u(\cdot)\big) \quad \textit{optimal payoff} \\ = \quad & V(t_0,x_0) \quad \textit{value function at } (t_0,x_0) \\ = \quad & \texttt{Crit}\big(\underbrace{x^\star(\cdot),u^\star(\cdot)}_{\text{optimal trajectory}}\big) \end{aligned}$$

“Life is lived forward but understood backward” (Søren Kierkegaard)



D. P. Bertsekas introduces his book
Dynamic Programming and Optimal Control
with a citation by Søren Kierkegaard

> *"Livet skal forstås baglaens, men leves forlaens"*
>
> *Life is to be understood backwards,*
> *but it is lived forwards*

- The value function and the optimal policies are computed backward and offline by means of the Bellman equation
- whereas the optimal trajectories are computed forward and online

Optimal strategies anticipate the future

Dam management optimal strategy

$$
\begin{aligned}
V(T,S) &= \overbrace{\mathtt{K}(S)}^{\text{final payoff}} , \\
V(t,S) &= \max_{0 \leq q \leq \min\{S,q^\sharp\}} \underbrace{p(t)q}_{\text{instant. payoff}} + V\big(t+1, \underbrace{\min\{S^\sharp, S-q+a(t)\}}_{\text{future stock volume}}\big)
\end{aligned}
$$

- The value function $S \mapsto V(t,S)$ at time t depends on the tails
  - $\big(a(t), \ldots, a(T-1)\big)$ of the water inflows
  - $\big(p(t), \ldots, p(T-1)\big)$ of the prices
- Therefore, an optimal policy $S \mapsto \mathtt{Pol}^\star(t,S)$ at time t depends on the same remaining future trajectories, hence anticipates the future
- Therefore, an optimal control trajectory $q(\cdot)$ depends on the whole trajectories $a(\cdot)$ and $p(\cdot)$

The adjoint state as a marginal value

Proposition

Suppose that the value function $V(t,x)$ associated to the maximization problem is smooth with respect to the state variable x
Assume that there exists an optimal trajectory $x^\star(\cdot)$ such that the optimal decision $\mathtt{Pol}^\star\big(t, x^\star(t)\big)$ is unique for all $t = t_0, \ldots, T-1$
Then, the sequence $q^\star(\cdot)$ defined by

$$q_i^\star(t) = \frac{\partial V}{\partial x_i}\big(t+1, x^\star(t+1)\big) \; , \quad t = t_0, \ldots, T-1$$

is a solution of the Hamilton equations

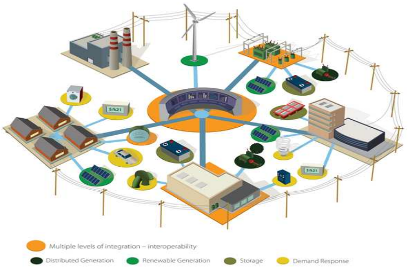
The adjoint state appears as the derivative of the value function with respect to the state variable along an optimal trajectory

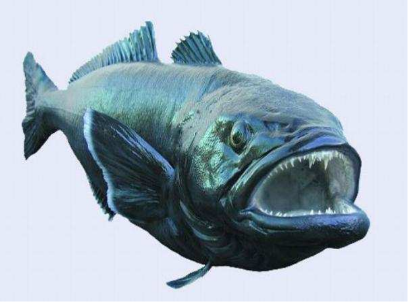
The curse of dimensionality is illustrated by the random access memory capacity on a computer: one, two, three, infinity (Gamov)

- On a computer
  - RAM: 8 GBytes $= 8(1\ 024)^3 = 2^{33}$ bytes
  - a double-precision real: 8 bytes $= 2^3$ bytes
  - $\Longrightarrow 2^{30} \approx 10^9$ double-precision reals can be handled in RAM
- If a state of dimension 4 is approximated
  by a grid with 100 levels by components,
  we need to manipulate $100^4 = 10^8$ reals
  (without even counting the discretization of the controls and the time loop)

Outline of the presentation

The issue of on-line information





- Can we centralize all the informations on stock values in a large power system?
- Can we measure on-line the abundances of all age-classes in a population model?
- What about measurement errors?

When decisions do not take into account on-line information and the clock time, we are in the stationary static case

Stationary (open-loop)

Stationary open-loop control is

$$u : \underbrace{t \in \mathbb{T}}_{\text{time}} \mapsto \underbrace{u(t) \equiv u_{\mathrm{E}} \in \mathbb{U}}_{\text{control}}$$

Harvest the same biomass every year, as in the maximum sustainable yield

When decisions do not take into account on-line information but depend on the clock, we are in the open-loop case

Open-loop ⏲

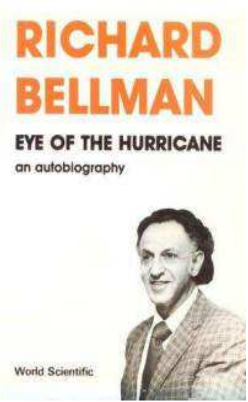
Open-loop control consists of time-dependent sequences (planning, scheduling)

$$u : \underbrace{t \in \mathbb{T}}_{\text{time}} \mapsto \underbrace{u(t) \in \mathbb{U}}_{\text{control}}$$

Examples of open-loop control

- Fixed cycle gears for traffic lights in traffic regulation
- Mine planning: extract a given sequence of blocks every year, whatever you learn of the metal prices or of the ore content
- Solutions to optimal control problems by Pontryagin's variational approach

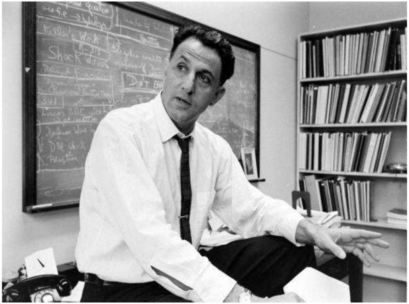
"I started work on control theory" (Richard Bellman)



The tool we used was the calculus of variations. What we found was that very simple problems required great ingenuity. A small change in the problem caused a great change in the solution.

Clearly, something was wrong. There was an obvious lack of balance. Reluctantly, I was forced to the conclusion that the calculus of variations was not an effective tool for obtaining a solution

"The thought was finally forced upon me that the desired solution in a control process was a policy" (Richard Bellman)



Richard Ernest Bellman
(August 26, 1920 – March 19, 1984)

From planning ⌚
to contingent planning ⌚× 👁

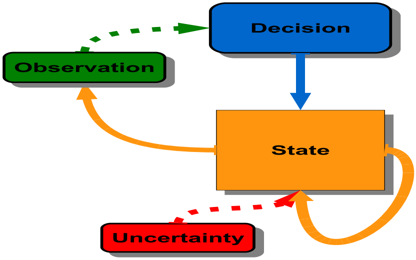
> *Again the intriguing thought: A solution is not merely a set of functions of time, or a set of numbers, but a rule telling the decisionmaker what to do; a policy (Richard Bellman)*

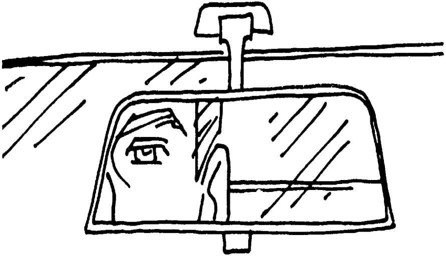
A computer code aboard a launcher embodies the concept of policy

```
if     state==0,
       do control=8
elseif state==1,
       do control=5.4
else do control=-15
```

On 4 June 1996, the maiden flight of the Ariane 5 launcher ended in a failure. (...) The attitude of the launcher and its movements in space are measured by an Inertial Reference System (SRI). (...) The data from the SRI are transmitted through the databus to the On-Board Computer (OBC), which executes the flight program (...) The Operand Error occurred due to an unexpected high value of an internal alignment function result called BH, Horizontal Bias, related to the horizontal velocity sensed by the platform

“The blind cat does not catch mice”

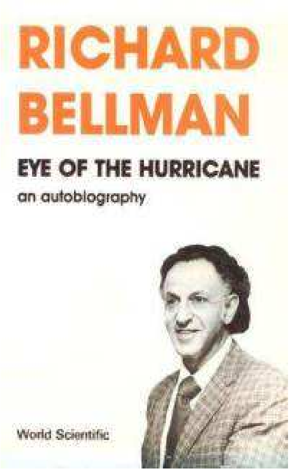




$$\underbrace{u(t)}_{\text{control}} = \texttt{Pol}\big(t, \underbrace{y(t)}_{\text{output}}\big)$$

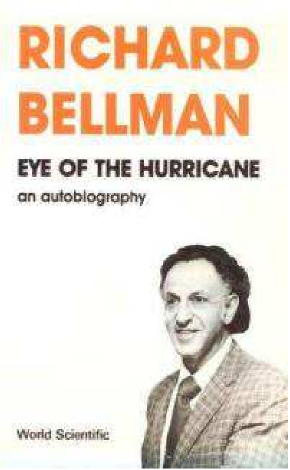
- adaptive
- adjustable
- feeback
- wait and see
- full recourse
- on-line management
- corrective (vs. preventive)

How clouded the crystal ball looks beforehand



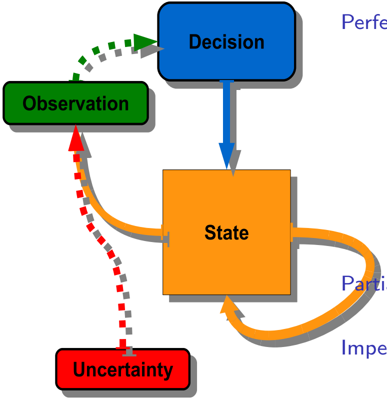
What is worth noting about the foregoing development is that I should have seen the application of dynamic programming to control theory several years before. I should have, but I didn't. It is very well to start a lecture by saying, 'Clearly, a control process can be regarded as a multistage decision process in which. . . ,' but it is a bit misleading.

How clouded the crystal ball looks beforehand



Scientific developments can always be made logical and rational with sufficient hindsight. It is amazing, however, how clouded the crystal ball looks beforehand. We all wear such intellectual blinders and make such inexplicable blunders that it is amazing that any progress is made at all.

There are different observation patterns



Perfect observation:

- Decision-hazard

$$y(t) = x(t)$$

- Hazard-decision

$$y(t) = \big(x(t), w(t)\big)$$

Partial observation:

$$y(t) = \texttt{Obs}\big(t, x(t)\big)$$

Imperfect observation:

$$y(t) = \texttt{Obs}\big(t, x(t), w(t)\big)$$

Dams management

Observing the stocks of all dams / the stocks and the water inflows / or only some stocks

State feedback policies correspond to perfect observation of the state

'Do thus-and-thus if you find yourself in this portion of state space with this amount of time left' (Richard Bellman)

Closed-loop control, state feedback (decision rule)

$$\mathtt{Pol} : \underbrace{(t, x) \in \mathbb{T} \times \mathbb{X}}_{\text{(time, state)}} \mapsto u = \underbrace{\mathtt{Pol}(t, x) \in \mathbb{U}}_{\text{control}}$$

Turbinate a fraction of the dam stock

$$\mathtt{Pol}(t, S) = \alpha(t) S \text{ with } 0 \leq \alpha(t) \leq 1$$

ICES precautionary approach

$$\lambda_{UA}(N) = \max\{\lambda \in \mathbb{R}_+ \mid \mathtt{SSB}(\mathtt{Dyn}(N, \lambda)) \geq B_{\text{lim}} \text{ and } F(\lambda) \leq F_{\text{lim}}\}$$

Going from planning to contingent planning, we have considerably enlarged the set of solutions

- Stationary (open-loop): stationary sequences

$$u : t \in \mathbb{T} \mapsto u(t) \equiv u_{\mathrm{E}} \;, \quad u_{\mathrm{E}} \in \mathbb{U}$$

Once the control space $\mathbb{U}$ is discretized in $N_{\mathbb{U}}$ elements,
the solution space cardinality is $N_{\mathbb{U}}$

- Open-loop: time-dependent sequences (planning, scheduling)

$$u : t \in \mathbb{T} \mapsto u(t) \;, \quad u(\cdot) = \big(u(t_0), \ldots, u(T-1)\big) \in \mathbb{U}^{\mathbb{T}}$$

With $N_{\mathbb{T}}$ time periods, the solution space cardinality is $N_{\mathbb{U}}^{N_{\mathbb{T}}}$

- Closed-loop: time and state-dependent sequences

$$\mathtt{Pol} : (t, x) \in \mathbb{T} \times \mathbb{X} \mapsto u = \mathtt{Pol}(t, x) \in \mathbb{U} \;, \quad \mathtt{Pol} \in \mathbb{U}^{\mathbb{T} \times \mathbb{X}}$$

Once the state space $\mathbb{X}$ is discretized in $N_{\mathbb{X}}$ elements,
the solution space cardinality is $N_{\mathbb{U}}^{N_{\mathbb{T}} \times N_{\mathbb{X}}}$

Summary

- Bellman's Principle of Optimality
  breaks an intertemporal optimization problem
  into a sequence of interconnected static optimization problems
- The payoff-to-go / value function / Bellman function
  is solution of a backward dynamic programming equation,
  or Bellman equation
- The Bellman equation provides an optimal policy,
  a concept of solution which will prove useful in the uncertain case
- In numerical practice, the curse of dimensionality
  forbids to use dynamic programming
  for a state with dimension more than three or four

Outline of the presentation

Outline of the presentation

Two stage optimal management of an exhaustible resource

- Time t runs from $t_0 = 0$ to $T = 2$, that is, $t = 0, 1, 2$
- We consider an exhaustible resource

$$S(t+1) = \underbrace{S(t)}_{\text{stock}} - \underbrace{h(t)}_{\text{extraction}} \ , \quad 0 \leq h(t) \leq S(t) \ , \quad t = 0, 1$$

- where one maximizes intertemporal discounted utility

$$\max_{h(0), h(1)} \left(\underbrace{\big(h(0)\big)^{\gamma}}_{\text{extraction utility}} + \overbrace{\delta}^{\text{discount factor}} \underbrace{\big(h(1)\big)^{\gamma}}_{\text{extraction utility}} + \delta^2 \underbrace{\big(S(2)\big)^{\gamma}}_{\text{final stock utility}} \right)$$

As an exercise, write the Bellman equation and find the optimal control at time $t = 1$

- The Bellman equation is

$$\begin{cases} V(2,S) &= \delta^2\sqrt{S} \\ V(1,S) &= \max_{0\leq h\leq S}\left(\delta\sqrt{h} + V\big(2, S-h\big)\right) \\ V(0,S) &= \max_{0\leq h\leq S}\left(\sqrt{h} + V\big(1, S-h\big)\right) \end{cases}$$

- In particular, we deduce that

$$V(1,S) = \max_{0\leq h\leq S}\left(\delta\sqrt{h} + \delta^2\sqrt{S-h}\right)$$

- By differenting with respect to h, we find that the maximum is achieved at $h^\star$ solution of $(S - h^\star) = \delta^2 h^\star$, that is,

$$h^\star = \frac{1}{1+\delta^2}S$$

Then find the Bellman functions and the optimal policies at times $t = 0$ and $t = 1$

- The Bellman functions at times $t = 0$, $t = 1$ and $t = 2$ are

$$\begin{cases} V(2, S) &= \delta^2 \sqrt{S} \\ V(1, S) &= \delta \sqrt{1 + \delta^2} \sqrt{S} \\ V(0, S) &= \end{cases}$$

- The optimal policies at times $t = 0$ and $t = 1$ are

$$\begin{cases} \mathtt{Pol}^\star(1, S) &= \frac{1}{1+\delta^2} S \\ \mathtt{Pol}^\star(0, S) &= \frac{1}{1+\delta^2+\delta^4} S \end{cases}$$

Finally, find the optimal trajectories

$$
\begin{cases}
S^\star(0) & = & S_0 \\
h^\star(0) & = & \texttt{Pol}^\star(0, S_0) = \frac{1}{1+\delta^2+\delta^4} S_0 \\
S^\star(1) & = & S^\star(0) - h^\star(0) = \frac{\delta^2(1+\delta^2)}{1+\delta^2+\delta^4} S_0 \\
h^\star(1) & = & \texttt{Pol}^\star\big(1, S^\star(1)\big) = \frac{\delta^2}{1+\delta^2+\delta^4} S_0 \\
S^\star(2) & = & S^\star(1) - h^\star(1) = \frac{\delta^4}{1+\delta^2+\delta^4} S_0
\end{cases}
$$

Optimal management of an exhaustible resource

We now consider the multi-period case where time t runs from t_0 to T

- We consider an exhaustible resource

$$S(t+1) = \underbrace{S(t)}_{\text{stock}} - \underbrace{h(t)}_{\text{extraction}} \quad , \quad 0 \leq h(t) \leq S(t)$$

- where one maximizes intertemporal discounted utility

$$\max_{h(t_0),\ldots,h(T-1)} \Big(\sum_{t=t_0}^{T-1} \delta^{t-t_0} \underbrace{\big(h(t)\big)^{\gamma}}_{\text{extraction utility}} + \delta^{T-t_0} \underbrace{\big(S(T)\big)^{\gamma}}_{\text{final stock utility}} \Big)$$

As an exercise, provide the dynamic programming equation attached to this specific problem

$$\begin{cases} V(T,x) & = & \mathtt{K}(x) \\ V(t,x) & = & \max_{u\in\mathbb{B}(t,x)} \Big(\underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}} + \overbrace{V\big(t+1, \underbrace{\mathtt{Dyn}(t,x,u)}_{\text{future state}}\big)}^{\text{optimal payoff}} \Big) \end{cases}$$

Data	Abstract	Specific
State	x	S
Control	u	h
Control constraints	$u \in \mathbb{B}(t,x)$	$h \in [0, S]$
Dynamic	$\mathtt{Dyn}(t,x,u)$	$S - h$
Instantaneous gain	$\mathtt{L}(t,x,u)$	$\delta^{t-t_0} h^\gamma$
Final gain	$\mathtt{K}(x)$	$\delta^{T-t_0} S^\gamma$

$$\begin{cases} V(T,S) & = & ? \\ V(t,S) & = & \max_? \big(? + V(t+1, ?) \big) \end{cases}$$

Dynamic programming equation

$$\begin{cases} V(T,x) &= \mathtt{K}(x) \\ V(t,x) &= \max_{u \in \mathbb{B}(t,x)} \Big[\mathtt{L}(t,x,u) + V\big(t+1, \mathtt{Dyn}(t,x,u)\big) \Big] \end{cases}$$

$$\begin{aligned} V(T,S) &= \delta^{T-t_0} S^{\gamma} \\ V(T-1,S) &= \max_{0 \leq h \leq S} \big[\delta^{T-t_0-1} h^{\gamma} + V\big(T, S-h\big) \big] \\ &= \max_{0 \leq h \leq S} \big[\delta^{T-t_0-1} h^{\gamma} + \delta^{T-t_0} (S-h)^{\gamma} \big] \\ &= \delta^{T-t_0-1} \max_{0 \leq h \leq S} \big[h^{\gamma} + \delta (S-h)^{\gamma} \big] \end{aligned}$$

What is the influence of the discount factor δ on optimal strategies and extraction paths?

One can prove by induction that

$$\begin{cases} V(t,S) & = & \delta^{t-t_0} b(t)^{\gamma-1} S^{\gamma} \\ h^{\star}(t,S) & = & b(t)S \end{cases}$$

where

$$\frac{1}{b(t)} = \frac{\theta - 1}{\theta - \theta^{t-T}} = \underbrace{1 + \theta + \cdots + \theta^{T-t}}_{T-t+1 \text{ terms}} \quad \text{with} \quad \theta = \delta^{\frac{1}{1-\gamma}}$$

- What happens to $b(t)$ when time goes on?
- When $\delta \downarrow 0$ (sensititivity to the present), one has $\theta \downarrow 0$: what happens to $b(t)$, hence to $h^{\star}(t,S)$?
- When $\delta \uparrow 1$ (sensititivity to the future), one has $\theta \uparrow 1$: what happens to $b(t)$, hence to $h^{\star}(t,S)$?
- Show that $h^{\star}(t+1) = \theta h^{\star}(t)$.

Influence of the discount factor δ

The optimal fraction of the stock S to be extracted at time t is

$$b(t) = \frac{\theta - \theta^{t-T}}{\theta - 1} = \frac{1}{\underbrace{1 + \theta + \cdots + \theta^{T-t}}_{T-t+1 \text{ terms}}} \text{ with } \theta = \delta^{\frac{1}{1-\gamma}}$$

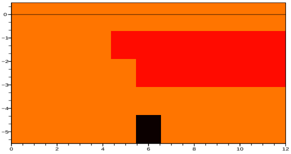
so that, in particular

$$b(T) = 1\,, \quad b(T-1) = \frac{1}{1+\theta}\,, \quad b(T-2) = \frac{1}{1+\theta+\theta^2}$$

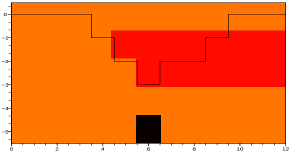
Discount factor δ	θ	optimal fraction $b(t)$	interpretation
$\delta \uparrow 1$	$\theta \uparrow 1$	$b(t) \downarrow \frac{1}{T-t+1}$	equity
$\delta \downarrow 0$	$\theta \downarrow 0$	$b(t) \uparrow 1$	myopism and greed

Comparing $\delta = 0.95$ (left) with $\delta = 0.99$ (right)

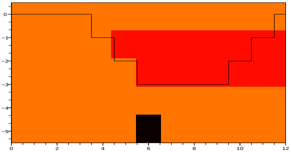
Optimal extraction profile at initial time t=1 for discount rate 0.95



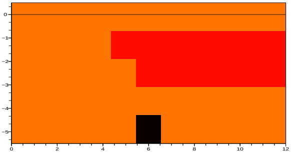
Optimal extraction profile at time t=12 for discount rate 0.95



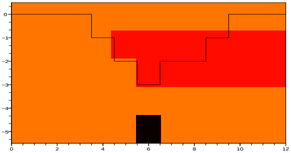
Optimal extraction profile at ultimate time t=20 for discount rate 0.95



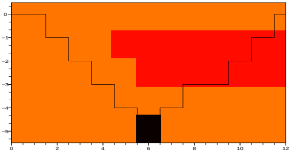
Optimal extraction profile at ultimate time t=1 for discount rate 0.99



Optimal extraction profile at ultimate time t=12 for discount rate 0.99



Optimal extraction profile at ultimate time t=30 for discount rate 0.99



Outline of the presentation

Discounting erases the future

The French public discount rate

En France, le rapport *Révision du taux d'actualisation des investissements publics* (Commissariat général du Plan, groupe d'experts présidé par Daniel Lebègue, janvier 2005) a conduit à diviser par deux (de 8% à 4%) le taux d'actualisation à retenir pour évaluer la rentabilité des choix d'investissements publics

$$\delta = \frac{1}{1+0.04} \approx 0.96$$

The future in one hundred years is valued, seen from today, 2%

$$(\frac{1}{1+0.04})^{10} \approx 0.68 \; , \quad (\frac{1}{1+0.04})^{50} \approx 0.14 \; , \quad (\frac{1}{1+0.04})^{100} \approx 0.02$$

The discount rate is not necessarily an interest rate

M. Boiteux, À propos de la "critique de la théorie de l'actualisation telle qu'employée en France", Revue d'Économie Politique, 1976.

- "Le taux d'actualisation optimal pour orienter les choix d'intérêt général" n'est pas "nécessairement égal dans la réalité au taux d'intérêt d'un quelconque marché monétaire et financier"
- "l'actualisation, instrument de cohérence des choix"

Discounting may be related to random final time

- Nicholas Stern. *The Economics of Climate Change*. Cambridge University Press, 2006.
  (...) following distinguished economists from Frank Ramsey in the 1920s to Amartya Sen and Robert Solow more recently, the only sound ethical basis for placing less value on the utility (as opposed to consumption) of future generations was the uncertainty over whether or not the world will exist, or whether those generations will all be present
- The discounted utility criterion can be written without discounting as a mathematical expectation

$$\sum_{t=0}^{+\infty} \delta^t \mathrm{L}(c(t)) = \mathbb{E}\left[\sum_{t=0}^{\tau-1} \mathrm{L}(c(t))\right]$$

  where the random final time τ follows a memoryless geometric distribution on the set $\{1, 2, 3 \dots\}$ with parameter $1 - \delta = \mathbb{P}(\tau = 1)$

We interpret discounting till infinity as no discounting till Geometric distributed final time

- Let τ denote a random variable having geometric distribution

$$\mathbb{P}(\tau \geq s) = \delta^{s-1} , \quad s = 1, 2, 3 \ldots$$

- or, equivalently, that

$$\mathbb{P}(\tau = s) = (1-\delta)\delta^{s-1} , \quad s = 1, 2, 3 \ldots$$

- Then, we have that

$$\underbrace{\mathbb{E}\left[\sum_{t=0}^{\tau-1} \mathtt{L}\big(c(t)\big)\right]}_{\text{no discounting}} = \mathbb{E}\left[\sum_{t=0}^{+\infty} \mathbf{1}_{\{\tau-1 \geq t\}} \mathtt{L}\big(c(t)\big)\right] = \underbrace{\sum_{t=0}^{+\infty} \delta^t \mathtt{L}\big(c(t)\big)}_{\text{discounting}}$$

We interpret the discount factor $\delta \in [0,1[$ in term of mean lifetime

	discount factor	discount rate	mean time
discount factor	δ	$r_e = \frac{1-\delta}{\delta}$	$\overline{\theta} = \frac{1}{1-\delta}$
discount rate	$\delta = \frac{1}{1+r_e}$	r_e	$\overline{\theta} = \frac{1+r_e}{r_e}$
mean time	$\delta = \frac{\overline{\theta}-1}{\overline{\theta}}$	$r_e = \frac{1}{\overline{\theta}-1}$	$\overline{\theta}$

discount factor δ	discount rate r_e	mean time $\overline{\theta}$	survival $\mathbb{P}(\theta \geq \overline{\theta})$
0.990	1%	100	0.370
0.952	5%	21	0.377
0.909	10%	11	0.386

"The pure time discount rate" and chance of extinction

Nicholas Stern. *The Economics of Climate Change*.
Cambridge University Press, 2006.
(...) we should interpret the factor $e^{-\Delta t}$ in $W = \int_0^\infty \mathtt{L}(c(t))e^{-\Delta t}dt$ as the probability that the world exists at that time

Pure time preference Δ	Probability of human race not surviving 10 years	Probability of human race not surviving 100 years
0,1	0,010	0,095
0,5	0,049	0,393
1,0	0,095	0,632
1,5	0,139	0,777

Outline of the presentation

Optimal management of a biomass linear growth model with catches

- We consider a biomass linear growth model with catches

$$B(t+1) = R\big(\underbrace{B(t)}_{\text{biomass}} - \underbrace{h(t)}_{\text{catches}} \big) , \quad 0 \leq h(t) \leq B(t)$$

where $R = 1 + r$, with biological growth rate $r > 0$,

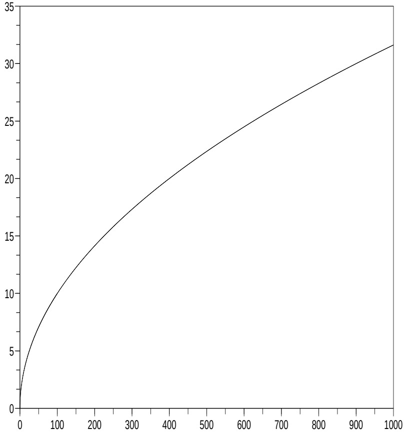
- and where one maximizes intertemporal discounted utility

$$\max_{h(t_0),\dots,h(T-1)} \Big(\sum_{t=t_0}^{T-1} \delta^{t-t_0} \underbrace{\mathtt{L}(h(t))}_{\text{catches utility}} + \delta^{T-t_0} \underbrace{\mathtt{L}(B(T))}_{\text{final biomass utility}} \Big)$$

- in the special case where, for the sake of simplicity, discount rate r_e and biological growth rate $r = R - 1$ are equal

$$r_e = r \iff \delta = \frac{1}{R}$$

The utility function L is supposed to be stricly increasing and strictly concave



The utility function $\mathtt{L}$ is

- strictly increasing: $\mathtt{L}' > 0$
- strictly concave: $\mathtt{L}'' < 0$

As an exercise, provide the dynamic programming equation attached to this specific problem

$$\begin{cases} V(T,x) & = & \mathtt{K}(x) \\ V(t,x) & = & \max_{u \in \mathbb{B}(t,x)} \Big(\underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}} + \overbrace{V\big(t+1, \underbrace{\mathtt{Dyn}(t,x,u)}_{\text{future state}}\big)}^{\text{optimal payoff}} \Big) \end{cases}$$

State	x	B
Control	u	h
Control constraints	$u \in \mathbb{B}(t,x)$	$h \in [0,B]$
Dynamic	$\mathtt{Dyn}(t,x,u)$	$R(B-h)$
Instantaneous gain	$\mathtt{L}(t,x,u)$	$\delta^{t-t_0}\mathtt{L}(h)$
Final gain	$\mathtt{K}(x)$	$\delta^{T-t_0}\mathtt{L}(B)$

$$\begin{cases} V(T,B) & = & ? \\ V(t,B) & = & \max_{?} \big(? + V(t+1, ?) \big) \end{cases}$$

Detail the dynamic programming equation for $t = T$ and $t = T - 1$

$$\begin{cases} V(T,x) = \mathtt{K}(x) \\ V(t,x) = \max_{u \in \mathbb{B}(t,x)} \Big(\underbrace{\mathtt{L}(t,x,u)}_{\text{instantaneous gain}} + \overbrace{V\big(t+1, \underbrace{\mathtt{Dyn}(t,x,u)}_{\text{future state}}\big)}^{\text{optimal payoff}} \Big) \end{cases}$$

$$V(T,B) = \overbrace{\delta^{T-t_0}\mathtt{L}(B)}^{\text{final utility}} \text{ with } R\delta = 1$$

$$\begin{aligned} V(T-1,B) &= \max_{0 \le h \le B} \Big[R^{t_0-T+1}\mathtt{L}(h) + V\big(T, \overbrace{R(B-h)}^{\text{future state}}\big) \Big] \\ &= R^{t_0-T} \max_{0 \le h \le B} \Big[R\mathtt{L}(h) + \mathtt{L}\big(R(B-h)\big) \Big] \end{aligned}$$

The optimal decision rule is linear in the biomass

- The maximum $V(T-1,B) = R^{T-t_0-1} \max_{0 \leq h \leq B} \Big[\mathtt{L}(h) + R\mathtt{L}\big(R(B-h)\big) \Big]$ is achieved for

$$h^{\star}(T-1,B) = \frac{R}{1+R} B$$

- and the maximum is thus

$$V(T-1,B) = \frac{1+R}{R^{T-t_0}} \times \mathtt{L}\big(\frac{R}{1+R} B\big)$$

The optimal trajectories have the property that catches are stationary

- Prove by induction that

$$h^{\star}(t,B) = \frac{R^{T-t}}{1+R+\cdots+R^{T-t}}B$$

- Compute the optimal trajectory given by

$$B(t+1) = R\big(B(t)-h(t)\big) \; , \quad h(t) = \frac{R^{T-t}}{1+R+\cdots+R^{T-t}}B(t)$$

- Show that $h(t+1) = h(t)$ for all times t

Outline of the presentation

Optimal management of a biomass linear growth model with catches

- We consider a biomass linear growth model with catches

$$B(t+1) = R\big(\underbrace{B(t)}_{\text{biomass}} - \underbrace{h(t)}_{\text{catches}} \big) , \quad 0 \leq h(t) \leq B(t)$$

  where $r_b = R - 1$ is the biological growth rate

- and where one maximizes intertemporal discounted profit (with fixed prices and without costs)

$$\max_{h(t_0),h(t_0+1),\ldots,h(T-1)} \sum_{t=t_0}^{T-1} \delta^{t-t_0} h(t)$$

  where r_e is the discount rate

Low discount rate $r_e < r_b$ (high preference for the future)

If $r_b > r_e$,

- there are no catches

$$h^\star(t_0) = h^\star(t_0 + 1) = \cdots = h^\star(T - 2) = 0$$

- except at the penultimate period where

$$h^\star(T - 1) = B_0 R^{T-1-t_0}$$

High discount rate $r_e > r_b$ (high preference for the present)

If $r_b < r_e$,

- all the biomass is captured at the initial time t_0

$$h^\star(t_0) = B_0$$

- so that the resource is immediately extinct

	yearly growth rate
whales	2 — 5%
money	5%

Special case $r_b = r_e$

If $r_b = r_e$,

- many solutions exist,
- and one optimal solution is given by equitable catches

$$h^\star(t_0) = h^\star(t_0 + 1) = \cdots = h^\star(T - 1) = \frac{B_0}{T - t_0}$$

Outline of the presentation

Outline of the presentation

Green Golden criterion and dictatorship of the future

- We consider a biomass growth model with catches

$$B(t+1) = \overbrace{\mathtt{Biol}}^{\text{dynamics}} \Big(\underbrace{B(t)}_{\text{biomass}} - \underbrace{h(t)}_{\text{catches}} \Big)$$

- where one maximizes the final utility

$$\max_{h(t_0),\ldots,h(T-1)} \mathtt{K}\big(B(T)\big)$$

Show that, when

- final utility $B \mapsto \mathtt{K}(B)$
- dynamics $B \mapsto \mathtt{Biol}(B)$

are increasing with biomass B, it is optimal never to capture

Outline of the presentation

Maximin dynamic programming equation

- For the optimization problem

$$\max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t_0,x_0)} \min_{t=t_0,\ldots,T-1} \mathtt{L}\big(t,x(t),u(t)\big)$$

- the value function

$$V(t,x) = \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t,x)} \Big(\min_{s=t,\ldots,T-1} \mathtt{L}\big(s,x(s),u(s)\big)\Big)$$

- is the solution of

$$\begin{cases} V(T,x) &= +\infty \\ V(t,x) &= \max_{u\in\mathbb{B}(t,x)} \min\Big(\mathtt{L}(t,x,u), V\big(t+1,\mathtt{Dyn}(t,x,u)\big)\Big) \end{cases}$$

Maximin dynamic programming equation

- In the general case, the value function

$$V(t,x) = \max_{\big(x(\cdot),u(\cdot)\big)\in\mathcal{T}^{\mathrm{ad}}(t,x)} \min\Big(\min_{s=t,\dots,T-1} \mathtt{L}\big(s,x(s),u(s)\big), \mathtt{K}\big(x(T)\big)\Big)$$

- is the solution of

$$\begin{aligned} V(T,x) &= \mathtt{K}\big(x\big)\ , \quad \forall x \in \mathbb{V}\mathrm{iab}(T) \\ V(t,x) &= \max_{u\in\mathbb{B}^{\mathsf{viab}}(t,x)} \min\Big(\mathtt{L}(t,x,u), V\big(t+1, \mathtt{Dyn}(t,x,u)\big)\Big)\ , \\ & \quad \forall x \in \mathbb{V}\mathrm{iab}(t) \end{aligned}$$

Maximin and viability

$$\mathbb{V}\mathrm{iab}(t, L^\flat) = \left\{ x \in \mathbb{X} \;\middle|\; \begin{array}{l} \exists (x(\cdot), u(\cdot)) \text{ such that } \forall s = t, \ldots, T \\ x(s+1) = \mathtt{Dyn}(s, x(s), u(s)) \\ x(t) = x \\ \mathtt{L}(s, x(s), u(s)) \geq L^\flat \end{array} \right\}$$

Proposition

The value function associated to the maximin problem

$$\max_{u(\cdot)} \min_{t=t_0,\ldots,T-1} \mathtt{L}\big(t, x(t), u(t)\big)$$

satisfies

$$V(t,x) = \max\{L^\flat \in \mathbb{R} \mid x \in \mathbb{V}\mathrm{iab}(t, L^\flat)\}$$

Maximin for an exhaustible resource

- Consider the exhaustible resource management

$$S(t+1) = \underbrace{S(t)}_{\text{stock}} - \underbrace{h(t)}_{\text{extraction}} \quad , \quad 0 \leq h(t) \leq S(t)$$

- where the utility of the least favoured generation is maximized

$$\max_{h(t_0),\ldots,h(T-1)} \min_{t=t_0,..,T-1} \mathtt{L}\big(h(t)\big)$$

The "maximin" approach leads to intergenerational equity

Assuming that the utility $h \mapsto \mathtt{L}(h)$ is increasing,

- the optimal decision rule is

$$h^\star(t, S) = \frac{S}{T - t}$$

- the maximin optimal stock path $S^\star(\cdot)$ is regularly decreasing

$$S^\star(t) = \frac{T - t}{T - t_0} S_0$$

- the maximin optimal extraction path $h^\star(\cdot)$ is stationary

$$h^\star(t) = h^\star\big(t, S^\star(t)\big) = \frac{S_0}{T - t_0}$$