Dynamic Programming and Sufficient Optimality Conditions

Continuous and discrete time dynamics

Alain Rapaport

Doctoral lectures, I2S, U. Montpellier

march 2016

Optimal Control Problems

Consider a state space $\mathrm{X} \subset \mathbb{R}^n$ and a control space $\mathrm{U} \subset \mathbb{R}^p$

$$\begin{cases} \dot{x} = f(t, x, u), \\ x(t_0) = x_0 \in \mathrm{X} \end{cases} \qquad \mathcal{U}_{[t_0,T]} := \{u(\cdot) \text{ meas. } [t_0, T] \mapsto \mathrm{U}\}$$

Assumption. *Given $(x_0, u(\cdot)) \in \mathrm{X} \times \mathcal{U}_{[t_0,T]}$, there exists an unique solution of $\dot{x} = f(t, x, u(t))$, $x(t_0) = x_0$ defined on $[t_0, T]$.*

$$J_T(t_0, x_0, u(\cdot)) = \phi_T(x(T)) + \int_{t_0}^{T} l(\tau, x(\tau), u(\tau))\, d\tau$$

$$\to \inf_{u(\cdot) \in \mathcal{U}_{[t_0,T]}} J_T(t_0, x_0, u(\cdot))$$

The Value Function

Definition. The *value function* $V : (-\infty, T] \times \mathrm{X} \mapsto \mathbb{R}$ is

$$V(t_0, x_0) = \inf_{u(\cdot)} J_T(t_0, x_0, u(\cdot))$$

Remark. One has the boundary condition

$$V(T, x) = \phi_T(x), \quad \forall x \in \mathrm{X}$$

Objectives :

- caracterize the function V on the whole state space
- when $V(t_0, x_0) = \min_{u(\cdot)} J_T(t_0, x_0, u(\cdot))$, caracterize an optimal control $u^\star(\cdot)$

Simpler framework

non autonomous	autonomous
$f(t,x,u),\ l(t,x,u)$	$f(x,u),\ l(x,u)$

- non autonomous $\rightarrow$ autonomous :

$$\begin{cases} \dot{x} &= f(z,x,u) \\ \dot{z} &= 1 \end{cases} \qquad \tilde{l}(z,x,u) = l(t,x,u)$$

	Bolza problem	Mayer problem
$J(t_0,x_0,u(\cdot))$:	$\phi_T(x(T)) + \int_{t_0}^{T} l(x(\tau),u(\tau))\,d\tau$	$\phi_T(x(T))$

- Bolza $\rightarrow$ Mayer :

$$\begin{cases} \dot{x} &= f(x,u) \\ \dot{z} &= l(x,u) \end{cases} \qquad \tilde{\phi}_T(x,z) = \phi_T(x) + z$$

Minimal time problem

$$T_{\mathcal{T}}(x(\cdot)) = \left| \begin{array}{ll} +\infty & \text{if } x(t) \notin \mathcal{T}\,,\ \forall t \geq t_0 \\ \inf\{t \geq t_0,\ x(t) \in \mathcal{T}\} - t_0 & \text{otherwise} \end{array} \right.$$

$$J_T(t_0, x_0, u(\cdot)) = \psi_{X\setminus\mathcal{C}}(x(T)) + \int_{t_0}^{T} d\tau \ \rightarrow \inf_{T} \left\{ \inf_{u(\cdot)\in\mathcal{U}_{[t_0,T]}} J_T(t_0, x_0, u(\cdot)) \right\}$$

$$\text{with } \psi_K(x) = \left| \begin{array}{ll} 0 & x \in K \\ +\infty & x \notin K \end{array} \right.$$

- We set $\phi_T = \psi_{X\setminus\mathcal{C}}$, $l = 1$ and T is free

Remarks.

- For autonomous dynamics and free terminal time, V does not depend on t_0.
- The boundary condition is $V(x) = 0,\ \forall x \in \mathcal{T}$

Dynamic programming for Mayer problem

Value function :

$$V(t_0, x_0) = \inf_{u(\cdot)} \phi(x(T))$$

Lemma. Fix an initial condition (t_0, x_0)

1. For any trajectory $x(\cdot)$, the map $t \mapsto V(t, x(t))$ is non-decreasing
2. If there exists an optimal trajectory $x^\star(\cdot)$, then $V(t, x^\star(t)) = V(t_0, x_0)$ for any $t \in [t_0, T]$.

Dynamic programming principle : For any $t_1 \in (t_0, T)$, one has

$$V(t_0, x_0) = \inf_{u(\cdot) \in \mathcal{U}_{[t_0, t_1]}} V(t_1, x(t_1))$$

Dynamic programming for more general problems

Non-autonomous Bolza problem. For any $t_1 \in (t_0, T)$, one has

$$V(t_0, x_0) = \inf_{u(\cdot)\in\mathcal{U}_{[t_0,t_1]}} \left\{ V(t_1, x(t_1)) + \int_{t_0}^{t_1} l(\tau, x(\tau), u(\tau))\, d\tau \right\}$$

Autonomous minimal time problem. For $t \in (0, V(t_0))$, one has

$$V(x_0) = \inf_{u(\cdot)\in\mathcal{U}_{[0,t]}} V(x(t)) + t$$

Discrete time framework

$$t_0, T \in \mathbb{N}, \quad \mathcal{U}_{\{t_0 \cdots T\}} := \{u(\cdot) : \{t_0, t_0+1, \cdots T\} \mapsto \mathrm{U}\}$$

$$\begin{cases} x(t+1) = F(t, x(t), u(t)), \\ x(t_0) = x_0 \in \mathrm{X} \end{cases}$$

$$J(t_0, x_0, u(\cdot)) = \phi_T(x(T)) + \sum_{t=t_0}^{T} L(t, x(t), u(t))$$

Value function : $V(t_0, x_0) = \inf_{u(\cdot) \in \mathcal{U}_{\{t_0 \cdots T\}}} J(t_0, x_0, u(\cdot))$

Dynamic Programming Principle :

$$V(t_0, x_0) = \inf_{u(\cdot) \in \mathcal{U}_{\{t_0 \cdots t_1\}}} \left\{ V(t_1, x(t_1)) + \sum_{t=t_0}^{t_1} L(t, x(t), u(t)) \right\}$$

The (autonomous) discrete time case

For $t_0 = t$ and $t_1 = t + 1$, one has

- for fixed terminal time :

$$\begin{cases} V(t,x) & = & \inf_{u \in \mathrm{U}} V(t+1, F(x,u)) + l(x,u), \quad t < T \\ V(T,x) & = & \phi_T(x) \end{cases}$$

Moreover $u^\star(t,x) \in \operatorname{Arg}\min_{u \in \mathrm{U}} V(t+1, F(x,u)) + l(x,u)$ is **optimal**.

- for free terminal time :

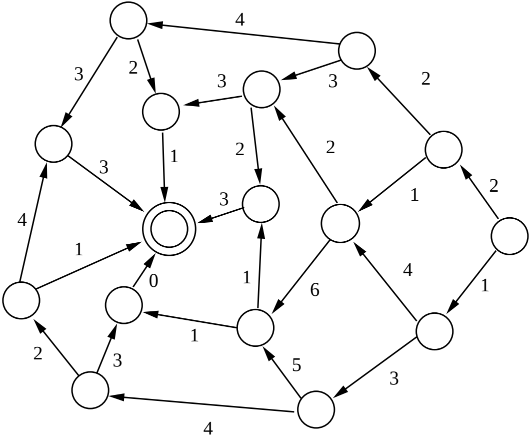
$$\begin{cases} V(x) & = & \inf_{u \in \mathrm{U}} V(F(x,u)) + l(x,u), \quad x \notin \mathcal{C} \\ V(x) & = & \phi(x), \quad x \in \mathcal{T} \end{cases}$$

Moreover $u^\star(x) \in \operatorname{Arg}\min_{u \in \mathrm{U}} V(F(x,u)) + l(x,u)$ is **optimal**.

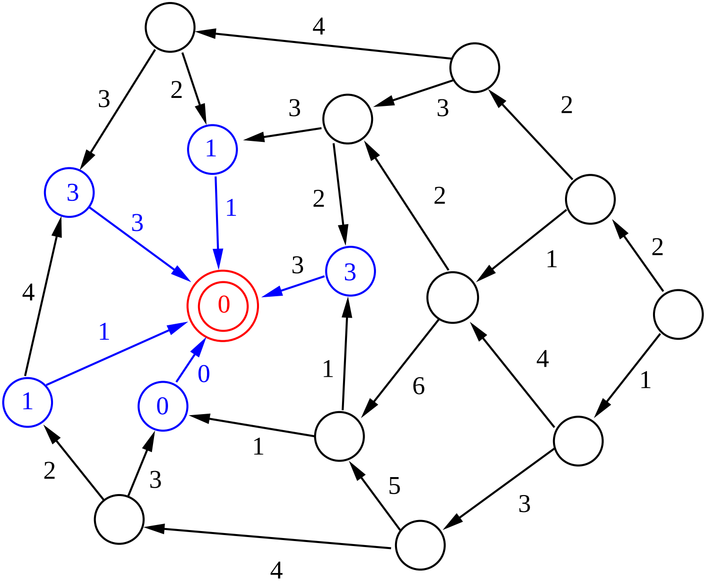
- **backward dynamic programming algorithm**

An example

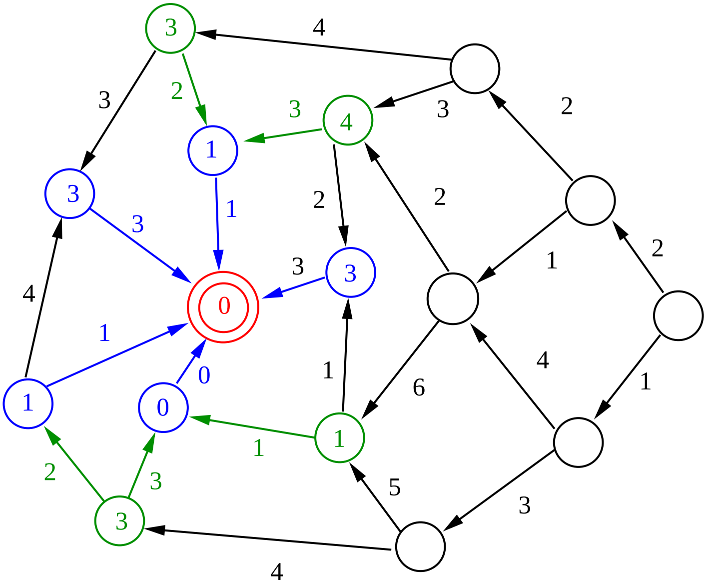
$$\mathrm{X} = \{x_0, x_1, x_2, \cdots\},\ \mathrm{U} = \{u_1, u_2, \cdots\},\ \mathcal{T} = \{X_0\},\ J = \sum_{l=0}^{m} C(x_{i_l}, u_{k_l})$$



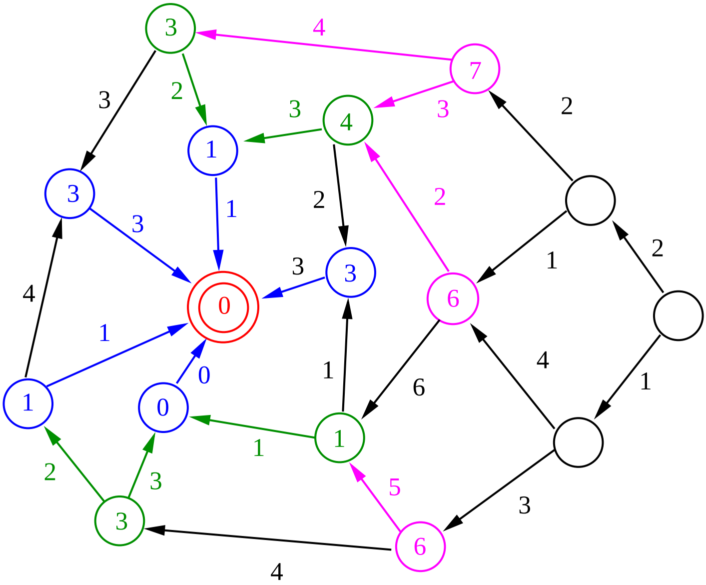
Step 1



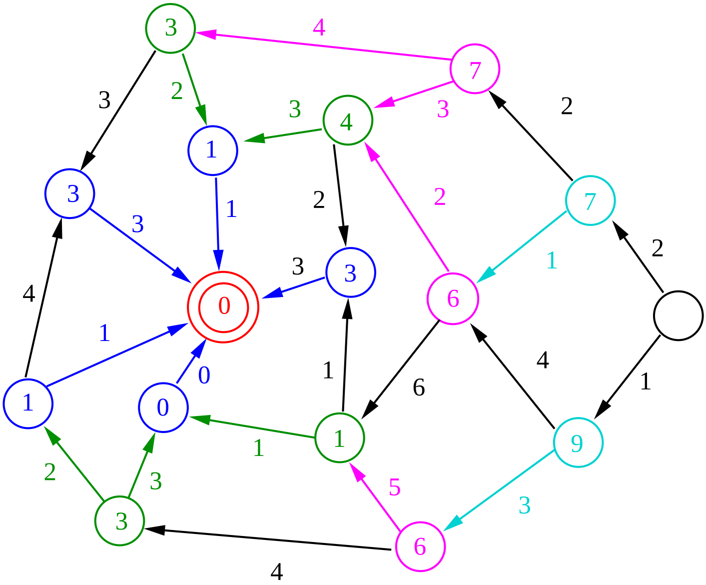
Step 2



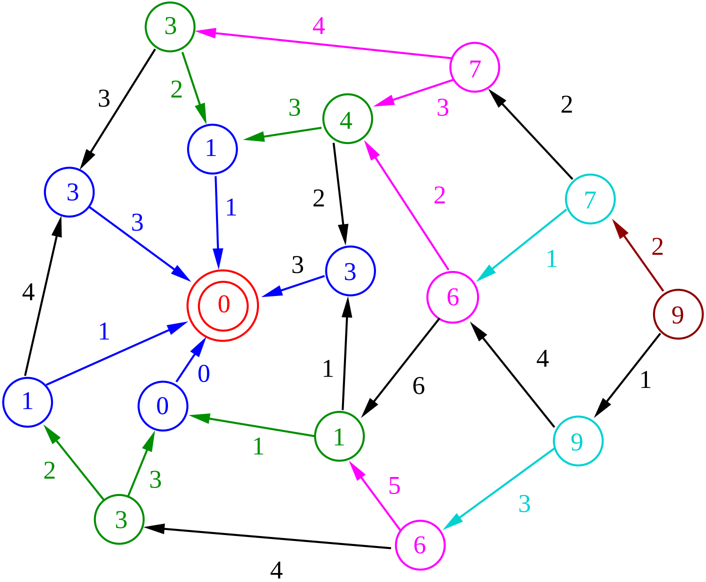
Step 3



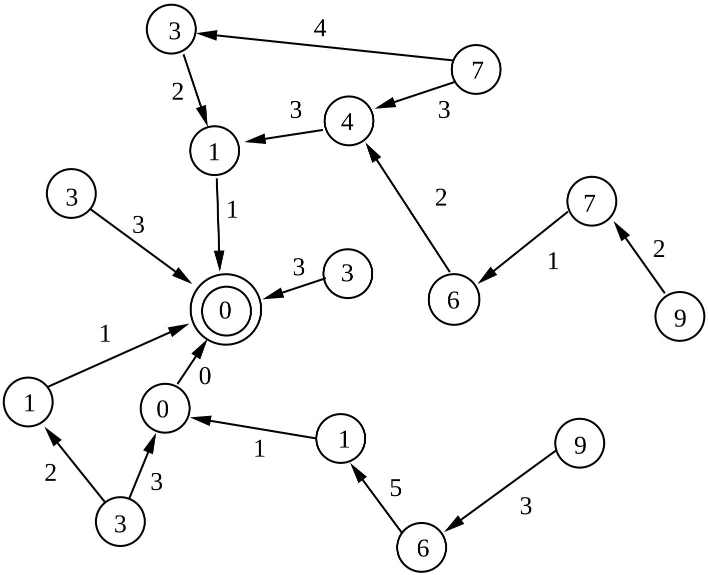
Step 4



Step 5

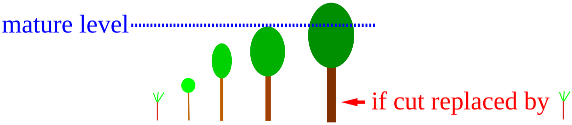


Solution

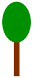


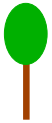
Example of forest harvesting

A forest is composed of S parcels where n is the number of years for a planted parcel to become *mature* :



State of the system *(here for* $n = 5$*)* :

x_1 : nb of parcels of age ≥ 4 

x_2 : nb of parcels of age $\in [3, 4[$ 

x_3 : nb of parcels of age $\in [2, 3[$ 

x_4 : nb of parcels of age $\in [1, 2[$ 

x_5 : nb of parcels of age < 1 

Control : nb of harvested parcels, mature in the coming year :

$$u(t) \in \{0, \cdots, x_1(t) + x_2(t)\}$$

Optimization problem

Criterion : $J_T(t_0, x_0, u(\cdot)) = \phi_T(x(T)) + \sum_{t=t_0}^{T} \delta^t U(u(t)) \to \max_{u(\cdot)}$

where $\delta \in]0,1[$: discount factor
$U(\cdot)$: utility function (with $U' > 0$ and $U'' < 0$)

Faustmann computation :

$$\begin{aligned}
\phi_T(x(T)) &= \max_{u(\cdot)} \phi_{2T}(x(2T) + \sum_{t=T}^{2T} \delta^T U(u(t)) \\
&= \max_{u(\cdot)} \phi_{2T}(x(3T) + \sum_{t=T}^{3T} \delta^T U(u(t)) = \cdots
\end{aligned}$$

- Infinite horizon criterion : $J(t_0, x_0, u(\cdot)) = \lim_{T \to +\infty} \sum_{t=t_0}^{T} \delta^t (u(t))$

Remarks.

- $V(t_0, x_0) = \delta^{t_0} V(0, x_0)$ when the dynamics is autonomous
- no terminal condition...

General setup

$$\begin{cases} x(t+1) = F(x(t), u), \;\; u \in \mathrm{U} \\ x(0) = x_0 \in \mathrm{X} \end{cases} \qquad J(x_0, u(\cdot)) = \sum_{t=0}^{+\infty} \delta^t L(x(t), u(t))$$

Hypothesis. U compact, F, L continuous w.r.t. u and

$$\exists M < +\infty \text{ s.t. } \max_{u \in \mathrm{U}} L(x, u) < M, \quad \forall x \in \mathrm{X}$$

Definitions.

- $\mathcal{B} := \{W : \mathrm{X} \mapsto \mathbb{R} \text{ bounded continuous}\}$
- Bellman operator $\Phi : \mathcal{B} \mapsto \mathcal{B}$ defined as

$$\Phi[W](x) = \max_{u \in \mathrm{U}} L(x, u) + \delta W(F(x, u)), \quad x \in \mathrm{X}$$

Dynamic programming. V satisfies $V = \Phi[V]$

Dynamic programming

$$\Phi[\gamma](x) = \max_{u \in \mathsf{U}} L(x,u) + \delta W(F(x,u)),\ W \in \mathcal{B},\ x \in \mathsf{X}$$

Properties.

- $\delta < 1 \Rightarrow \Phi$ is contractif i.e.

$$||\Phi[W_1] - \Phi[W_2]||_\infty \leq \delta ||W_1 - W_2||_\infty$$

- V is the unique fixed point of Φ
- for any $V_0 \in \mathcal{B}$, one has

$$\lim_{k \to +\infty} \Phi^k[W_0] = V$$

Example of forest harvesting

$$\begin{cases} x(t+1) &= Ax(t) + Bu(t) \\ x(0) &= x_0 \end{cases} \qquad u(t) \in \{0, \cdots, Cx(t)\}$$

$$\text{with } A = \begin{bmatrix} 1 & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$\text{and } C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

Example of forest harvesting

Let S be the number of parcels, and n the number of age classes

$$\mathrm{X} := \left\{ x \in \mathbb{N}^n \text{ s.t. } \sum_{i=1}^{n} x_i = S \right\} \text{ with } \mathrm{card}(\mathrm{X}) = C_{S+n-1}^{n-1}$$

Transition matrix. $T : \mathrm{X} \times \mathrm{U} \mapsto \mathrm{X}$

u	0	1	2	$\cdots$
① $\begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	① $\begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	② $\begin{bmatrix} 8 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	③ $\begin{bmatrix} 7 \\ 0 \\ 0 \\ 2 \end{bmatrix}$	$\cdots$
② $\begin{bmatrix} 8 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	① $\begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	② $\begin{bmatrix} 8 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	③ $\begin{bmatrix} 7 \\ 0 \\ 0 \\ 2 \end{bmatrix}$	$\cdots$
$\vdots$				

Computer implementation

X : matrix of the card(X) state vectors
U : column vector of the values of $U(u)$ for $u = 0, 1, \cdots S$
T : transition matrix of size card(X) $\times$ $(S+1)$

```
function NW=operator(W)
  for i=1:nb_states
     nc=C*X(:,i);
     NW(i)=max(U(1:nc)+delta*W(T(i,1:nc));
  end
endfunction
```

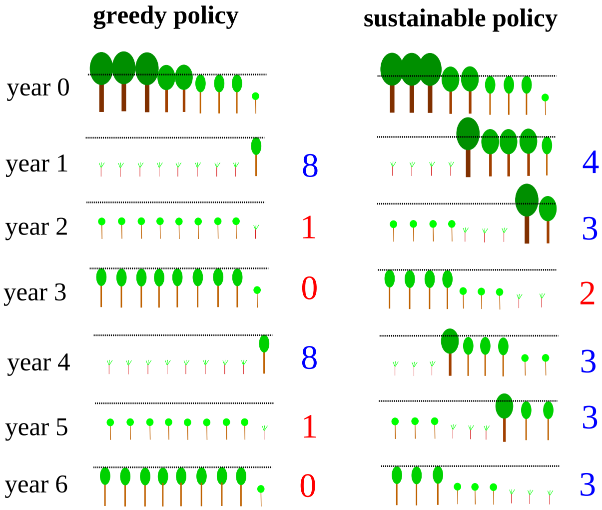
Numerical simulations

Launch the Scilab script

`foret.sce`

with different choices of δ and β (U concavity)

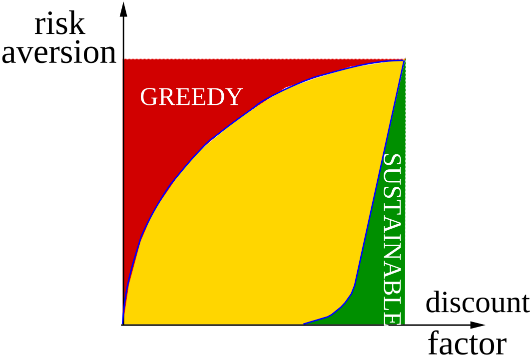
Optimal policies



Inverse optimality

$$\text{Greedy cost : } J_G(x) = \frac{U(x_1 + x_2) + \sum_{j=2}^{n} \delta^j U(x_j)}{1 - \delta^{n-1}}$$

Question : when J_G satisfies $\Phi[J_G] = J_G$?



The Hamilton-Jacobi-Bellman equation

For an autonomous Mayer problem, define the **Hamiltonians**

$$H(x,p,u) = p.f(x,u)$$
$$\bar{H}(x,p) = \min_{u \in \mathrm{U}} H(x,p,u) \qquad (x,p) \in \mathrm{X} \times \mathbb{R}^n$$

Proposition If the value function V is C^1, then V is solution of the Hamilton-Jacobi-Bellman (HJB) equation

$$\partial_t V(t,x) + \bar{H}(x, \nabla V(t,x)) = 0 \ , \quad \forall (t,x) \in (-\infty, T] \times \mathrm{X}$$

with the terminal condition :

$$V(T,x) = \phi_T(x) \ , \quad \forall x \in \mathrm{X} \ .$$

Sketch of proof. Let $v(t) = V(t, x(t))$. One can write

$$v(t+h) = v(t) + \int_t^{t+h} \partial_t V(t,x) + \partial_x V(t,x) f(x, u(\tau)) + r(\tau)\, d\tau$$

Bolza and minimal time problems

- Bolza (autonomous) control problem

$$\begin{cases} \partial_t V(t,x) + \bar{H}(x, \nabla V(t,x)) = 0 \ , & \forall (t,x) \in (-\infty, T] \times \mathrm{X} \\ V(T,x) = \phi_T(x) \ , & \forall x \in \mathrm{X} \end{cases}$$

$$\text{with } H(x,p,u) = p.f(x,u) + l(x,u)$$

- Minimal time (autonomous) control problem

$$\begin{cases} 1 + \bar{H}(x, \nabla V(x)) = 0 \ , & \forall x \notin \mathcal{T} \\ V(x) = 0 \ , & \forall x \in \mathcal{T} \end{cases}$$

$$\text{with } H(x,p,u) = p.f(x,u)$$

Example 1

$$\dot{x} = \cos(t)u\,, \quad x(t_0) = x_0 \in \mathbb{R}, \quad u \in [-1,1] \quad \to \sup_{u(\cdot)} x(T)$$

Clearly, one has $V(t_0, x_0) = x_0 + \int_{t_0}^{T} |\cos(\tau)|\, d\tau$

The function V is C^1 with $\partial_t V(t,x) = -|\cos(t)|$, $\partial_x V(t,x) = 1$ and is solution of the HJB equation :

$$\begin{cases} \partial_t V(t,x) + |\partial_x V(t,x)\cos(t)| = 0, & \forall (t,x) \in (-\infty, T] \times \mathbb{R} \\ V(T,x) = x, & \forall x \in \mathbb{R} \end{cases}$$

Remark. The optimal control $u^\star(t) = \mathrm{sgn}(\cos(t))$ is discontinuous.

Example 2

$$\dot{x} = ux\,, \quad x(t_0) = x_0 \in \mathbb{R}, \quad u \in [-1,1] \quad \to \inf_{u(\cdot)} x(1)$$

Clearly, $u^\star = -\mathrm{sgn}(x_0)$ is optimal and the value function is

$$V(t_0, x_0) = \left| \begin{array}{ll} x_0 e^{t_0 - 1} & x_0 \geq 0 \\ x_0 e^{1 - t_0} & x_0 \leq 0 \end{array} \right.$$

$$\Rightarrow (\partial_t V(t,x), \partial_x V(t,x)) = \left| \begin{array}{ll} (xe^{t-1}, e^{t-1}), & x > 0 \\ (-xe^{1-t}, e^{1-t}), & x < 0 \end{array} \right.$$

V is not differentiable on $(-\infty, 1) \times \{0\}$ but the HJB equation

$$\begin{cases} \partial_t V(t,x) - |\partial_x V(t,x)\, x| = 0, & \forall (t,x) \in (-\infty, 1] \times \mathbb{R} \\ V(1,x) = x, & \forall x \in \mathbb{R} \end{cases}$$

is fulfilled at (t,x) where V is differentiable.

Example 3

$$\dot{x} = u \in [-1,1]\,, \quad x(t_0) = x_0 \in \mathbb{R} \quad \to \inf_{u(\cdot)} \{t_f | x(t_f) \in \{-1,1\}\}$$

Clearly, $u^\star = \left| \begin{array}{ll} 1, & x_0 \in]-\infty,-1[\bigcup[0,1[\\ -1, & x_0 \in]-1,0]\bigcup]1,+\infty[\end{array} \right.$ is optimal

and the value function is $V(x) = ||x|-1|$ which is piecewise C^1, and fullfills

$$\begin{cases} 1 - |V'(x)| = 0, & x \in \mathbb{R} \setminus \{-1,1\} \\ V(-1) = V(1) = 0 \end{cases}$$

at its differentable points.

Let $w(\cdot)$ be piecewise constant in $L_1(\mathbb{R}, \{-1,1\})$ such that

$$\int_{-1}^{1} w(\xi) d\xi = 0 \ .$$

Then $W(x) = \displaystyle\int_{-1}^{x} w(\xi) d\xi$ is piecewise C^1 and satisfies also the HJB equation at differentiable points, with $W(-1) = W(1) = 0$.

Open-loop and closed-loop controls

Open-loops :

$$\mathcal{U}_{[t_0,T]} := \{u(\cdot) \text{ meas. } [t_0, T] \mapsto \mathrm{U}\}$$

Closed-loops :
$\psi : (-\infty, T] \times \mathrm{X} \mapsto \mathrm{U}$ is an **admissible** **state feedback** if for any (t_0, x_0), there exists an **unique absolutely continuous** sol. of

$$\dot{x} = f(x, \psi(t, x))\,,\; x(t_0) = x_0$$

For any $(t_0, x_0) \in (-\infty, T] \times \mathrm{X}$, the time function $u(\cdot) \in \mathcal{U}_{[t_0,T]}$ s.t.

$$u(t) = \psi(t, x(t)) \text{ p.p. } t \in [t_0, T]$$

is called an **open-loop representation** (denoted $u_{t_0,x_0,\psi}$).

$\mathbf{C^1}$ Sufficient Optimality Conditions

Proposition. If there exists a C^1 function W sol. of

$$\begin{cases} \partial_t W(t,x) + \bar{H}(x, \nabla W(t,x)) = 0 \ , \quad \forall (t,x) \in (-\infty, T] \times \mathsf{X} \\ W(T,x) = \phi_T(x) \ , \quad \forall x \in \mathsf{X} \end{cases}$$

and an admissible state-feedback $\psi^\star$ s.t.

$$\bar{H}(x, \nabla W(t,x)) = \partial_x W(t,x) f(x, \psi^\star(t,x)) \, , \quad \forall (t,x) \in (-\infty, T] \times \mathsf{X}$$

then W is the value function.

Moreover, for any $(t_0, x_0) \in (-\infty, T] \times \mathsf{X}$,

$$u^\star(t) = u_{t_0, x_0, \psi^\star}(t) \, , \quad \forall t \in [t_0, T]$$

is an open-loop control that realizes the minimum of $J(t_0, x_0, u(\cdot))$.

The Linear-Quadratic case

$$\dot{x} = Ax + Bu,$$

$$\inf_{u(\cdot)\in\mathcal{U}_{[t_0,T]}} \frac{1}{2}x^t(T)Sx(T) + \frac{1}{2}\int_{t_0}^{T} x^t(\tau)Qx(\tau) + u^t(\tau)Ru(\tau)\, d\tau$$

with S, Q symmetric positive and R symmetric **definite** positive.

Proposition.

1. There exists an unique C^1 symmetric positive solution of the **backward Riccati equation**

$$\dot{P} + PA + A^tP + Q - PBR^{-1}B^tP = 0\,, \quad P(T) = S(T)$$

2. The value function is $V(t,x) = \frac{1}{2}x^tP(t)x$ and an optimal *admissible* feedback is $\psi^\star(t,x) = -R^{-1}B^tP(t)x$

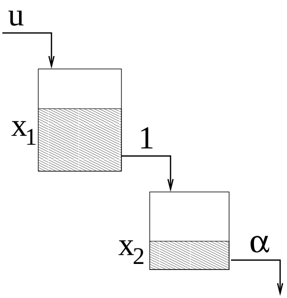
Remark. Q definite positive $\Rightarrow$ $P(t)$ definite positive for any $t < T$

Example of lake regulation

The case of double lake or cascade-lake :

Objective : drive the water levels to desired reference values without *too large* excursions of the input variable.

Modelling of cascade reservoirs



$$\begin{cases} \dot{x}_1 & = & -x_1 + u \\ \dot{x}_2 & = & x_1 - \alpha x_2 \end{cases}$$

Fix a reference point $(\bar{x}_1, \bar{x}_2, \bar{u})$ and consider the problem

$$\inf_{u(\cdot)} \frac{1}{2\varepsilon} \left((x_1(T) - \bar{x}_1)^2 + (x_2(T) - \bar{x}_2)^2 \right) + \frac{1}{2} \int_0^T (u(\tau) - \bar{u})^2 \, d\tau$$

Optimal control of cascade reservoirs

$$\dot{z} = \underbrace{\begin{bmatrix} -1 & 0 \\ 1 & -\alpha \end{bmatrix}}_{A} z + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{B} v \quad \text{with } z = \begin{bmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \end{bmatrix} \text{ and } v = u - \bar{u}$$

$$\frac{1}{2} z^t(T) \underbrace{\begin{bmatrix} 1/\varepsilon & 0 \\ 0 & 1/\varepsilon \end{bmatrix}}_{S} z(T) + \frac{1}{2}\int_0^T z^t(\tau) \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}}_{Q} z(\tau) + v(\tau) \underbrace{1}_{R} v(\tau)\, d\tau$$

$$\text{Riccati eq. : } \dot{P} + PA + A^tP + Q - PBR^{-1}B^tP = 0, \; P(T) = S(T)$$

$$\Rightarrow \begin{cases} \dot{p}_1 = 2(p_1 - p_{12}) + p_1^2\,, & p_1(T) = 1/\varepsilon \\ \dot{p}_{12} = (-p_2 + p_{12} + \alpha p_{12}) + p_{12}p_1\,, & p_{12}(T) = 0 \\ \dot{p}_2 = 2\alpha p_2 + p_{12}^2\,, & p_2(T) = 1/\varepsilon \end{cases}$$

$$\Rightarrow v^\star(t,z) = -R^{-1}B^tP(t)z = p_1(t)z_1 - p_{12}(t)z_2$$

i.e. $\boxed{u^\star(t,x_1,x_2) = \bar{u} - p_1(t)(x_1 - \bar{x}_1) - p_{12}(t)(x_2 - \bar{x}_2)}$ is optimal

Numerical simulations

Launch the Scilab script

`lac_optimal.sce`

with different choice of penality $1/\epsilon$.

The tracking problem

Given a reference output function $\bar{y}(\cdot)$, minimize the criterion

$$\frac{1}{2}x^t(T)Sx(T)+\frac{1}{2}\int_{t_0}^{T}(Cx(\tau)-\bar{y}(\tau))^tQ(Cx(\tau)-\bar{y}(\tau))+u^t(\tau)Ru(\tau)\,d\tau$$

that can be re-written as follows

$$\frac{1}{2}x^t(T)Sx(T)+$$

$$\frac{1}{2}\int_{t_0}^{T}\begin{bmatrix} x(\tau) \\ 1 \end{bmatrix}^t \underbrace{\begin{bmatrix} C^tQC & -C^tQ\tilde{y}(\tau) \\ -\tilde{y}^t(\tau)QC & \tilde{y}^t(\tau)Q\tilde{y}(\tau) \end{bmatrix}}_{\tilde{Q}(\tau)} \begin{bmatrix} x(\tau) \\ 1 \end{bmatrix} + u^t(\tau)Ru(\tau)\,d\tau$$

$\rightarrow$ we extend the dynamics : $\tilde{x} = \begin{bmatrix} x \\ z \end{bmatrix}$ with $\begin{cases} \dot{z} = 0 \\ z(0) = 1 \end{cases}$

The tracking problem

$$\text{With } \tilde{A} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad \tilde{S} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{R} = R,$$

we associate the Riccati equation

$$\dot{\tilde{P}} + \tilde{P}\tilde{A} + \tilde{A}^t\tilde{P} + \tilde{Q}(t) - \tilde{P}\tilde{B}R^{-1}\tilde{B}^tP = 0\,, \quad \tilde{P}(T) = \tilde{S}(T)$$

and write its solution as $\tilde{P}(t) = \begin{bmatrix} P(t) & \beta(t) \\ \beta^t(t) & \gamma(t) \end{bmatrix}$

$$\Rightarrow \begin{cases} \dot{P} + PA + A^tP + C^tQC - PBR^{-1}B^tP = 0\,, & P(T) = S(T) \\ \dot{\beta} + A^t\beta - PBR^{-1}B^t\beta - C^tQ\bar{y}(t) = 0\,, & \beta(T) = 0 \end{cases}$$

and $u^\star(t,x) = -\tilde{R}^{-1}\tilde{B}^t\tilde{P}(t)\tilde{x} = -R^{-1}B^tP(t)x - R^{-1}B^t\beta(t)$ is an optimal (time-dependent) feedback.

Tracking of cascade reservoirs

Let $\tilde{x}_2(t) = x_2(0) + (\bar{x}_2 - x_2(0))\dfrac{t}{T}$ be a desired output reference.

We consider the doubly penalized criterion :

$$\frac{1}{2\varepsilon}(z_1(T)^2 + z_2(T)^2) + \frac{1}{2}\int_0^T \sigma(z_2(\tau) - c(\tau))^2 + v(\tau)^2 \, d\tau$$

$$\text{with } c(t) = z_2(0)\left(1 - \frac{t}{T}\right).$$

$$\Rightarrow \begin{cases} \dot{p}_1 = 2(p_1 - p_{12}) + p_1^2 \,, & p_1(T) = 1/\varepsilon \\ \dot{p}_{12} = (-p_2 + p_{12} + \alpha p_{12}) + p_{12}p_1 \,, & p_{12}(T) = 0 \\ \dot{p}_2 = 2\alpha p_2 + p_{12}^2 - \sigma \,, & p_2(T) = 1/\varepsilon \\ \dot{\beta}_1 = (1 + p_1)\beta_1 - \beta_2 \,, & \beta_1(T) = 0 \\ \dot{\beta}_2 = p_{12}\beta_1 + \alpha\beta_2 + \sigma\tilde{x}_2(t) \,, & \beta_2(T) = 0 \end{cases}$$

and $\boxed{u^\star(t, x_1, x_2) = \bar{u} - p_1(t)(x_1 - \bar{x}_1) - p_{12}(t)(x_2 - \bar{x}_2) - \beta_1(t)}$

Numerical simulations

Launch the Scilab script

`lac_suivi.sce`

with different choice of penalities $1/\epsilon$ *and* σ.

Deterministic Kalman filter

Problem. Given an observation $y(\cdot)$, find an estimation $\hat{x}(t)$ of $x(t)$ as a solution of the *corrected* dynamics

$$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu(t) \\ \hat{x}(0) = \hat{x}_0 \end{cases}$$

such that

- the error $e(\cdot) = y(\cdot) - C\hat{x}(\cdot)$
- the correction term $u(\cdot)$
- the norm of the unknown $\hat{x}_0$

are small ?

$$\rightarrow \quad \min_{\hat{x}_0, u(\cdot)} \hat{x}_0^t S \hat{x}_0 + \int_0^T e(\tau)^t Q e(\tau) + u(\tau)^t R u(\tau) d\tau$$

Deterministic Kalman filter

$$\min_{u(\cdot)} \ \hat{x}_0^t S \hat{x}_0 + \int_0^T e(\tau)^t Q e(\tau) + u(\tau)^t R u(\tau) d\tau$$

is a tracking problem in reverse time.

Denote $\tilde{x}(t) = \hat{x}(T-t)$ and $\tilde{V}$ the value function :

$$\tilde{V}(0, x_0) = \tilde{x}_0^t \tilde{P}(0) \tilde{x}_0 + 2\tilde{x}_0^t \tilde{\beta}(0) + \tilde{\gamma}(0)$$

$$Q, S >> 0 \Rightarrow \tilde{x}_0^\star \in \arg\min_{x_0} \tilde{V}(0, x_0) \text{ satisfies } \tilde{x}_0^\star = -\tilde{P}(0)^{-1}\tilde{\beta}(0)$$

Then, $\hat{x}^\star(t) = -\tilde{P}(T-t)^{-1}\tilde{\beta}(T-t)$, with

$$\begin{cases} \dot{\tilde{P}} - \tilde{P}A - A^t\tilde{P} + C^tQC + \tilde{P}BR^{-1}B^t\tilde{P} = 0\,, & \tilde{P}(T) = S \\ \dot{\tilde{\beta}} - A^t\tilde{\beta} + \tilde{P}BR^{-1}B^t\tilde{\beta} + C^tQy(T-t)\,, & \tilde{\beta}(T) = 0 \end{cases}$$

is the best estimator.

Recursive Kalman filter

$\Pi(t) = \tilde{P}(T-t)^{-1}$ is solution of the **dual Riccati equation** :

$$\dot{\Pi} - A\Pi - A^t\Pi + \Pi C^t Q C \Pi - \Pi R^{-1}\Pi = 0, \quad \Pi(0) = S^{-1}$$

and one has :

$$\begin{aligned} \frac{d}{dt}\hat{x}^\star(t) &= -\dot{\Pi}(t)\tilde{\beta}(T-t) + \Pi(t)\dot{\tilde{\beta}}(T-t) \\ &= (A - \Pi(t)C^t QC)\hat{x}^\star(t) + \Pi(t)C^t Q y(t) \\ &= A\hat{x}^\star(t) + L(t)(C\hat{x}^\star(t) - y(t)) \end{aligned}$$

where $L(t) = -\Pi(t)C^t Q$.

The Infinite-Horizon case

$$\dot{x} = Ax + Bu, \qquad \inf_{u(\cdot)\in\mathcal{U}_{[0,+\infty)}} \lim_{T\to+\infty} \int_0^T x^t(\tau)Qx(\tau) + u^t(\tau)Ru(\tau)\, d\tau$$

Proposition.
1. If (A, B) is controllable, $P_\infty = \lim_{t\to+\infty} P(-t)$ where P is the positive solution of

$$\dot{P} + PA + A^tP + Q - PBR^{-1}B^tP = 0\,, \quad P(0) = S(0)$$

is the smallest positive solution of the algebraic Riccati equation

$$PA + A^tP + Q - PBR^{-1}B^tP = 0$$

Then, $V(x) = x'P_\infty x$ and $\psi^\star(x) = -R^{-1}B^tP_\infty x$ are the value function and optimal feedback of the infinite horizon problem.

2. If Q is definite positive or $Q = C^tC$ with (A, C) observable, then any optimal trajectory satisfies $\lim_{t\to+\infty} x^\star(t) = 0$.

Link with the Pontryagin Maximum Principle

Assume that V is C^2, and consider an optimal trajectory $x^\star(\cdot)$:

$$\partial_t V(t, x^\star(t)) + H(x^\star(t), \nabla V(t, x^\star(t)), u^\star(t)) = 0 \ ,$$
$$\partial_t V(t, \xi) + H(\xi, \nabla V(t, \xi), u^\star(t)) \geq 0 \, , \quad \forall \xi \in \mathrm{X} \ .$$

i.e. $x^\star(t)$ minimizes $\Gamma : \xi \mapsto \partial_t V(t, \xi) + H(\xi, \partial_x V(t, \xi)^t, u^\star(t))$

If $x(t) \in \mathrm{int}\ \mathrm{X}$ and $t < T$, one has $\partial_x \Gamma(x^\star(t)) = 0$.

Let $p(t) = \nabla V(t, x^\star(t))$. Then

$$\partial_x \Gamma(x^\star(t)) = 0 \quad \Longleftrightarrow \quad \dot{p}(t) = -\partial_x H(x^\star(t), p(t), u^\star(t))^t$$

$$V(T, x) = \phi_T(x) \quad \Longrightarrow \quad p(T) = \partial_x \phi(x^\star(T))^t$$

References

- D. BERTSEKAS . Dynamic Programming and Optimal Control. 4th Edition, Athena Scientific, 2012.
- D. LIBERZON. Calculus of Variations and Optimal Control Theory : A Concise Introduction. Princeton University Press, 2012.