Optimal Control of Ordinary Differential Equations SOD 311

Laurent Pfeiffer
Inria and CentraleSupélec, Université Paris-Saclay

Ensta-Paris
Paris-Saclay University
October 7, 2021

Lecture 2:
Linear-quadratic optimal control problems

- *Goal:* investigating linear-quadratic optimal control problems and their associated linear optimality system.
- *Issues:* existence of a solution, optimality conditions, connexion with the Riccati equation (a non-linear ODE which yields a feedback law).

Bibliography

The following references are related to Lecture 2:

- E. Trélat, Contrôle optimal : théorie et applications. Version électronique, 2013. (Chapitre 3).
- E. Sontag, Mathematical Control Theory, Springer, 1998. (Chapter 8).
- U. Boscain and Y. Chitour. Introduction à l'automatique (Chapitre 5) / Introduction to automatic control (Chapter 8). Available on U. Boscain's webpage.

LQ problem

Consider the **linear-quadratic (LQ) optimization problem**:

$$\inf_{x \in \mathbb{R}^n} f(x) := \frac{1}{2} \langle x, Qx \rangle, \quad \text{subject to: } Ax = b. \qquad (P[b])$$

Data:

- $Q \in \mathbb{R}^{n \times n}$, assumed symmetric positive definite
- $A \in \mathbb{R}^{m \times n}$, assumed surjective (i.e. $\text{rank}(A) = m$)
- $b \in \mathbb{R}^m$.

First goal: characterizing the solution to $(P[b])$ with a linear system (the optimality system), analyzing this system.

Second goal: extending the techniques for solving $(P[b])$ to a linear-quadratic optimal control problem.

LQ problem

Consider the **linear-quadratic (LQ) optimization problem**:

$$\inf_{x \in \mathbb{R}^n} f(x) := \frac{1}{2}\langle x, Qx \rangle, \quad \text{subject to: } Ax = b. \qquad (P[b])$$

Data:

- $Q \in \mathbb{R}^{n \times n}$, assumed symmetric positive definite
- $A \in \mathbb{R}^{m \times n}$, assumed surjective (i.e. $\text{rank}(A) = m$)
- $b \in \mathbb{R}^m$.

First goal: characterizing the solution to $(P[b])$ with a linear system (the optimality system), analyzing this system.

Second goal: extending the techniques for solving $(P[b])$ to a linear-quadratic optimal control problem.

Elementary remarks

Lemma 1

The map f is strictly convex *and continuously* differentiable, *with*

$$\nabla f(x) = Qx.$$

Proof. Let $x \in \mathbb{R}^n$, let $\delta x \in \mathbb{R}^n$. We have:

$$\begin{aligned} f(x+\delta x) &= \frac{1}{2}\Big(\langle x, Qx \rangle + \langle \delta x, Qx \rangle + \langle x, Q\delta x \rangle + \langle \delta x, Q\delta x \rangle \Big) \\ &= f(x) + \langle Qx, \delta x \rangle + \frac{1}{2} \underbrace{\langle \delta x, Q\delta x \rangle}_{=\mathcal{O}(\|\delta x\|^2)}. \end{aligned}$$

Thus, $\nabla f(x) = Qx$ and $f(x+\delta x) > f(x) + \langle \nabla f(x), \delta x \rangle$ if $\delta x \neq 0$. Therefore f is strictly convex.

Elementary remarks

Lemma 2

Let $\alpha > 0$ denote the smallest eigenvalue of Q. Then

$$\langle x, Qx\rangle \geq \alpha \|x\|^2, \quad \forall x \in \mathbb{R}^n.$$

Proof. By the spectral theorem, there exists an orthonormal basis $(e_i)_{i=1,\dots,n}$ of eigenvectors of Q. Let $(\lambda_i)_{i=1,\dots,n}$ denote the associated eigenvalues.

Let $x = \sum_{i=1}^n x_i e_i$. We have

$$\langle x, Qx\rangle = \left\langle \sum_{i=1}^n x_i e_i, \sum_{i=1}^n \lambda_i x_i e_i \right\rangle = \sum_{i=1}^n \lambda_i x_i^2 \geq \alpha \sum_{i=1}^n x_i^2 = \alpha \|x\|^2.$$

Elementary remarks

Lemma 3

There exists a matrix $\tilde{A} \in \mathbb{R}^{m\times n}$ such that

$$A\tilde{A} = I.$$

Proof. Let $(e_i)_{i=1,\dots,m}$ denote a basis of $\mathbb{R}^m$. For all $i = 1, \dots, m$, let u_i be such that $Au_i = e_i$. Given $x = \sum_{i=1}^m x_i e_i$, define

$$\tilde{A}x = \sum_{i=1}^m x_i u_i.$$

Obviously $\tilde{A}$ is linear and

$$A\tilde{A}x = \sum_{i=1}^m x_i A u_i = \sum_{i=1}^m x_i e_i = x.$$

Elementary remarks

Corollary 4

There exists a constant M_1 such that for all $b \in \mathbb{R}^m$, there exists $x \in \mathbb{R}^n$ satisfying

$$Ax = b \quad \text{and} \quad \|x\| \le M_1 \|b\|.$$

Proof. Take $x = \tilde{A}b$ and $M_1 = \|\tilde{A}\|$.

Existence of a solution

Lemma 5

For all $b \in \mathbb{R}^m$, the problem $(P[b])$ has a unique solution $\bar{x}[b]$. *Moreover, there exists a constant $M_2 > 0$, depending only on Q and A, such that*

$$\|\bar{x}[b]\| \leq M_2 \|b\|.$$

Proof. Let $z \in \mathbb{R}$ denote the value of problem $(P[b])$. Let $\tilde{x} = \tilde{A}b$. Let $(x_k)_{k \in \mathbb{N}}$ be a **minimizing sequence**, i.e. a sequence such that

$$Ax_k = b, \quad \forall k \in \mathbb{N} \quad \text{and} \quad f(x_k) \to z.$$

Without loss of generality, we assume that $f(x_k) \leq f(\tilde{x})$. We have

$$\frac{1}{2}\alpha \|x_k\|^2 \leq f(x_k) \leq f(\tilde{x}) \leq \frac{1}{2}\|Q\| \cdot \|\tilde{x}\|^2 \leq \frac{1}{2}\|Q\| \big(\|\tilde{A}\| \cdot \|b\|\big)^2.$$

Existence of a solution

It follows that

$$\|x_k\| \leq \underbrace{\left(\frac{\|Q\|}{\alpha}\right)^{1/2}\|\tilde{A}\|}_{=:M_2} \cdot \|b\|.$$

By the **Bolzano-Weierstrass theorem**, there exists an accumulation point $\bar{x}[b]$ such that

$$A\bar{x}[b] = b, \quad f(\bar{x}[b]) = z, \quad \|\bar{x}[b]\| \leq M_2\|b\|.$$

Thus $\bar{x}[b]$ is **optimal**.

Uniqueness: follows of the strict convexity of f and the linearity of the constraints.

Optimality conditions

Lemma 6

For all $b \in \mathbb{R}^m$, there exists a unique $\lambda[b] \in \mathbb{R}^m$ such that

$$Q\bar{x}[b] + A^\top \lambda[b] = 0.$$

Moreover, $(\bar{x}[b], \lambda[b])$ is the unique solution to the following linear system:

$$\begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}.$$

The variable $\lambda[b]$ is referred to as **Lagrange multiplier**.

The above system is referred to as the **optimality system**.

Optimality conditions

Proof. *Step 1:* existence of the Lagrange multiplier.
Define the **Lagrangian**

$$L(x,\lambda) = f(x) + \langle \lambda, Ax - b \rangle = f(x) + \langle A^\top \lambda, x \rangle - \langle \lambda, b \rangle.$$

By the **Karush-Kuhn-Tucker conditions**, $\exists\, \lambda[b]$ such that

$$0 = \nabla_x L(\bar{x}[b], \lambda[b]) = Q\bar{x}[b] + A^\top \lambda[b].$$

Step 2: uniqueness of the Lagrange multiplier.
A direct consequence of the injectivity of $A^\top$.

Optimality conditions

Step 3: uniqueness of the solution to the optimality system.

Take a pair (x, λ) such that $\begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$.

Let x' be such that $Ax' = b$. By the convexity of the Lagrangian with respect to its first variable, we have

$$f(x') = L(x', \lambda) \geq \underbrace{L(x, \lambda)}_{=f(x)} + \langle \underbrace{\nabla_x L(x, \lambda)}_{=Qx+A^\top \lambda=0}, x' - x \rangle = f(x).$$

Therefore x is optimal for $(P[b])$ and thus $x = \bar{x}[b]$ and $\lambda = \lambda[b]$.

Analytic solution

Lemma 7

For all $b \in \mathbb{R}^m$, we have

$$\begin{pmatrix} \bar{x}[b] \\ \lambda[b] \end{pmatrix} = \begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ b \end{pmatrix} = \begin{pmatrix} Q^{-1}A^\top(AQ^{-1}A^\top)^{-1}b \\ -(AQ^{-1}A^\top)^{-1}b. \end{pmatrix}$$

Proof. For all (x, λ), we have

$$\left\{ \begin{aligned} Qx + A^\top\lambda &= 0 \\ Ax &= b \end{aligned} \right. \iff \left\{ \begin{aligned} x &= -Q^{-1}A^\top\lambda \\ -AQ^{-1}A^\top\lambda &= b \end{aligned} \right.$$

$$\iff \left\{ \begin{aligned} x &= Q^{-1}A^\top(AQ^{-1}A^\top)^{-1}b \\ \lambda &= -(AQ^{-1}A^\top)^{-1}b. \end{aligned} \right.$$

Note that $\tilde{Q} = AQ^{-1}A^\top$ is sym. positive definite (thus regular).

Value function

Lemma 8

For all $b \in \mathbb{R}^m$,

$$V(b) := f(\bar{x}[b]) = \frac{1}{2}\langle b, (AQ^{-1}A^\top)^{-1} b\rangle,$$

$$\nabla V(b) = -\lambda[b].$$

Proof. Direct calculation following Lemma 7. We have

$$\begin{aligned} V(b) &= \frac{1}{2}\langle Q^{-1}A^\top \tilde{Q}^{-1} b, QQ^{-1}A^\top \tilde{Q}^{-1} b\rangle \\ &= \frac{1}{2}\langle b, \tilde{Q}^{-1}(AQ^{-1}A^\top)\tilde{Q}^{-1} b\rangle = \frac{1}{2}\langle b, \tilde{Q}^{-1} b\rangle. \end{aligned}$$

Thus, $\nabla V(b) = \tilde{Q}^{-1} b = -\lambda[b]$.

Summary

Remember:

- Problem $(P[b])$ has a **unique solution** with a unique associated Lagrange multiplier.
- The pair $(\bar{x}[b], \lambda[b])$ is characterized by a **well-posed linear system**.
- In the analytical expression, we get a relation between $\lambda[b]$ and b, involving a **symmetric matrix**.

We adapt next the previous analysis to LQ optimal control problems.

Linear quadratic optimal control

Consider the following LQ optimal control problem:

$$\inf_{\substack{y\in H^1(0,T;\mathbb{R}^n)\\ u\in L^2(0,T;\mathbb{R}^m)}} \quad \frac{1}{2}\int_0^T \Big(\langle y(t), Wy(t)\rangle + \|u(t)\|^2 \Big)\,\mathrm{d}t + \frac{1}{2}\langle y(T), Ky(T)\rangle$$

$$\text{subject to:} \quad \begin{cases} \dot{y}(t) = & Ay(t) + Bu(t) \\ y(0) = & y_0. \end{cases} \qquad (P(y_0))$$

Data and assumptions:

- Time horizon: $T > 0$.
- Dynamics coefficients: $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$
- Cost coefficients: $W \in \mathbb{R}^{n\times n}$ and $K \in \mathbb{R}^{n\times n}$, both assumed symmetric positive semi-definite.

The initial condition $y_0 \in \mathbb{R}^n$ is seen as a *parameter* of the problem.

The generic constant M

Convention.

All constants M appearing in forthcoming lemmas will depend on A, B, W, K, and T only. They **will not depend** on y_0.

We use **the same name** for (a priori) independent constants. This is acceptable in so far as all statements remain true if the value of M is **increased**.

The Sobolev space $H^1(0, T; \mathbb{R}^m)$

The space $H^1(0, T; \mathbb{R}^n)$ is defined as follows:

$$H^1(0, T; \mathbb{R}^n) = \left\{ y \in L^2(0, T; \mathbb{R}^n) \,|\, \dot{y} \in L^2(0, T; \mathbb{R}^n) \right\}$$

where $\dot{y}$ denotes the weak derivative of y. It is a **Hilbert space**, equipped with the scalar product:

$$\langle y_1, y_2 \rangle = \int_0^T \langle y_1(t), y_2(t) \rangle \, \mathrm{d}t + \int_0^T \langle \dot{y}_1(t), \dot{y}_2(t) \rangle \, \mathrm{d}t$$

and the norm $\|y\|_{H^1(0,T;\mathbb{R}^n)} = \left(\|y\|^2_{L^2(0,T;\mathbb{R}^n)} + \|\dot{y}\|^2_{L^2(0,T;\mathbb{R}^n)} \right)^{1/2}$.

Lemma 9

The space $H^1(0, T; \mathbb{R}^m)$ is contained in the set of continuous functions from $[0, T]$ to $\mathbb{R}^n$. Moreover, all usual calculus rules are valid (in particular, integration by parts*).*

State equation

Given $u \in L^2(0,T;\mathbb{R}^m)$ and $y_0 \in \mathbb{R}^n$, let $y[u,y_0] \in H^1(0,T;\mathbb{R}^n)$ denote the solution to the state equation

$$\dot{y}(t) = Ay(t) + Bu(t), \quad y(0) = y_0.$$

Lemma 10

The map $(u,y_0) \in L^2(0,T;\mathbb{R}^m) \times \mathbb{R}^n \mapsto y[u,y_0] \in H^1(0,T;\mathbb{R}^n)$ *is* linear. *There exists* $M > 0$ *such that for all* $u \in L^2(0,T;\mathbb{R}^m)$ *and for all* $y_0 \in \mathbb{R}^n$,

$$\|y[u,y_0]\|_{L^\infty(0,T;\mathbb{R}^n)} \leq M\big(\|y_0\| + \|u\|_{L^2(0,T;\mathbb{R}^n)}\big),$$
$$\|y[u,y_0]\|_{H^1(0,T;\mathbb{R}^n)} \leq M\big(\|y_0\| + \|u\|_{L^2(0,T;\mathbb{R}^n)}\big).$$

Proof. A direct application of Duhamel's formula and Cauchy-Schwarz inequality.

Reduced problem

Let $J\colon L^2(0,T;\mathbb{R}^m) \to \mathbb{R}$ be defined by

$$J(u) = J_1(u) + J_2(u) + J_3(u),$$

where

$$J_1(u) = \frac{1}{2}\int_0^T \langle y[u,y_0](t), Wy[u,y_0](t)\rangle \,\mathrm{d}t$$

$$J_2(u) = \frac{1}{2}\int_0^T \|u(t)\|^2 \,\mathrm{d}t$$

$$J_3(u) = \frac{1}{2}\langle y[u,y_0](T), Ky[u,y_0](T)\rangle.$$

Consider the **reduced problem**, equivalent to $(P(y_0))$,

$$\inf_{u\in L^2(0,T;\mathbb{R}^m)} J(u). \qquad (P'(y_0))$$

Weak lower semi-continuity

Definition 11

A map $F\colon L^2(0,T;\mathbb{R}^m) \to \mathbb{R}$ is said to be **weakly lower semi-continuous** (resp. **weakly continuous**) if for any weakly convergent sequence $(u_k)_{k\in\mathbb{N}}$ with weak limit $\bar{u}$, it holds

$$F(\bar{u}) \leq \liminf_{k\in\mathbb{N}} F(u_k) \quad \Big(\text{resp. } F(\bar{u}) = \lim_{k\in\mathbb{N}} F(u_k)\Big).$$

Lemma 12

The map J is strictly convex *and* weakly lower semi-continuous.

Proof.

- J_1, J_2, and J_3 are convex, J_2 is strictly convex
- J_1 and J_3 are weakly continuous, J_2 is weakly lower semi-continuous

Regularity of J

Let $(u_k)_{k\in\mathbb{N}}$ be a sequence in $L^2(0,T;\mathbb{R}^m)$, let $\bar{u}\in L^2(0,T;\mathbb{R}^m)$. Assume that $u_k \rightharpoonup \bar{u}$. Let $y_k = y[u_k,y_0]$ and $\bar{y} = y[\bar{u},y_0]$. Then,

- $(u_k)_{k\in\mathbb{N}}$ is bounded in $L^2(0,T;\mathbb{R}^m)$
- by Lemma 10, y_k is bounded in $L^\infty(0,T;\mathbb{R}^n)$.

With the help of Duhamel's formula, we obtain that

$$y[u_k,y_0](t) \to y[\bar{u},y_0](t), \quad \text{for all } t\in[0,T].$$

Step 1: This directly implies that

$$J_3(u_k) = \frac{1}{2}\langle y_k(T), K y_k(T)\rangle \to \frac{1}{2}\langle \bar{y}(T), K\bar{y}(T)\rangle = J_3(\bar{u}).$$

Thus J_3 is weakly continuous.

Regularity of J

Let $(u_k)_{k\in\mathbb{N}}$ be a sequence in $L^2(0,T;\mathbb{R}^m)$, let $\bar{u} \in L^2(0,T;\mathbb{R}^m)$. Assume that $u_k \rightharpoonup \bar{u}$. Let $y_k = y[u_k, y_0]$ and $\bar{y} = y[\bar{u}, y_0]$. Then,

- $(u_k)_{k\in\mathbb{N}}$ is bounded in $L^2(0,T;\mathbb{R}^m)$
- by Lemma 10, y_k is bounded in $L^\infty(0,T;\mathbb{R}^n)$.

With the help of Duhamel's formula, we obtain that

$$y[u_k, y_0](t) \to y[\bar{u}, y_0](t), \quad \text{for all } t \in [0,T].$$

Step 1: This directly implies that

$$J_3(u_k) = \frac{1}{2}\langle y_k(T), K y_k(T)\rangle \to \frac{1}{2}\langle \bar{y}(T), K\bar{y}(T)\rangle = J_3(\bar{u}).$$

Thus J_3 is weakly continuous.

Regularity of J

Let $(u_k)_{k\in\mathbb{N}}$ be a sequence in $L^2(0,T;\mathbb{R}^m)$, let $\bar{u}\in L^2(0,T;\mathbb{R}^m)$. Assume that $u_k \rightharpoonup \bar{u}$. Let $y_k = y[u_k, y_0]$ and $\bar{y} = y[\bar{u}, y_0]$. Then,

- $(u_k)_{k\in\mathbb{N}}$ is bounded in $L^2(0,T;\mathbb{R}^m)$
- by Lemma 10, y_k is bounded in $L^\infty(0,T;\mathbb{R}^n)$.

With the help of Duhamel's formula, we obtain that

$$y[u_k, y_0](t) \to y[\bar{u}, y_0](t), \quad \text{for all } t\in[0,T].$$

Step 1: This directly implies that

$$J_3(u_k) = \frac{1}{2}\langle y_k(T), K y_k(T)\rangle \to \frac{1}{2}\langle \bar{y}(T), K\bar{y}(T)\rangle = J_3(\bar{u}).$$

Thus J_3 is weakly continuous.

Regularity of J

Step 2: By the dominated convergence theorem,

$$J_1(u_k) = \frac{1}{2}\int_0^T \langle y_k(t), W y_k(t)\rangle \, \mathrm{d}t \to \frac{1}{2}\int_0^T \langle \bar{y}(t), W\bar{y}(t)\rangle \, \mathrm{d}t = J_1(\bar{u}).$$

Step 3: Finally, we have:

$$\begin{aligned} J_2(u_k) - J_2(\bar{u}) &= \frac{1}{2}\int_0^T \|u_k(t)\|^2 - \|\bar{u}(t)\|^2 \, \mathrm{d}t \\ &= \underbrace{\int_0^T \langle \bar{u}(t), u_k(t) - \bar{u}(t)\rangle \, \mathrm{d}t}_{\to 0} + \underbrace{\frac{1}{2}\int_0^T \|u_k(t) - \bar{u}(t)\|^2 \, \mathrm{d}t}_{\geq 0}. \end{aligned}$$

Therefore, $\liminf J_2(u_k) - J_2(\bar{u}) \geq 0$ and J_2 is weakly lower semi-continuous.

Regularity of J

Step 2: By the dominated convergence theorem,

$$J_1(u_k) = \frac{1}{2}\int_0^T \langle y_k(t), W y_k(t)\rangle \,\mathrm{d}t \to \frac{1}{2}\int_0^T \langle \bar{y}(t), W\bar{y}(t)\rangle \,\mathrm{d}t = J_1(\bar{u}).$$

Step 3: Finally, we have:

$$\begin{aligned} J_2(u_k) - J_2(\bar{u}) &= \frac{1}{2}\int_0^T \|u_k(t)\|^2 - \|\bar{u}(t)\|^2 \,\mathrm{d}t \\ &= \underbrace{\int_0^T \langle \bar{u}(t), u_k(t) - \bar{u}(t)\rangle \,\mathrm{d}t}_{\to 0} + \underbrace{\frac{1}{2}\int_0^T \|u_k(t) - \bar{u}(t)\|^2 \,\mathrm{d}t}_{\geq 0}. \end{aligned}$$

Therefore, $\liminf J_2(u_k) - J_2(\bar{u}) \geq 0$ and J_2 is weakly lower semi-continuous.

Existence result

Lemma 13

For all $y_0 \in \mathbb{R}^n$, the problem $(P'(y_0))$ has a unique solution $\bar{u}[y_0]$. Moreover, there exists a constant M, independent of y_0, such that

$$\|\bar{u}[y_0]\|_{L^2(0,T;\mathbb{R}^m)} \leq M\|y_0\|.$$

Proof. Let $(u_k)_{k\in\mathbb{N}}$ be a **minimizing sequence**. W.l.o.g.,

$$\frac{1}{2}\|u_k\|^2_{L^2(0,T;\mathbb{R}^m)} = J_2(u_k) \leq J(u_k) \leq J(0) \leq \frac{1}{2}\big(M\|y_0\|\big)^2.$$

Extracting a subsequence, we can assume that $u_k \rightharpoonup \bar{u}$, for some $\bar{u} \in L^2(0,T;\mathbb{R}^m)$. We have $\|\bar{u}\|_{L^2(0,T;\mathbb{R}^m)} \leq M\|y_0\|$, moreover

$$J(\bar{u}) \leq \liminf J(u_k) = \inf_{u\in L^2(0,T;\mathbb{R}^m)} J(u).$$

Thus, $\bar{u}$ is optimal. Strict convexity of $J \Longrightarrow$ uniqueness.

Fréchet differentiability

Definition 14

The map J is said to be **Fréchet differentiable** if for any $u \in L^2(0,T;\mathbb{R}^m)$, there exists a continuous linear form $DJ(u)\colon L^2(0,T;\mathbb{R}^m) \to \mathbb{R}$ such that

$$\frac{|J(u+v) - J(u) - DJ(u)v|}{\|v\|_{L^2(0,T;\mathbb{R}^m)}} \underset{\|v\|_{L^2}\downarrow 0}{\longrightarrow} 0.$$

Remark. A sufficient condition for Fréchet differentiability is to have

$$|J(u+v) - J(u) - DJ(u)v| \leq M\|v\|^2_{L^2(0,T;\mathbb{R}^m)},$$

for all v and for some M independent of v.

Fréchet differentiability

Lemma 15

The map J is Fréchet differentiable. Let $\bar{u}$ and $v \in L^2(0,T;\mathbb{R}^m)$. Let $\bar{y} = y[\bar{u}, y_0]$ and let $z \in y[v,0]$. Omitting the time variable,

$$DJ(\bar{u})v = \int_0^T \langle W\bar{y}, z\rangle + \langle \bar{u}, v\rangle \, dt + \langle K\bar{y}(T), z(T)\rangle.$$

Proof. First, $y[u+v,y_0] - y[u,y_0] = y[v,0] = z$.
We have

$$J_1(\bar{u}+v) - J_1(\bar{u}) = \underbrace{\int_0^T \langle W\bar{y}, z\rangle \,\mathrm{d}t}_{=DJ_1(\bar{u})v} + \underbrace{\frac{1}{2}\int_0^T \langle z, Wz\rangle \,\mathrm{d}t}_{\substack{=\mathcal{O}\left(\|z\|^2_{L^\infty(0,T;\mathbb{R}^n)}\right) \\ =\mathcal{O}\left(\|v\|^2_{L^2(0,T;\mathbb{R}^m)}\right)}}.$$

Fréchet differentiability

Similarly, we have

$$J_2(\bar{u}+v) - J_2(\bar{u}) = \underbrace{\int_0^T \langle \bar{u}, v\rangle \,\mathrm{d}t}_{=DJ_2(\bar{u})v} + \frac{1}{2}\|v\|^2_{L^2(0,T;\mathbb{R}^m)}.$$

and

$$J_3(\bar{u}+v) - J_3(\bar{u}) = \underbrace{\langle K\bar{y}(T), z(T)\rangle}_{=DJ_3(\bar{u})v} + \langle z(T), Kz(T)\rangle.$$

Riesz representative

Pre-hamiltonian: given $(u, y, p) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$,

$$H(u, y, p) = \frac{1}{2}\left(\langle y, Wy\rangle + \|u\|^2\right) + \langle p, Ay + Bu\rangle \in \mathbb{R}.$$

We have

$$\nabla_y H(u, y, p) = Wy + A^\top p \quad \text{and} \quad \nabla_u H(u, y, p) = u + B^\top p.$$

Lemma 16

Let $\bar{u} \in L^2(0, T; \mathbb{R}^m)$. Let $\bar{y} = y[\bar{u}, y_0]$. Let $p \in H^1(0, T; \mathbb{R}^n)$ be the solution to

$$-\dot{p}(t) = \nabla_y H(\bar{u}(t), \bar{y}(t), p(t)), \quad p(T) = K\bar{y}(T).$$

Then,

$$DJ(\bar{u})v = \int_0^T \langle \nabla_u H(\bar{u}(t), \bar{y}(t), p(t)), v(t)\rangle dt.$$

Riesz representative

Pre-hamiltonian: given $(u, y, p) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n$,

$$H(u, y, p) = \frac{1}{2}\left(\langle y, Wy \rangle + \|u\|^2\right) + \langle p, Ay + Bu \rangle \in \mathbb{R}.$$

We have

$$\nabla_y H(u, y, p) = Wy + A^\top p \quad \text{and} \quad \nabla_u H(u, y, p) = u + B^\top p.$$

Lemma 16

Let $\bar{u} \in L^2(0, T; \mathbb{R}^m)$. Let $\bar{y} = y[\bar{u}, y_0]$. Let $p \in H^1(0, T; \mathbb{R}^n)$ be the solution to

$$-\dot{p}(t) = \nabla_y H(\bar{u}(t), \bar{y}(t), p(t)), \quad p(T) = K\bar{y}(T).$$

Then,

$$DJ(\bar{u})v = \Big\langle \nabla_u H(\bar{u}(\cdot), \bar{y}(\cdot), p(\cdot)), v \Big\rangle_{L^2(0,T;\mathbb{R}^m)}.$$

Riesz representative

Proof. We have

$$
\begin{aligned}
\langle K\bar{y}(T), z(T)\rangle &= \langle p(T), z(T)\rangle - \langle p(0), z(0)\rangle \\
&= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\langle p(t), z(t)\rangle \,\mathrm{d}t \\
&= \int_0^T \langle \dot{p}(t), z(t)\rangle + \langle p(t), \dot{z}(t)\rangle \,\mathrm{d}t \\
&= \int_0^T \langle -A^\top p - W\bar{y}, z\rangle + \langle p, Az + Bv\rangle \,\mathrm{d}t \\
&= \int_0^T -\langle W\bar{y}, z\rangle + \langle B^\top p, v\rangle \,\mathrm{d}t.
\end{aligned}
$$

Combined with Lemma 15 and the expression of $\nabla_u H(u, y, p)$, we obtain the result.

Pontryagin's principle

Theorem 17

Let $\bar{u} \in L^2(0,T;\mathbb{R}^m)$. Let $\bar{y} = y[\bar{u}, y_0]$. Let $\bar{p}$ be defined by the adjoint equation

$$\begin{aligned} -\dot{\bar{p}}(t) &= \nabla_y H(\bar{u}(t), \bar{y}(t), \bar{p}(t)) = A^\top \bar{p}(t) + W\bar{y}(t), \\ \bar{p}(T) &= K\bar{y}(T). \end{aligned}$$

Then, $\bar{u}$ is a solution *to* $(P'(y_0))$ if and only if

$$\bar{u}(t) + B^\top \bar{p}(t) = \nabla_u H(\bar{u}(t), \bar{y}(t), \bar{p}(t)) = 0, \quad \textit{for a.e. } t \in (0,T).$$

Proof. Since J is convex, $\bar{u}$ is optimal if and only if $DJ(\bar{u}) = 0$.

Remark. By convexity of $H(\cdot, \bar{y}(t), \bar{p}(t))$,

$$\nabla_u H(\bar{u}(t), \bar{y}(t), \bar{p}(t)) = 0 \iff \bar{u}(t) \in \underset{v \in \mathbb{R}^m}{\operatorname{argmin}}\ H(v, \bar{y}(t), \bar{p}(t)).$$

Estimate of p

Lemma 18

Let $\bar{u}$ denote the solution to $(P'(y_0))$, let $\bar{y} = y[\bar{u}, y_0]$, and let $\bar{p}$ be the associated costate. Then, there exists a constant M, independent of y_0, such that

$$\|\bar{p}\|_{L^\infty(0,T;\mathbb{R}^m)} \leq M\|y_0\| \quad \text{and} \quad \|\bar{p}\|_{H^1(0,T;\mathbb{R}^m)} \leq M\|y_0\|.$$

Proof. We know that

$$\|\bar{u}\|_{L^2(0,T;\mathbb{R}^m)} \leq M\|y_0\| \quad \text{and} \quad \|\bar{y}\|_{L^\infty(0,T;\mathbb{R}^m)} \leq M\|y_0\|.$$

Denote $\tilde{p}(t) = \bar{p}(T-t)$. Then $\tilde{p}$ is solution to

$$\tilde{p}(t) = A^\top \tilde{p}(t) + W\bar{y}(T-t), \quad \tilde{p}(0) = K\bar{y}(T).$$

Duhamel $\Longrightarrow$ bounds of $\tilde{p}$ in $L^\infty(0,T;\mathbb{R}^n)$ and $H^1(0,T;\mathbb{R}^n)$.

A last formula

Lemma 19

Let $\bar{u} = \bar{u}[y_0]$, let $\bar{y} = y[\bar{u}, y_0]$, and let $\bar{p}$ be the associated costate. Then,

$$V(y_0) := \Big(\inf_{u \in L^2(0,T;\mathbb{R}^m)} J(u) \Big) = J(\bar{u}) = \frac{1}{2} \langle \bar{p}(0), y_0 \rangle.$$

Proof. We have

$$\begin{aligned} 2J_3(\bar{u}) &= \langle \bar{y}(T), K\bar{y}(T) \rangle = \langle \bar{p}(T), \bar{y}(T) \rangle \\ &= \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \langle \bar{p}(t), \bar{y}(t) \rangle \, \mathrm{d}t + \langle \bar{p}(0), y_0 \rangle \, \mathrm{d}t. \end{aligned}$$

A last formula

We further have

$$
\begin{aligned}
\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\langle \bar{p}(t), \bar{y}(t)\rangle \,\mathrm{d}t &= \int_0^T \langle \dot{p}, \bar{y}\rangle + \langle \bar{p}, \dot{\bar{y}}\rangle \,\mathrm{d}t \\
&= \int_0^T \langle -A^\top \bar{p} - W\bar{y}, \bar{y}\rangle + \langle \bar{p}, A\bar{y} + B\bar{u}\rangle \,\mathrm{d}t \\
&= \int_0^T -\langle W\bar{y}, \bar{y}\rangle + \langle B^\top \bar{p}, \bar{u}\rangle \,\mathrm{d}t \\
&= \int_0^T -\langle W\bar{y}, \bar{y}\rangle - \|\bar{u}\|^2 \,\mathrm{d}t \\
&= -2J_1(\bar{u}) - 2J_2(\bar{u}).
\end{aligned}
$$

Combining the last two equalities, we obtain

$$
J(\bar{u}) = J_1(\bar{u}) + J_2(\bar{u}) + J_3(\bar{u}) = \frac{1}{2}\langle \bar{p}(0), y_0\rangle.
$$

Linear optimality system

The numerical resolution of $(P'(y_0))$ boils down to the numerical resolution of the following **linear optimality system**:

$$\begin{cases} \dot{y}(t) - Ay(t) - Bu(t) = & 0 & \text{State equation} \\ \dot{p}(t) + A^\top p(t) + Wy(t) = & 0 & \text{Adjoint equation} \\ u(t) + B^\top p(t) = & 0 & \text{Minimality condition} \\ p(T) - Ky(T) = & 0 & \text{Initial condition} \\ y(0) = & y_0. & \text{Terminal condition} \end{cases}$$

More precisely: Theorem 17 ensures that the optimality system has a unique solution, which is $(\bar{y}, \bar{u}, \bar{p})$.

Linear optimality system

After elimination of $u = -B^\top p$, we obtain the **coupled** system:

$$\left\{ \begin{aligned} \dot{y}(t) - Ay(t) + BB^\top p(t) &= 0 \\ \dot{p}(t) + A^\top p(t) + Wy(t) &= 0 \\ p(T) - Ky(T) &= 0 \\ y(0) &= y_0. \end{aligned} \right. \qquad (OS(y_0))$$

Key idea

A key idea is to **decouple** the linear system, by constructing a map

$$E\colon [0,T] \to \mathbb{R}^{n\times n},$$

independent of y_0, such that for any solution (y,p) to $(OS(y_0))$, we have

$$p(t) = -E(t)y(t).$$

Roadmap. Once E has been constructed, we have:

$$\dot{y} = Ay + Bu = Ay - BB^\top p = (A + BB^\top E)y$$

together we the initial condition $y(0) = y_0$. Thus, y can be computed by solving a linear differential system. Then, p and u are obtained via

$$p = -Ey \quad \text{and} \quad u = -B^\top p.$$

Key idea

A key idea is to **decouple** the linear system, by constructing a map

$$E \colon [0, T] \to \mathbb{R}^{n \times n},$$

independent of y_0, such that for any solution (y, p) to $(OS(y_0))$, we have

$$p(t) = -E(t)y(t).$$

Roadmap. Once E has been constructed, we have:

$$\dot{y} = Ay + Bu = Ay - BB^\top p = (A + BB^\top E)y$$

together we the initial condition $y(0) = y_0$. Thus, y can be computed by solving a linear differential system. Then, p and u are obtained via

$$p = -Ey \quad \text{and} \quad u = -B^\top p.$$

Derivation of the Riccati equation

Wanted: $p = -Ey$. The terminal condition $p(T) = Ky(T)$ yields

$$E(T) = -K.$$

Next, by differentiation, we have:

$$\dot{p} = -\dot{E}y - E\dot{y},$$

therefore,

$$\begin{aligned} -\dot{E}y &= \dot{p} + E\dot{y} \\ &= \left[-A^\top p - Wy \right] + \left[E(Ay - BB^\top p) \right] \\ &= \left[A^\top Ey - Wy \right] + \left[E(Ay + BB^\top Ey) \right] \\ &= \left(A^\top E + EA - W + EBB^\top E \right) y. \end{aligned}$$

Riccati equation

Theorem 20

There exists a unique smooth solution to the following matrix differential equation, called Riccati equation:

$$\begin{cases} -\dot{E}(t) = & A^\top E(t) + E(t)A - W + E(t)BB^\top E(t) \\ E(T) = & -K. \end{cases} \tag{RE}$$

Moreover, for all $y_0 \in \mathbb{R}^n$, the optimal trajectory $\bar{y}$ for $(P'(y_0))$ is the solution to the closed-loop system

$$\dot{y}(t) = (A + BB^\top E(t))y(t), \quad y(0) = y_0.$$

It also holds:

$$p(t) = -E(t)\bar{y}(t) \quad \text{and} \quad \bar{u}[y_0](t) = \underbrace{B^\top E(t)\bar{y}(t)}_{\textit{Feedback law!}}. \tag{1}$$

Riccati equation

Proof. Step 1. The only difficulty is to prove that (RE) is **well-posed**. Once we have a solution E, the closed-loop system and relation (1) define a triplet (u, y, p) which satisfies the linear optimality system:

- (y, u) satisfies the state equation
- u satisfies the minimality condition
- p satisfies the adjoint equation:

$$-\dot{p} = \dot{E}y + E\dot{y} = \ldots = A^\top p + Wy.$$

Thus $(u, y, p) = (\bar{u}, \bar{y}, \bar{p})$.

Riccati equation

Step 2. The Riccati equation has the abstract form:

$$-\dot{E}(t) = \mathcal{F}(E(t)), \quad E(T) = -K.$$

The map $\mathcal{F}\colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is polynomial, thus **locally Lipschitz** continuous (but not globally Lipschitz continuous!).

By the Picard-Lindelöf theorem, there exists $\tau \in [-\infty, T)$ such that (RE) has a unique solution on $(\tau, T]$. If $\tau \in \mathbb{R}$, then

$$\lim_{t \downarrow \tau} \|E(t)\| = \infty.$$

Riccati equation

Step 2. The Riccati equation has the abstract form:

$$-\dot{E}(t) = \mathcal{F}(E(t)), \quad E(T) = -K.$$

The map $\mathcal{F} \colon \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is polynomial, thus **locally Lipschitz** continuous (but not globally Lipschitz continuous!).

By the Picard-Lindelöf theorem, there exists $\tau \in [-\infty, T)$ such that (RE) has a unique solution on $(\tau, T]$. If $\tau \in \mathbb{R}$, then

$$\lim_{t \downarrow \tau} \|E(t)\| = \infty.$$

Riccati equation

Step 3. Assume that $\tau \geq 0$. Let $s \in (\tau, T]$. Let $y_s \in \mathbb{R}^n$, consider

$$\inf_{\substack{y \in H^1(s,T;\mathbb{R}^n)\\ u \in L^2(s,T;\mathbb{R}^m)}} \quad \frac{1}{2}\int_s^T \Big(\langle y(t), Wy(t) \rangle + \|u(t)\|^2 \Big)\,\mathrm{d}t + \frac{1}{2}\langle y(T), Ky(T) \rangle$$

$$\text{subject to:} \quad \begin{cases} \dot{y}(t) = & Ay(t) + Bu(t) \\ y(s) = & y_s. \end{cases}$$

Adapting the theory developed previously, we prove the existence of a unique solution $(\bar{u}, \bar{y})$ with associated costate p, such that

$$p(s) = -E(s)y_s \quad \text{and} \quad \|p(s)\| \leq M\|y_s\|. \tag{2}$$

Here the constant M is independent of y_s, $(\bar{u}, \bar{y})$ and p, it can also be shown to be independent of s.

Riccati equation

Step 3. Assume that $\tau \geq 0$. Let $s \in (\tau, T]$. Let $y_s \in \mathbb{R}^n$, consider

$$\inf_{\substack{y \in H^1(s,T;\mathbb{R}^n) \\ u \in L^2(s,T;\mathbb{R}^m)}} \quad \frac{1}{2} \int_s^T \Big(\langle y(t), Wy(t) \rangle + \|u(t)\|^2 \Big) \, \mathrm{d}t + \frac{1}{2} \langle y(T), Ky(T) \rangle$$

$$\text{subject to:} \quad \begin{cases} \dot{y}(t) = & Ay(t) + Bu(t) \\ y(s) = & y_s. \end{cases}$$

Adapting the theory developed previously, we prove the existence of a unique solution $(\bar{u}, \bar{y})$ with associated costate p, such that

$$p(s) = -E(s)y_s \quad \text{and} \quad \|p(s)\| \leq M\|y_s\|. \tag{2}$$

Here the constant M is independent of y_s, $(\bar{u}, \bar{y})$ and p, it can also be shown to be independent of s.

Riccati equation

Conclusion. Let $s > 0$ be such that

$$\|E(s)\| \geq M + 2,$$

where $\| \cdot \|$ denotes the operator norm and where M is the constant appearing in (2).

Let $y_s \in \mathbb{R}^n \backslash \{0\}$ be such that

$$\|E(s) y_s\| \geq (M+1)\|y_s\|.$$

Therefore,

$$\|p(s)\| \geq (M+1)\|y_s\| > M\|y_s\|.$$

A contradiction.

Riccati equation

Conclusion. Let $s > 0$ be such that

$$\|E(s)\| \geq M + 2,$$

where $\|\cdot\|$ denotes the operator norm and where M is the constant appearing in (2).

Let $y_s \in \mathbb{R}^n \setminus \{0\}$ be such that

$$\|E(s)y_s\| \geq (M+1)\|y_s\|.$$

Therefore,

$$\|p(s)\| \geq (M+1)\|y_s\| > M\|y_s\|.$$

A contradiction.

Additional properties

Lemma 21

1. *For all* $y_0 \in \mathbb{R}^n$,

$$V(y_0) := \Big(\inf_{u \in L^2(0,T;\mathbb{R}^m)} J(u) \Big) = -\frac{1}{2} \langle y_0, E(0) y_0 \rangle.$$

2. *For all* $t \in [0, T]$, $E(t)$ *is* symmetric negative semi-definite.
3. *For all* $y_0 \in \mathbb{R}^n$, $\nabla V(y_0) = \bar{p}(0)$.

Proof.

1. We have $V(y_0) = \frac{1}{2} \langle \bar{p}(0), y_0 \rangle = -\frac{1}{2} \langle y_0, E(0) y_0 \rangle$.
2. Verify that $E^\top$ is the solution (RE). Moreover, $V(y_0) \geq 0$.
3. We have $\nabla V(y_0) = -E(0) y_0 = \bar{p}(0)$.

Optimality system

Recall the **optimality system** to be solved:

$$\begin{cases} \dot{y} = & Ay - BB^\top p, & y(0) = & y_0, \\ \dot{p} = & -A^\top p - Wy, & p(T) = & Ky(T). \end{cases}$$

Equivalently:

$$\begin{pmatrix} \dot{y} \\ \dot{p} \end{pmatrix} = \underbrace{\begin{pmatrix} A & -BB^\top \\ W & A^\top \end{pmatrix}}_{=:R} \begin{pmatrix} y \\ p \end{pmatrix}, \quad y(0) = y_0, \quad p(T) = Ky(T).$$

The optimality system is a **two-point boundary value problem**.

If $p(0)$ was known, then the differential system could be solved numerically.

Shooting method: find $p(0)$ such that $p(T) = Ky(T)$.

Setting $\begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} = e^{TR}$, we have the equivalent formulation:

$$\begin{pmatrix} y(T) \\ p(T) \end{pmatrix} = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix} \begin{pmatrix} y(0) \\ p(0) \end{pmatrix}, \quad y(0) = y_0, \quad p(T) = Ky(T).$$

The optimality system reduces to the **shooting equation**:

$$\begin{aligned} & X_3 y_0 + X_4 p(0) = K(X_1 y_0 + X_2 p(0)) \\ & \iff p(0) = (X_4 - KX_2)^{-1}(KX_1 - X_3) y_0. \end{aligned} \qquad (SE)$$

Shooting algorithm

In the LQ case, the shooting algorithm consists then in the following steps:

- Compute e^{TR}, by solving the **matrix differential equation**

$$\dot{X}(t) = RX(t), \quad X(0) = I,$$

  in $\mathbb{R}^{2n\times 2n}$.
- Solve the **shooting equation** (SE) and find p_0.
- Solve the **differential equation**

$$\begin{pmatrix} \dot{y} \\ \dot{p} \end{pmatrix} = R \begin{pmatrix} y \\ p \end{pmatrix}, \quad \begin{pmatrix} y(0) \\ p(0) \end{pmatrix} = \begin{pmatrix} y_0 \\ p_0 \end{pmatrix}.$$

- The optimal control is given by $u = -B^\top p$.