Markov jump linear systems
Optimal Control

Pantelis Sopasakis

IMT Institute for Advanced Studies Lucca

February 5, 2016

Abbreviations

1. MJLS: Markov Jump Linear Systems
2. FHOC: Finite Horizon Optimal Control
3. IHOC: Infinite Horizon Optimal Control
4. CARE: Coupled Algebraic Ricatti Equations

Outline

1. LQR (deterministic case) – A quick revision
2. FHOC for MJLS
3. IHOC for MJLS (CARE)

I. Dynamic programming

Finite horizon optimal control

We have a (deterministic) LTI system

$$x(k+1) = Ax(k) + Bu(k),$$

with $x(0) = x_0$. For a given sequence of input values of length N, that is, $\pi_N = (u(0), u(1), \ldots, u(N-1))$ we define the cost function

$$J_N(\pi_N; x_0) = \sum_{k=0}^{N-1} \ell(x(k), u(k)) + \ell_N(x_N).$$

Assume

$$\ell(x,u) = x'Qx + u'Ru, \text{ and } \ell_N(x) = x'P_N x.$$

for some $Q \in S^n_+$, $P_f \in S^n_{++}$, $R \in S^m_{++}$.

Finite horizon optimal control

We need to determine a finite sequence π_N to minimise $J_N(\pi_N)$:

$$J_N^\star(x_0) = \min_{\pi_N} J_N(\pi_N; x_0)$$

subject to the system dynamics and $x(0) = x_0$. **DP recursion**[1]:

$$\begin{aligned} V_N(x(N)) &= x(N)' P_N x(N), \\ V_k(x(k)) &= \min_{u_k} \ell(x(k), u(k)) + V_{k+1}(x(k+1)), \end{aligned}$$

for $k = N-1, \ldots, 0$.

[1]See for instance: F. Borelli, *Constrained Optimal Control of Linear and Hybrid Systems*, Springer, 2003.

Why DP?

DP facts:

- We may **decompose** a complex optimisation problem into simpler subproblems

Why DP?

DP facts:

- We may **decompose** a complex optimisation problem into simpler subproblems
- Here, we solve for one u_k at a time

Why DP?

DP facts:

- We may **decompose** a complex optimisation problem into simpler subproblems
- Here, we solve for one u_k at a time
- DP used Bellman's **principle of optimality**

Why DP?

DP facts:

- We may **decompose** a complex optimisation problem into simpler subproblems
- Here, we solve for one u_k at a time
- DP used Bellman's **principle of optimality**
- It can be applied the same way to **stochastic** optimal control problems

Why DP?

DP facts:

- We may **decompose** a complex optimisation problem into simpler subproblems
- Here, we solve for one u_k at a time
- DP used Bellman's **principle of optimality**
- It can be applied the same way to **stochastic** optimal control problems
- It is a powerful tool to study the **MSS** of MJLS and Markovian switching systems (next class)

Finite horizon optimal control

Let $\pi^\star(x_0)$ be the respective minimiser with

$$\pi^\star(x_0) = \{u^\star(1), u^\star(2), \ldots, u^\star(N-1)\}.$$

Using DP we derive

$$\begin{aligned} V_k(x) &= x' P_k x, \\ u^\star(k) &= F(P_{k+1}) x(k) \end{aligned}$$

where P_k is determined as follows:

$$P_k = A' P_{k+1} A + Q + A' P_{k+1} F(P_{k+1})$$

and

$$F(P) = -B(B'PB + R)^{-1} B'PA.$$

Infinite horizon optimal control

What happens as $N \to \infty$? Let us define

$$J_\infty(\pi; x_0) = \sum_{k=0}^{\infty} \ell(x(k), u(k)),$$

where π is a sequence of inputs $\{u(k)\}_{k \in \mathbb{N}}$. For the series to converge it is of course required that

$$\|x(k)\|^2, \|u(k)\|^2 \to 0, \text{ as } k \to \infty.$$

Infinite horizon optimal control

We can show that – under certain conditions[2] – the IHOC problem is solvable and

$$\begin{aligned} J^\star_\infty(x) &= x' P_\infty x, \\ u^\star(k) &= F(P_\infty) x(k), \end{aligned}$$

where P_∞ is a fixed point of the DP recursion of the FHOC problem (Algebraic Ricatti Equation), that is

$$P_\infty = A' P_\infty A + Q - A' P_\infty B (B' P_\infty B + R)^{-1} B' P_\infty A.$$

[2]Provided that (A, B) is stabilisable and $(Q^{1/2}, A)$ is detectable. Then the matrix $A + BF(P_\infty)$ is stable. *Proof.* See D.P. Bertsekas, Dynamic programming and optimal control, Vol. 1, 2005, Prop. 4.4.1.

End of first section

- Revision of FHOC and DP
- We solved the LQR problem

II. FHOC for MJLS

FHOC for MJLS

Consider a MJLS

$$x(k+1) = A_{\theta(k)}x(k) + B_{\theta(k)}u(k) + M_{\theta(k)} \underbrace{v(k)}_{\text{noise}},$$

with $x(0) = x_0$, and let $z(k) = C_{\theta(k)}x(k) + D_{\theta(k)}u(k)$ be the quantity that will be penalised. We define the following cost functional

$$J(\theta_0, x_0, \pi_N) := \sum_{k=0}^{N-1} \mathbb{E}\left[\|z(k)\|^2\right] + \mathbb{E}\left[x(T)'V_{\theta(N)}x(T)\right].$$

Where π_N is a **policy** $\pi = (u(0), \ldots, u(N-1))$ with

$$u(k) = \mu_k(x(k), \theta(k)).$$

FHOC assumptions

Let $\mathfrak{G}_k$ be the σ-algebra generated by $\{x(t), \theta(t); t = 0, \ldots, N-1\}$.

Assumptions on v:

1. $v(k)$ are random variables with $\mathbb{E}\left[v(k)v(k)' 1_{\{\theta(k)=i\}}\right] = \Xi_i(k)$
2. For every f, g, $f(v(k))$ and $g(\theta(k))$ are independent w.r.t $\mathfrak{G}_k$
3. $\mathbb{E}\left[v(0)x(0)' 1_{\{\theta(0)=i\}}\right] = 0$

Assumptions on $z(k)$:

1. $C_i(k)' D_i(k) = 0$ – no penalties of the form $x(k)' S_{\theta(k)} u(k)$
2. $D_i(k)' D_i(k) > 0$

Control laws and policies for MJLS

A measurable function

$$\mu : \mathbb{R}^n \times \mathcal{N} \to \mathbb{R}^m$$

is called a **control law**.

A (finite of infinite) sequence of control laws

$$\pi = \{\mu_0, \mu_1, \ldots\},$$

where μ_k is $\mathfrak{G}_k$-measurable, called a **control policy**.

FHOC – Dynamic programming recursion

To perform DP we introduce the cost functional

$$J_\kappa(\theta(\kappa), x(\kappa), u_\kappa) := \sum_{k=\kappa}^{N-1} \mathbb{E}\left[\|z(k)\|^2 \mid \mathfrak{G}_\kappa\right] + \mathbb{E}\left[x(T)' V_{\theta(N)} x(T) \mid \mathfrak{G}_\kappa\right],$$

for $\kappa \in \{0, \ldots, N-1\}$ where $u_\kappa = (u(\kappa), \ldots, u(N-1))$ so that each $u(k)$ is $\mathfrak{G}_k$-measurable. The optimal value of $J_\kappa(\theta(\kappa), x(\kappa), u_\kappa)$ is then given by

$$J_\kappa^\star(i, x) = x' X_i(\kappa) x + \alpha(\kappa),$$

where X_i is given by a Ricatti-like equation.

FHOC – Dynamic programming recursion

We have

$$J_\kappa^\star(i,x) = x' X_i(\kappa) x + \alpha(\kappa),$$

where

$$X_i(N) = V_i,$$
$$X_i(k) = A_i'\mathcal{E}(X(k+1))A_i - A_i\mathcal{E}(X(k+1))B_iF_i(X(k+1)) + C_i'C_i,$$

where $\mathcal{E}_i(X) = \sum_{j=1}^N p_{ij}X_j$, $R_i(X) := D_i'D_i + B_i'\mathcal{E}(X)B_i$ and

$$F_i(X) := -R_i^{-1}B_i'\mathcal{E}(X)A_i.$$

The respective optimisers are given by

$$u^\star(k) = F_{\theta(k)}(X(k+1))x(k).$$

End of second section

- Formulation of FHOC for MJLS considering also an additive noise term
- Control policies and control laws
- Solution of FHOC: piecewise linear control laws

$$u^{\star}(k) = \kappa(x(k), \theta(k)) = F_{\theta(k)} x(k).$$

III. IHOC for MJLS and MSS

IHOC for MJLS

Consider a MJLS without additive noise

$$x(k+1) = A_{\theta(k)}x(k) + B_{\theta(k)}u(k),$$

with $x(0) = x_0$, and let $z(k) = C_{\theta(k)}x(k) + D_{\theta(k)}u(k)$ be the quantity that will be penalised. We are now looking for *sequences* $\pi = \{u(k)\}_{k\in\mathbb{N}}$ in

$$\mathcal{U} = \left\{ \pi \,\middle|\, \begin{array}{l} u(k) \text{ is } \mathfrak{G}_k\text{-measurable}, \forall k \in \mathbb{N} \\ \lim_{k\to\infty} \mathbb{E}\left[\|x(k)\|^2\right] = 0. \end{array} \right\}$$

IHOC for MJLS

With $\pi \in \mathcal{U}$ the following is a well-defined infinite horizon cost function

$$J(\theta_0, x_0, \pi) := \sum_{k=0}^{\infty} \mathbb{E}\left[\|z(k)\|^2\right],$$

and the IHOC problem amounts to determining

$$J^\star(\theta_0, x_0) := \inf_{\pi \in \mathcal{U}} J(\theta_0, x_0, \pi),$$

and we define $\pi^\star$ to be the respective optimiser with elements

$$u^\star(k) = \psi_k(\theta(k), x(k)).$$

Objectives

1. Under what conditions does the IHOC problem have a solution?
2. How can this solution be determined?
3. Can we derive a MS-stabilising controller by solving the IHOCP?

Control CARE

Assume that there is $X \in \mathbb{H}^n_+$ satisfying the *control CARE*:

$$X_i = A_i' \mathcal{E}_i(X) A_i - A_i \mathcal{E}_i(X) B_i (D_i' D_i + B_i' \mathcal{E}_i(X) B_i)^{-1} B_i' \mathcal{E}_i(X) A_i + C_i' C_i$$

and let

$$F_i(X) := -(D_i' D_i + B_i' \mathcal{E}_i(X) B_i)^{-1} B_i' \mathcal{E}_i(X) A_i.$$

The IHOC problem solution is given by

$$u^\star(k) = F_{\theta(k)}(X) x(k)$$

and the value function is

$$J^\star(\theta_0, x_0) = \mathbb{E}\left[x_0' X_{\theta_0} x_0\right].$$

Control CARE $\Rightarrow$ MSS

The control CARE, when solvable, yields a MS-stabilising control law, i.e., the closed-loop system

$$x(k+1) = (A_{\theta(k)} + B_{\theta(k)} F_{\theta(k)}(X))x(k),$$

is mean square stable.

Solvability conditions

The following conditions entail the solvability of the control CARE:

1. (A, B) – with $A \in \mathbb{H}^n$ and $B \in \mathbb{H}^{n,m}$ – is stabilisable,
2. (C, A) – with $C \in \mathbb{H}^{n,n_z}$ is detectable.

Proof. Book of Costa *et al.*, 2005, Corollary A.16.

End of third section

- We formulated the infinite horizon optimal control problem
- The solution of IHOC produces a MS-stabilising control law
- IHOC is solved by a CARE which can be formulated as an LMI
- Solvability conditions: (A, B) is stabilisable, (C, A) is detectable

References

1. For an introduction to DP: D. P. Bertsekas, Dynamic Programming and Optimal Control. Athena Scientific, 2nd ed., 2000.
2. Chapter 4 of: O.L.V. Costa, M.D. Fragoso and R.P. Marques, Discrete-time Markov Jump Linear Systems, Springer 2005.
3. Chapter 6 of: M.H.A. Davis and R.B. Vinter, Stochastic modelling and control, Chapman and Hall, New York 1985.
4. M.D. Fragoso, Discrete-time jump LQG problem, Int. J. Systems Sci., 20(12), pp. 2539–2545, 1989.