Project Instructions for MA5232
(Part IV: Optimal Control Theory)

Due: 30 April 2022

Instructions. The project consists of some theoretical derivations and some numerical investigations. Please submit your work (derivation+numerics/plots) in pdf format and your code in a folder. You can zip the pdf and the code folder into one zip file. Pleae adhere to the following format.

```
part4_YourStudentNumber_YourName.zip
├── report_YourStudentNumber_YourName.pdf
└── src
    └── (all your code files)...
```

Project Description. In this project, we will investigate the linear quadratic regulator problem, which is one of the canonical examples for optimal control. This is well studied because of its linear structure, making exact analysis more tractable. Furthermore, due to linearization arguments, many optimal control problems can be analyzed as a linear quadratic problem near the optima.

Consider a linear ODE with condition $x_0 \in \mathbb{R}^d$

$$\dot{x}_t = A(t)x_t, \qquad A(t) \in \mathbb{R}^{d\times d}, \quad t \in [0, T]. \tag{0.1}$$

This models a generic linear dynamical system. While this equation is simple, it finds applications in a large variety of fields, because many nonlinear systems under small perturbations from its steady state can be understood via linearization.

We will consider the *controlled* version of (0.1)

$$\dot{x}_t = A(t)x_t + B(t)u_t \qquad A(t) \in \mathbb{R}^{d\times d}, \quad B(t) \in \mathbb{R}^{d\times m}, \quad u_t \in \mathbb{R}^m \quad t \in [0, T]. \tag{0.2}$$

Here, the linear dynamics can be modified by a control signal $\boldsymbol{u} \in L^\infty([0, T], \mathbb{R}^m)$, and interacts with the system in time.

Cost functionals. To specify the description of the control problem, we introduce the cost functional. We will consider a quadratic terminal cost

$$\Phi(x) = x^\top Mx, \tag{0.3}$$

where $M \in \mathbb{R}^{d\times d}$ is symmetric positive semi-definite. For the running cost, we will also consider a quadratic one given by

$$L(t, x, u) = x^\top Q(t)x + u^\top R(t)u, \tag{0.4}$$

where $\{Q(t), R(t) : t \in [0, T]\}$ are given symmetric positive definite matrices.

The LQR problem. The linear quadratic regulator problem is formulated as minimizing the above defined cost functionals subject to the dynamics in (0.2). More precisely

$$\begin{aligned} &\min_{\boldsymbol{u}\in L^\infty([0,T],\mathbb{R}^d)} \quad \underbrace{x_T^\top M x_T}_{\Phi(x_T)} + \int_0^T \underbrace{x_t^\top Q(t)x_t + u_t^\top R(t)u_t}_{L(t,x_t,u_t)}\, dt \\ &\qquad\qquad \text{Subject to} \\ &\dot{x}_t = A(t)x_t + B(t)u_t \end{aligned} \tag{0.5}$$

In this project, we will explore some theoretical properties and implement some practical solvers for this problem.

1. We begin with investigating necessary conditions via the Pontryagin's maximum principle (PMP).

    a) Write down the Hamiltonian corresponding to the optimal control problem specified by (0.2).

    b) Write down the state, co-state equations and the Hamiltonian maximization conditions for the PMP corresponding to the LQR problem.

    c) Using the PMP, show that if an optimal control exists (call it $\boldsymbol{u}^*$), with $\boldsymbol{x}^*$ the corresponding optimally controlled trajectory (solution of the state equation), then for each t,

    $$u_t^* = \frac{1}{2}R^{-1}(t)B^\top(t)p_t^*, \tag{0.6}$$

    and furthermore, p_t^* is a linear function of x_t^*. This shows that the optimal control, if exists, is a linear function of the state.

2. Next, let us investigate this linear relationship in detail. From the results in (c), we may set

    $$p_t^* = -2P(t)x_t^*, \tag{0.7}$$

    where $P(t) \in \mathbb{R}^{d\times d}$ is a matrix whose value is to be determined. If we can solve P, we have solved the PMP.

    a) By differentiating (0.7) with respect to t and using the PMP equations for the state and co-state, show that the PMP is satisfied if $P(t)$ satisfies

    $$\dot{P}(t) = -P(t)A(t) - A^\top(t)P(t) - Q(t) + P(t)B(t)R^{-1}(t)B^\top(t)P(t), \quad P(T) = M. \tag{0.8}$$

    This is known as the *Ricatti Differential equation* (RDE).

b) Let $X(t), Y(t)$ be $d \times d$ time-dependent matrices. Suppose they satisfy

$$\frac{d}{dt}\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} A(t) & \frac{1}{2}B(t)R^{-1}(t)B^\top(t) \\ 2Q(t) & -A^\top(t) \end{pmatrix}\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}, \qquad \begin{pmatrix} X(T) \\ Y(T) \end{pmatrix} = \begin{pmatrix} I \\ -2M \end{pmatrix}. \tag{0.9}$$

Show that if we define $P(t) := -\frac{1}{2}Y(t)X^{-1}(t)$, then $P(t)$ satisfies the RDE (0.8). This shows that we can also find a solution to the non-linear $d \times d$ RDE by solving a larger, but linear matrix differential equation of size $2d \times d$.

3. We saw that the PMP allowed us to distill some insights: the optimal control, if exists, should be a linear function of the state. This is also known as a *feedback* or *closed-loop* control. Moreover, if we can find $P(t)$ that satisfies the RDE (0.8), then we can solve the optimal control problem. However, what we cannot conclude at this point are 1) whether an optimal control exists; and 2) whether any solution of the RDE gives rise to an optimal control. To answer these questions, we need to use the Hamilton-Jacobi-Bellman (HJB) framework.

   a) Let us use $V(t,x)$ to denote the value function corresponding to the LQR problem (0.5). Write down the Hamilton-Jacobi-Bellman equation that the value function satisfies.

   b) Show that it simplifies to

$$\begin{aligned} -\partial_t V(t,x) =& x^\top Q(t)x + (\partial_x V(t,x))^\top A(t)x \\ &- \frac{1}{4}(\partial_x V(t,x))^\top B(t)R^{-1}(t)B^\top(t)(\partial_x V(t,x)), \\ V(T,x) =& x^\top Mx. \end{aligned} \tag{0.10}$$

   c) We solve (0.10) using the *ansatz* $V(t,x) = x^\top P(t)x$ where $P(t)$ is to be determined. Show that if $P(t)$ is chosen to satisfy the RDE (0.8), then $V(t,x)$ so defined solves the HJB (0.10). This allows us to conclude that the control $u_t^* = -2P(t)x_t^*$ is indeed optimal. *Remark: A remaining gap here is to prove that the RDE* (0.8) *admits a unique solution. This can be done by combining the usual local existence and uniqueness theorem for ODEs with comparison principles. We omit these issues here.*

4. Let us now investigate the numerical solution of LQR problems. For this purpose, let us consider a simple application. Suppose we are trying to control a vehicle/aircraft travelling in a straight line with time varying air resistance. Let us denote its position by $x_t \in \mathbb{R}$ and velocity by $v_t \in \mathbb{R}$ at time t. Its equation of motion is

$$\begin{aligned} \dot{x}_t &= v_t, \\ \dot{v}_t &= -\alpha(t)v_t + u_t, \end{aligned} \tag{0.11}$$

where $u_t \in \mathbb{R}$ is the applied acceleration of the vehicle, and $\alpha(t) \in \mathbb{R}$ is the coefficient of resistance. The initial position of the vehicle is at $x_0 = 1, v_0 = 0$, and our goal is to control the vehicle (by supplying acceleration u_t) such that at time $t = 1$, x_1 is close to the origin.

Supplying acceleration has a cost, and we measure it by introducing the running cost λu_t^2. Together, this defines the following optimal control problem

$$\min_{\boldsymbol{u}\in L^\infty([0,1],\mathbb{R})} x_1^2 + \lambda \int_0^1 u_t^2 dt \quad \text{subject to (0.11)} \tag{0.12}$$

Let us now explore the numerical solution of this problem.

a) Write programs to solve (0.12), given $\alpha(t)$, in two ways:

   i. using the RDE (0.8)

   ii. Using one other method introduced in class (or another method of your choice).

   Your method should work for a general time-varying resistance $\alpha(t)$. You should test your methods on $\alpha(t) = \sin(10t)$, $\alpha(t) = t^2$, and feel free to explore other choices.

b) Compare the two methods and discuss, using numerical examples, some advantages and disadvantages of these methods. You may consider solution accuracy, running time, memory cost, etc in your discussion.