ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES DEPARTMENT OF MATHEMATICS

QUADRATIC OPTIMAL CONTROL PROBLEMS WITH SHOOTING METHOD

PREPARED BY:

Mr. ASRAT MAMO

Advisor:

Dr. BERHANU GUTA

November 25, 2022

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES DEPARTMENT OF MATHEMATICS

QUADRATIC OPTIMAL CONTROL PROBLEMS WITH SHOOTING METHOD

A Project Work Submitted to the School of Graduate Studies of Addis Ababa University in Partial Fulfilment of the Requirements for the Degree of Master of Optimization.

PREPARED BY:

Mr. ASRAT MAMO

Advisor:

Dr. BERHANU GUTA

November 25, 2022

Dedication

I would like to dedicate this project to the department of mathematics, college of natural and computational science, Addis Ababa University in partial fulfillment of the requirements for the degree of master of science in mathematics. Therefore; I declare that this report done by me during my study has not been submitted to any other higher institutions for the award of any academic certificate.

Advisor Approval sheet

This is the project work entitled by "**Quadratic Optimal Control Problems with Shooting Method**" done by Asrat Mamo under my supervision and guidance.

Name of the Chairperson	Signature	Date
______________________	____________	______________________
Name of Advisor	Signature	Date
Dr. Berhanu Guta	____________	______________________
Name of Internal Examiner One	Signature	Date
______________________	____________	______________________
Name of Internal Examiner Two	Signature	Date
______________________	____________	______________________

ACKNOWLEDGMENT

Foremost, I would like to express my sincere gratitude to my advisor **Dr. Berhanu Guta** for the continuous support and giving motivation and immense knowledge in doing my project. Secondly, I would like to extend my thanks to my families and all who encouraged me to propose my project and their hurtful help while writing the project. Next I would like to thank department of mathematics professors; **Dr. Addisalem Abathun, Dr. Tesfa Biset, Dr. Tilahun Abebaw, Dr. Hunduma Legesse, Dr. Bahiru, Dr. Yibeltal, and others** at Addis Ababa University, for giving good knowledge in my staying years. Lastly I would like to thank GSS-Gibson School Systems staff members for all your considerations. Above all I would like to thank my Almighty God who gives the strength to finish my second step and encouragement me for the third step.

LIST OF ABBREVIATIONS

$C^1[a, b]$	The space consisting of all real valued continuously differentiable functions defined on the bounded and closed interval [a, b]
$C^k[a, b]$	The space consisting of all real valued continuously k-times differentiable functions defined on the bounded and closed interval [a, b]
R^n	n-dimensional Euclidean Space
BVP	Boundary Value Problem
C[a,b]	The space consisting of all continuou real valued functions defined on the bounded and closed interval [a, b]
ELDE	Euler Lagrange Differential Equations
IVP	Initial Value Problem
OCP	Optimal Control Problem
ODE	Ordinary Differential Equations
PWC[a,b]	The space consisting of all piecewise continuous functions defined on the bounded and closed interval [a, b]
PWS[a,b]	The space consisting of all piecewise smooth functions defined on the bounded and closed interval [a, b]
QOCP	Quadratic Optimal Control Problem

TR	Transversality Relation
V	Vector Space

List of Tables

List of Figures

ABSTRACT

Control theory is an area of applied mathematics that deals with principles,laws, and desire of dynamic systems. Optimal control problems are generalized form of variation problems. A very important tool in variational calculus is the notion of Gateaux-differentiability. It is the basis of the development of necessary optimality conditions.
ELDE is a necessary optimality condition to solve variational problems. The solution of ELDE is an extremal function of a variational problem. Characterizing theorem of convex optimization is the necessary and sufficient condition of many convex problems. i.e Let (P) be given, S is convex set, f is convex function and $x_0 \in S$. Then $x_0 \in M(f, S)$ if and only if $f'(x_0, x-x_0) \geq 0, \forall x \in S$. In an OCP our aim is to find the optimal state function $x^*(t)$ and the optimal control function $u^*(t)$ which optimize the objective functional in $t \in [a, b]$ by using necessary and sufficient optimality conditions. The necessary optimality conditions for $(x^*(t), u^*(t))$ to be extremal solutions of optimal control problem is the validity of :- Pontryagin minimum principle, of ELDE with TR, and ODE conditions. To determine whether the extremals are optimal solutions of OCP or not; sufficient optimality conditions are required; (e.g checking the convexity of the objective functional and the convexity of the feasible set). QOCP is a non linear optimization where the cost function is quadratic but the differential equation is linear. In quadratic control problem since the objective function is convex then the extremals are the optimal solution of the problem. Linear QOCP is an important type of quadratic control problem that simples the work of feed back control system. OCP can be solved by different methods depending on the type of the problem. This paper mainly considers solving QOCP by using the method of lagrange multiplier and shooting method.

Contents

Introduction

Optimal control is a part of dynamic optimization which aims to determine control variables so that an optimal value of the objective function is obtained while satisfying some constraints related to the problem system. OCP is formulated in the form of differential equations. It consists of state variables which represent condition of the system over time and control variables which is the decision variables to be determined in order to obtain the optimal value of the objective function. There are three kinds of necessary optimality conditions for (x^*, u^*) to be extremal solutions of an OCP. These are Pontryagin minimum principle, validity of ELDE with TR, and ODE conditions. To determine whether the extremals are optimal solutions of OCP or not sufficient optimality conditions are required like checking the convexity of the objective functional and the convexity of the feasible set. This paper briefly tries to see how to solve linear OCP and QOCP. Linear OCP is an optimal problem where both the objective function and the differential equation are linear. QOCP is an optimal problem where the objective function is quadratic and the differential equation is linear. Linear optimal control is widely applied in manufacturing firms and quadratic optimal control is applied in space science. This paper also focuses on a special case of the general nonlinear optimal control problem called linear QOCP. This paper consists of four chapters.

1. chapter-1: It consists of definiteness of matrices, convex analyses, necessary and sufficient optimality conditions, to solve optimization problems, and shooting method for BVP.

2. Chapter-2: It includes variational problems with fixed end points, variational problem with free right end point. The chapter particularly focuses on how to find optimal solution of a variational problem, using optimality conditions and it also contains applications of variational calculus.

3. Chapter-3: It is the main chapter of this paper which includes definitions and theorems on OCP. This chapter focuses on how to find optimal

control function and optimal state function for a given separated linear and QOCP by using lagrange multiplier function.

4. Chapter-4: It is the last chapter in this project and introduces about the numerical methods that are used to solve a given OCP by using shooting method.

Chapter 1

Preliminaries

1.1 Vector or linear Space

Definition 1.1.1. [**Vector Space**]

A **vector space** is a non-empty set $\boldsymbol{V}$ of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in $\boldsymbol{V}$ and for all scalars $\boldsymbol{c}$ and $\boldsymbol{d}$. [1]

(a) The sum of $\boldsymbol{u}$ and $\boldsymbol{v}$, denoted by $\boldsymbol{u} + \boldsymbol{v}$, is in $\boldsymbol{V}$.

(b) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$.

(c) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$.

(d) There is a **zero** vector $\mathbf{0}$ in $\boldsymbol{V}$ such that $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$.

(e) For each $\boldsymbol{u}$ in $\boldsymbol{V}$, there is a vector **-$\boldsymbol{u}$** in $\boldsymbol{V}$ such that $\boldsymbol{u}$ + (-$\boldsymbol{u}$) = $\mathbf{0}$

(f) The scalar multiple of $\boldsymbol{u}$ by c, denoted by $c\boldsymbol{u}$, is in $\boldsymbol{V}$.

(g) $c(\boldsymbol{u} + \boldsymbol{v}) = c\mathbf{u} + c\mathbf{v}$

(h) $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$

(i) $(cd)\boldsymbol{u} = c(d\boldsymbol{u})$

(j) $\boldsymbol{1u} = \boldsymbol{u}$

Definition 1.1.2. [**Subspace**]

A subset U of the vector space V is a subspace of V if it is closed under the linear operations. That is $x + y \in U$ whenever $x, y \in U$ and $\alpha x \in U$ for each $\alpha \in \mathbb{K}$ and $x \in U$.

Example 1. The set, R^n is a vector space.

Example 2. C[a,b] is a vector space.

Example 3. $C^1[a, b]$ is a vector space.

Example 4. $C^k[a, b]$ is a vector space.

Example 5. PWS[a,b] is a vector space.

Example 6. PWC[a,b] is a vector space.

Example 7. The vector space $C^k[a, b]$ is a subspace of $C^1[a, b]$ or the vector space PWS[a,b] is a subspace of C[a,b].

1.2 Normed space

Definition 1.2.1. [Normed space]
Let V be Linear space over the field $\mathbb{R}$. Then the real valued functional $\|.\| : V \to \mathbb{R}$ is said to be norm on V if and only if for each $f, g \in V$ and $\lambda \in \mathbb{R}$

N1. $\|f\| = 0 \iff f = 0$ (**Nonnegativity**)

N2. $\|\lambda f\| = |\lambda|\|f\|$ (**Positive Homogeneity**)

N3. $\|f + g\| \leq \|f\| + \|g\|$ (**Triangular Inequality**)

The pair $(V, \|.\|)$ is called normed linear space simply normed space. [4] Each normed space V can be consider as metric space by definition of the canonical metric in the following way.

$$d(x, y) = \|x - y\|, \ x, y \in V$$

Example 8. Let $V = C[0, 1]$ be the space consisting of all continuous real valued function defined on the bounded and closed interval $[0, 1]$ with $\|f\|_2 = \left(\int_0^1 |f(t)|^2 \, dt\right)^{1/2}$ is a normed space.

1.3 Definiteness of matrices

A symmetric matrix is a matrix **B** such that $B^T = B$. Such a matrix is necessarily square. Its main diagonal entries are arbitrary, but its other entries occur in pairs—on opposite sides of the main diagonal. [1] and [6]

Definition 1.3.1. [Definiteness of matrices]
An $n \times n$ symmetric matrix, **B** is said to be:

a. positive definite (PD) iff $y^T By > 0$ for any nonzero vector $y \in R^n$.

b. positive semi-definite (PSD) iff $y^T By \geq 0$ for any nonzero vector $y \in R^n$.

c. negative semi-definite iff $-B$ is PSD.

d. negative definite iff $-B$ is PD.

Remark 1.

- If all the leading determinants of symmetric matrix, B are non-negatives, then B is PSD and vice versa.
- If all the leading determinants of symmetric matrix, B are positives, then B is PD and vice versa.
- If all the eigenvalues of symmetric matrix, B are non-negatives, then B is PSD and vice versa.
- If all the eigenvalues of symmetric matrix, B are positives, then B is PD and vice versa.

1.4 The concept of convex sets and convex functions

Definition 1.4.1. [Convex sets]
A set **M** in the Euclidean n-dimensional vector space is said to be ***convex*** if the line segment joining any two points of the set also belongs to the set. In other words, if m_1 and m_2 are in **M**, then $\alpha m_l + (1 - \alpha)m_2$, must also belong to **M** for each $\alpha \in [0, 1]$. [5]

Example 9. M = {x: x solves Problem P below} is a convex set.

$$\text{Problem P: minimize } c^T x$$
$$\text{subject to: Ax = b, } x \geq 0$$

; where c is an n-vector, b is an m-vector, A is an $m \times n$ matrix, and x is an n-vector.

Definition 1.4.2. [Convex functions]
A real-valued function ϕ on (a, b) is said to be **convex** provided for each pair of points x_1, x_2 in (a, b) and each α with $0 < \alpha < 1$, [4] and [5]

$$\phi(\alpha x_1 + (1-\alpha)x_2) \leq \alpha\phi(x_1) + (1-\alpha)\phi(x_2).$$

If we look at the graph of ϕ, the convexity inequality can be formulated geometrically by saying that each point on the chord between $(x_1, \phi(x_1))$ and $(x_2, \phi(x_2))$ is above the graph of ϕ.

Example 10. Let $f : [1, 7] \to \mathbb{R}$. Then $f(x) = x^3 - 2x^2 - x + 2$ is a convex function in the interval $[1, 7]$.

1.5 Notions on derivatives of functionals

1.5.1 Functions of several variables

Definition 1.5.1. [Partial Derivatives]
Let $L : \mathbb{R}^3 \to \mathbb{R}$. Then the partial derivatives of L at the point (x_o, y_o, z_o) are defined as follows. [2]

1. $L_x(x_o, y_o, z_o) = \lim_{h\to 0} \frac{L(x_o+h, y_o, z_o) - L(x_o, y_o, z_o)}{h} = \frac{\partial L}{\partial x}$
2. $L_y(x_o, y_o, z_o) = \lim_{h\to 0} \frac{L(x_o, y_o+h, z_o) - L(x_o, y_o, z_o)}{h} = \frac{\partial L}{\partial y}$
3. $L_z(x_o, y_o, z_o) = \lim_{h\to 0} \frac{L(x_o, y_o, z_o+h) - L(x_o, y_o, z_o)}{h} = \frac{\partial L}{\partial z}$

provided that these limits exist.

Remark 2. **Second-Order Partial Derivatives**
Based on definition 1.5.1, the partial derivatives would generate 9 new second order partial derivatives.

- $(L_x)_x$ usually denoted by L_{xx} or $\frac{\partial^2 L}{\partial x^2}$
- $(L_y)_y$ usually denoted by L_{yy} or $\frac{\partial^2 L}{\partial y^2}$
- $(L_z)_z$ usually denoted by L_{zz} or $\frac{\partial^2 L}{\partial z^2}$

- $(L_x)_y$ usually denoted by L_{xy} or $\frac{\partial^2 L}{\partial y \partial x}$
- $(L_x)_z$ usually denoted by L_{xz} or $\frac{\partial^2 L}{\partial z \partial x}$
- $(L_y)_x$ usually denoted by L_{yx} or $\frac{\partial^2 L}{\partial x \partial y}$
- $(L_y)_z$ usually denoted by L_{yz} or $\frac{\partial^2 L}{\partial z \partial y}$
- $(L_z)_x$ usually denoted by L_{zx} or $\frac{\partial^2 L}{\partial x \partial z}$
- $(L_z)_y$ usually denoted by L_{zy} or $\frac{\partial^2 L}{\partial y \partial z}$

1.5.2 Differentiation under the integral sign

Sometimes a function **L** of a single variable is defined as a definite integral of a function $\boldsymbol{l}$ of two variables like $L(x) = \displaystyle\int_c^d l(x,y)dy$.

Theorem 1.5.1. Let R be a rectangle consisting of all points (x, y) for which $a \le x \le b$ and $c \le y \le d$. Suppose that L is a function of two variables defined on R such that L and L_x are continuous on R. Then $\displaystyle\int_c^d L(x,y)dy$ is a differentiable function of x, and

$$\frac{d}{dx}\int_c^d L(x,y)dy = \int_c^d \frac{\partial L}{\partial x}(x,y)dy.$$

Example 11. Find $\displaystyle\frac{d}{dx}\int_1^2 \frac{e^{xy}}{y}dy$

Solution: By theorem 1.5.1, $\displaystyle\frac{d}{dx}\int_1^2 \frac{e^{xy}}{y}dy = \int_1^2 \frac{\partial}{\partial x}\left(\frac{e^{xy}}{y}\right)dy$

$$= \int_1^2 e^{xy}dy = \frac{e^{xy}}{x}\bigg|_1^2 = \frac{1}{x}(e^{2x} - e^x)$$

□

1.5.3 Total differential

Definition 1.5.2. [Total Differential]

- Let $f : \mathbb{R}^2 \to \mathbb{R}$ be partially differentiable with respect to x and y where x = x(t) and y = y(t). Then the total derivative of L = f(x, y) with respect to t is:

$$\frac{dL}{dt} = \frac{\partial L}{\partial x}\frac{dx}{dt} + \frac{\partial L}{\partial y}\frac{dy}{dt}.$$

- Let $f : \mathbb{R}^3 \to \mathbb{R}$ be partially differentiable with respect to x, y, and z where x = x(t), y = y(t), and z = z(t). Then the total derivative of L = f(x, y, z) with respect to t is:

$$\frac{dL}{dt} = \frac{\partial L}{\partial x}\frac{dx}{dt} + \frac{\partial L}{\partial y}\frac{dy}{dt} + \frac{\partial L}{\partial z}\frac{dz}{dt}.$$

The proof of the total differential in two variables is found in [8].

Remark 1.5.1. $[\nabla L(\cdot), DL(\cdot), \nabla^2 L(\cdot)]$

Let $L : \mathbb{R}^3 \to \mathbb{R}$ be partially differentiable with respect to each of its arguments. For a point $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ and a nonzero vector $h = (h_1, h_2, h_3) \in \mathbb{R}^3$, then:

i. **The gradient of L:** The gradient of L at p_o, denoted by $\nabla L(p_o)$, is a vector given by $\nabla L^T(p_o) = \left(\frac{\partial L}{\partial x}, \frac{\partial L}{\partial y}, \frac{\partial L}{\partial z}\right)$.

ii. **The directional derivative of L:** The directional derivative of L at p_o in the direction of h, denoted by $DL(p_o; h)$, is given by $DL(p_o; h) = \lim_{\epsilon \to 0} \frac{L(p_o + \epsilon h) - L(p_o)}{\epsilon}$. If so, $DL(p_o; h) = \nabla L(p_o) h^T = L_x h_1 + L_y h_2 + L_z h_3$.

iii. **The generalized chain rule of L:** Let $L : \mathbb{R}^3 \to \mathbb{R}$ be partially differentiable with respect to each of its arguments, a point $P_o = (x_o, y_o, z_o) \in \mathbb{R}^3$, and a nonzero vector $h = (h_1, h_2, h_3) \in \mathbb{R}^3$. Then generalized chain rule:

$$\begin{aligned} \frac{d}{d\epsilon} L(p_o + \epsilon h &= L_x(.)h_1 + L_y(.)h_2 + L_z(.)h_3 \\ &= \nabla L(p_o + \epsilon h)^T h \end{aligned}$$

iv. **The Hessian of L:** The Hessian of L at p_o, denoted by $\nabla^2 L(p_o)$ is an $n \times n$ symmetric matrix given by $\nabla^2 L(p_o) = \begin{bmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{xy} & L_{yy} & L_{yz} \\ L_{xz} & L_{yz} & L_{zz} \end{bmatrix}$.

Example 12. Consider a quadratic function $q : \mathbb{R}^n \to R$ given by $q(x) = \frac{1}{2}x^T Ax - B^T x$ where A is $n \times n$ symmetric matrix and B is a column matrix. Show that:

(a) $\nabla q(x) = Ax - B$.

(b) $\nabla^2 q(x) = A$.

Proof.
Let vector x be the usual column vector such as $x = (x_i)_{i=1}^n$, B is a column matrix such as $B = (b_{i1})_{i=1}^n$, and symmetric matrix $A = (a_{ij})_{i=j=1}^n$. Here we could take the derivative of the scalar sum in

$$q(x) = x^T Ax - B^T x = \sum_{j=1}^n x_j \sum_{i=1}^n x_i a_{ji} - \sum_{j=1}^n b_{1j} \sum_{i=1}^n x_i$$

.

(a) To show $\nabla q(x) = Ax - B$: The derivative with respect to the k^{th} variable (product and sum rules of derivative) is then:

$$\begin{aligned}
\nabla q(x) &= \frac{1}{2}\frac{\partial}{\partial x_k}\left(x^T Ax - 2B^T x\right) \\
&= \frac{1}{2}\left(\sum_{j=1}^n \frac{\partial}{\partial x_k} x_j \sum_{i=1}^n x_i a_{ji} + \sum_{j=1}^n x_j \sum_{i=1}^n \frac{\partial}{\partial x_k} x_i a_{ji} - 2\sum_{j=1}^n b_{1j} \sum_{i=1}^n \frac{\partial}{\partial x_k} x_i\right) \\
&= \frac{1}{2}\left(\sum_{i=1}^n x_i a_{ki} + \sum_{i=1}^n x_j a_{jk} - 2(b_{i1})_{i=1}^n\right)
\end{aligned}$$

If we arrange these derivatives into a column vector, we get:

$$\begin{aligned}
\frac{1}{2}\frac{\partial}{\partial x_k}\left(x^T Ax - 2B^T x\right) &= \frac{1}{2}\begin{bmatrix} \sum_{i=1}^n x_i a_{1i} + \sum_{j=1}^n x_j a_{j1} \\ \sum_{i=1}^n x_i a_{2i} + \sum_{j=1}^n x_j a_{j2} \\ \sum_{i=1}^n x_i a_{3i} + \sum_{j=1}^n x_j a_{j3} \\ . \\ . \\ . \\ \sum_{i=1}^n x_i a_{ni} + \sum_{j=1}^n x_j a_{jn} \end{bmatrix} - \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \\ . \\ . \\ . \\ b_{n1} \end{bmatrix} \\
&= \frac{1}{2}\left(Ax + (x^T A)^T\right) - B \\
&= \frac{1}{2}\left((A + A^T)x\right) - B = \frac{1}{2}(2Ax) - B = Ax - B
\end{aligned}$$

(b) To show $\nabla^2 q(x) = A$: We can take the partial derivatives of $\nabla q(x)$.

$$\nabla^2 q(x) = \frac{\partial}{\partial x_k}(Ax - B) = \frac{\partial}{\partial x_k}\sum_{i=1}^{n} a_{ij}x_i$$
$$= (a_{kj})_{k=j=1}^{n} = A.$$

$\square$

1.5.4 Derivatives of functionals-functions of functions

Definition 1.5.3. [Gateaux Derivative]
Let $V = C^1[a,b]$ be a real vector space. The **Gateaux derivative** of a functional $I : V \to \mathbb{R}$ at $x_o \in V$ in the direction of a nonzero $\delta \in V$ is given by $I'(x_o;\delta) = \lim_{\epsilon \to 0} \dfrac{I(x_o + \epsilon\delta) - I(x_o)}{\epsilon}$ if the limit exists. [3]

<u>Remark</u> **1.5.2.** i. $I'(x_o; -\delta) = -I(x_o;\delta)$

ii. Let $\delta > 0$ and define $G : (-\delta, \delta) \to \mathbb{R}$ by $G(\gamma) = I(x_o + \gamma\delta)$. Then $G'(0) = I'(x_o;\delta)$.

iii. $\dfrac{d}{d\gamma}\displaystyle\int_m^n L(\gamma,t)dt = \int_m^n \frac{\partial}{\partial\gamma} L(\gamma,t)dt$

Example 13. Let $I : C^1[0,1] \to \mathbb{R}$ given by $\displaystyle\int_0^1 \left(x^2 + \dot{x}^2\right) dt$. Find the Gateaux derivative of I at $p_o(t) = e^t - 1$ in the direction of $\delta(t) = t$.

<u>Solution:</u>

Here by the above remark, $I'(p_o;\delta) = \displaystyle\int_0^1 \left(2p_o\delta + 2\dot{p_o}\dot{\delta}\right) dt = 2\int_0^1 \left(e^t t + e^t - t\right) dt = \frac{2e-1}{2}$.

1.5.5 Optimality conditions

Derivatives help us to find optimality conditions for optimization problems. [3]

Definition 1.5.4. Let problem, **P**, be unconstrained optimization problem subject to $x \in \mathbb{R}^n$; where $L : \mathbb{R}^n \to \mathbb{R}$.

Problem **P**: minimize $L = c^T x$
subject to: $x \in \mathbb{R}^n$

; where c is an n-vector and x is an n-vector.

i. A point x_o is said to be a local minimizer of L, if there is a neighbourhood N of x_o such that $L(x_o) \leq L(x); \forall(x) \in N$.

ii. A point x_o is said to be a global minimizer of L, if $L(x_o) \leq L(x); \forall(x) \in \mathbb{R}^n$.

Theorem 1.5.2 (Necessary optimality conditions). i. **First order necessary optimality condition**

Let L be differentiable function in N and x_o be a local minimizer of L. Then $\nabla L(x_o) = 0$.

ii. **Second order necessary optimality condition**

Let L be twice differentiable function in N and x_o be a local minimizer of L. Then $\nabla L(x_o) = 0$ and $\nabla^2 L(x_o) \geq 0$.

A point x_o is said to be a global minimizer of L, if $L(x_o) \leq L(x); \forall(x) \in \mathbb{R}^n$.

Theorem 1.5.3 (Sufficient optimality condition). Let L be twice continuously differentiable function over N, and x_o be a local minimizer of L over N. If $\nabla L(x_o) = 0$ and $\nabla^2 L(x_o) > 0$, then x_o is strict local minimizer of L means L is locally convex over N.

Remark 3. If L is convex function, then any local minimizer is a global minimizer.

Definition 1.5.5. Let problem, **P**, be unconstrained optimization problem subject to $x \in \mathbb{S}$; where $L : \mathbb{S} \to \mathbb{R}$.

$$\text{Problem } \mathbf{P}\text{: minimize } L = c^T x$$
$$\text{subject to: } x \in \mathbb{S}$$

; where c is an n-vector and x is an n-vector.

i. A point x_o is said to be a local minimizer of L on S, if there is a neighbourhood N of x_o such that $L(x_o) \leq L(x); \forall(x) \in N \cap S$.

ii. A point x_o is said to be a global minimizer of L, if $L(x_o) \leq L(x); \forall(x) \in N \cap S$.

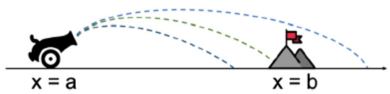

Figure 1.1: Shooting Method

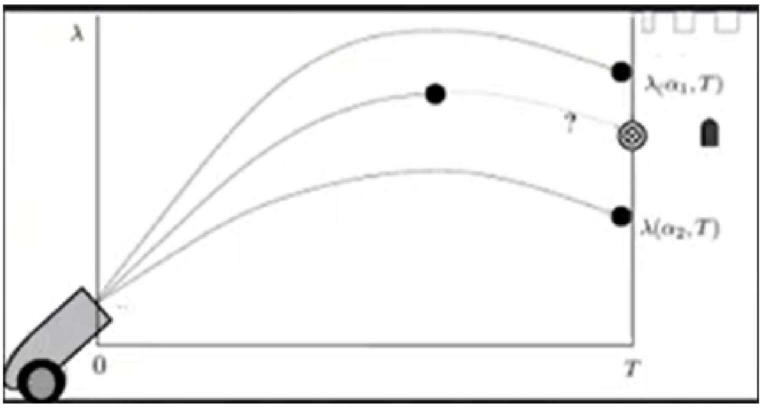

Figure 1.2: Shooting Method

1.6 Shooting method

In numerical analysis, the shooting method is a method for solving a BVP by reducing it to an IVP. It involves finding solutions to the IVP for different initial conditions until one finds the solution that also satisfies the boundary conditions of the BVP. In layman's terms, one "shoots" out trajectories in different directions from one boundary until one finds the trajectory that "hits" the other boundary condition. The shooting method was developed with the goal of transforming an ODE BVP into an equivalent IVP so that we can solve it using the Range-Kuta Method. With initial value problems, we start at the initial value and march forward to get the solution. This method does not work for BVP because there are not enough initial value conditions to solve the ODE to get a unique solution. The shooting method was developed to overcome this difficulty.

The name "shooting method" is analogous with the target shooting: as shown in Fig.1.1 and 1.2, we shoot at the target and observe where we hit the target. Based on the errors, we can adjust our aim and shoot again hoping that we will hit closer to the target. We can see from the analogy that the shooting method is an iterative optimization method. Let us see how the shooting method works using the second-order ODE given $f(a) = f_a$ and

$f(b) = f_b$, as well as

$$F\left(x, f(x), \frac{df(x)}{dx}\right) = \frac{d^2 f(x)}{dx^2}$$

.

1. The first step is we start the whole process by guessing $f'(a) = \alpha$, then together with $f(a) = f_a$, we turn the above problem into an IVP with two conditions all at the value $x = a$. This is the aim step.

2. Second by using Runge–Kutta method, to integrate to the other boundary b to find $f(b) = f_\beta$. This is called the **shooting step**.

3. Thirdly now we compare the value of f_β with f_b. Usually, our initial guess is not good, and $f_b \neq f_\beta$, but what we want is $f_\beta - f_b = 0$; therefore, we adjust our initial guesses and repeat the process until the error is acceptable, at which time we can stop. This is the iterative step.

Due to the applications of BVP in real-life phenomena, the shooting method has proven itself useful and efficient in handling BVP. The shooting method works by first reducing the BVP to an IVP, then one/two initial value guesses are made. The IVPs are then solved using an iterative solution, and this process is then repeated until the second boundary condition is reached to a satisfactory level.

We know that a smooth function $Y : [a, b] \to R^n$ is said to be an approximate solution of the differential equation $\frac{dX}{dz} = f(z, X(z)), X(0) = X_o$ with an error at most τ if

$$||Y(z) - X(z)|| < \tau; \forall(z) \in [a, b]$$

with deviation at most ϵ if $Y(z)$ is continuous and satisfies the differential inequality

$$||\frac{dY}{dz} - f(z, X(z))|| \leq \epsilon$$

.

Range-Kutta Method

Let $\mathbb{P} : a = z_o < z_1 < \cdots < z_{s-1} < z_s = b$ be a partition of the interval $[a, b]$. An approximate solution of the IVP

$$\frac{dX(z)}{dz} = f(z, X(z)); X(a) = X_o$$

on $[a, b]$ is given by the following iterative scheme.

$$X^o = X_o, X^{i+1} = X^i + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4); i = 0, 1, 2, \cdots$$

where $h = \frac{b-a}{s}$ and

$$k_1 = f(z_i, X^i), k_2 = f(z_i + \frac{h}{2}, X^i + h\frac{k_1}{2}), k_3 = f(z_i + \frac{h}{2}, X^i + h\frac{k_2}{2}), k_4 = f(z_i + h, X^i + hk_3)$$

which is an error $E = O(h^5)$ at each iterative step. Then we can use ode45 in MATLAB to solve the IVP at a discrete set of points in the interval $[a, b]$. [22], [23], and [24]

Chapter 2

Variational Calculus

Variation calculus is a field of analytic mathematics that study with optimization problem maximizing or minimizing functional, which are mapping from a set of functions to the real numbers. Functional are often expressed as definite integrals involving functions and their derivatives. The aim is to find an extremal function at which the functional attain a maximum or minimum value, where the rate of change of the functional is zero.

2.1 Variational problems with fixed end points

Definition 2.1.1. Let $a,\ b,\ \alpha,\ \beta \in \mathbb{R}$. Assume, $a < b$ and $\mathrm{L} : \mathbb{R} \times \mathbb{R} \times [a, b] \to R$ be continuous and continuously partially differentiable with respect the first two components. Further more, let J be a functional such that $\mathrm{J} : C^1[a, b] \to \mathbb{R}$, where

$$\begin{aligned} \min\ \mathrm{J(x)} \ &= \int_a^b \mathrm{L(x(t), \dot{x}(t), t)} \\ &s.t\ S = \{x \in C^1[a, b] \mid x(a) = \alpha,\ x(b) = \beta\} \end{aligned}$$

then this type of optimization problem is called variational problem with fixed end points.

Assume the linear subspace U in the vector $C^1[a, b]$

$$U \ = \ \{u \in C^1[a, b] \mid u(a) \ = \ u(b \ = 0)\}$$

Consider the function, $y_0(t) = \frac{\beta-\alpha}{b-a}(t - a) \ + \ \alpha$ then we get $S = y_0 + U$ that is S is linear manifold in $C^1[a, b]$. $\mathrm{J}'(x_0, u) \ = \ 0, \forall u \in U$ is the necessary condition x_0 to be the optimal(minimize) point of J on the region S.

Theorem 2.1.1. Let $a,\ b,\ \alpha,\ \beta \in \mathbb{R}$. Assume, $a < b$ and $\mathrm{L} : \mathbb{R} \times \mathbb{R} \times [a,b] \to R$ be continuous and continuously partially differentiable with respect the first two components. Further more, let J be a functional such that J : $C^1[a,b] \to \mathbb{R}$, where

$$\min\ \mathrm{J(x)} = \int_a^b \mathrm{L(x(t), \dot{x}(t), t)}$$
$$s.t\ S = \{x \in C^1[a,b] \mid x(a) = \alpha,\ x(b) = \beta\}$$

And assume $u \in U$ then the Gateaux derivative of J at $x_0 \in U$ in the direction of u is

$$\mathrm{J}'(x_0,\ u) = \int_a^b (\mathrm{L}_x(x_0, \dot{x}_0, t)u + \mathrm{L}_{\dot{x}}(x_0, \dot{x}_0, t)\dot{u})dt$$

Proof. suppose $x \in S$, such that $\Theta(\varepsilon) = \mathrm{J}(x + \varepsilon u)$

$$\Theta'(\varepsilon) = \mathrm{J}'(x + \varepsilon u)u = \mathrm{J}'(x + \varepsilon u, u)$$
$$\Theta'(0) = \mathrm{J}(x, u)$$
$$\Theta'(0) = \frac{d\Theta}{d\varepsilon}|_{\varepsilon=0} = \int_a^b \frac{d\mathrm{L}}{d\varepsilon}|_{\varepsilon=0} \tag{2.1}$$
$$\mathrm{J}'(x,u) = \int_a^b \frac{d}{d\varepsilon}\mathrm{L}(y + \varepsilon u,\ \dot{x} + \varepsilon\dot{u}, t)dt$$
$$\mathrm{J}'(x,u) = \int_a^b (\frac{\partial \mathrm{L} d}{dx d\varepsilon}(x + \varepsilon u) + \frac{\partial \mathrm{L} d}{d\dot{x} d\varepsilon}(\dot{x} + \varepsilon\dot{u}) + \frac{\partial \mathrm{L} dt}{dt d\varepsilon})dt$$
$$\mathrm{J}'(x,u) = \int_a^b (\mathrm{L}_x(x, \dot{x}, t)u + L_{\dot{x}}(x, \dot{x}, t)\dot{u})dt \tag{2.2}$$

Take $x = x_0$, then the equation 2.2 become

$$\mathrm{J}'(x_0,u) = \int_a^b (\mathrm{L}_x(x_0, \dot{x_0}, t)u + L_{\dot{x}}(x_0, \dot{x_0}, t)\dot{u})dt$$

Therefor, $\mathrm{J}(x_0, u)$ is the Gateaux derivative of J in the direction u. □

Corollary 2.1. Let $a,\ b,\ \alpha,\ \beta \in \mathbb{R}$. Assume, $a < b$ and $\mathrm{L} : \mathbb{R} \times \mathbb{R} \times [a,b] \to R$ be continuous and continuously partially differentiable with respect the first two components.

$$\min_x\ \mathrm{J(x)} = \int_a^b \mathrm{L(x(t), \dot{x}(t), t)}$$
$$s.t\ S = \{x \in C^1[a,b] \mid x(a) = \alpha,\ x(b) = \beta\}$$

And assume any vector $u \in U$ then the Gateaux derivative of J at $x_0 \in U$ in the direction of u is

$$\mathrm{J}(x_0,\ u) \ = \ \int_a^b ((\mathrm{L}_x(x_0, \dot{x}_0, t) \ - \ \mathrm{L}_{\dot{x}}(x_0, \dot{x}_0, t))u)dt \ + \ \mathrm{L}_{\dot{x}}(x_0, \dot{x}_0, t)u|_a^b, \ \forall u \in U.$$

Proof. By using by part integration on the integral $\int_a^b (L_{\dot{x}}(x_0, \dot{x_0}, t)\dot{u})dt$ we have get,

$$\int_a^b (L_{\dot{x}}(x_0, \dot{x_0}, t)\dot{u})dt \ = \ \mathrm{L}_{\dot{x}}(x, \dot{x}, t)u|_a^b \ - \ \int_a^b \frac{d}{dt}\mathrm{L}_{\dot{x}}udt$$

$$\mathrm{J}(x_0, u) \ = \ \int_a^b \mathrm{L}_x(x_0, \dot{x_0}, t)udt \ + \ \mathrm{L}_{\dot{x}}(x, \dot{x}, t)u|_a^b \ - \ \int_a^b \frac{d}{dt}\mathrm{L}_{\dot{x}}udt$$

then,

$$\mathrm{J}(x_0, u) \ = \ \int_a^b (\mathrm{L}_x(x_0, \dot{x_0}, t) \ - \ \frac{d}{dt}\mathrm{L}_{\dot{x}})udt \ + \ \mathrm{L}_{\dot{x}}(x, \dot{x}, t)u|_a^b, \ \forall u \in U$$

□

Fundamental Lemma of Variational Problems

Lemma 2.1.1. If J is continuous function on an open interval (a, b) and h is a smooth function on (a, b) which compactly supports f,such that

$$\int_a^b \mathrm{J}(x)h(x)dx = 0, \ \text{then } \mathrm{J}(x) \ = \ 0$$

2.1.1 Optimality condition for VP-with fixed end points

Theorem 2.1.2. If $y_0 \in S$ is a solution of variational problem with fixed end points, and if J is Gateaux-differentiable at x_0 in the direction of u, $\forall u \in U$, then $\mathrm{J}(x_0, u) \ = \ 0, \forall u \in U$.

Proof. Since U is a sub space of $C^1[a, b]$, then $u \in U$ and $-u \in U$.

$$\mathrm{J}'(x_0, -u) \ = \ \lim_{t \to 0} \frac{\mathrm{J}(x_0 + t(-u)) \ - \ \mathrm{J}(x_0)}{t}$$

$$\mathrm{J}'(x_0, -u) \ = \ \lim_{\lambda \to 0} \frac{\mathrm{J}(x_0 + \lambda u) \ - \ \mathrm{J}(x_0)}{-\lambda}, \ \text{for } \lambda \ = \ -t$$

$$\mathrm{J}'(x_0, -u) \ = \ -\lim_{\lambda \to 0} \frac{\mathrm{J}(x_0 + \lambda u) \ - \ \mathrm{J}(x_0)}{\lambda}$$

$$\mathrm{J}'(x_0, -u) \;=\; \mathrm{J}'(x_0, -u)$$

If $x_0 \in S$ is a minimizer of J over the space S, then $\mathrm{J}(x_0) \;\leq\; \mathrm{J}(x_0{+}tu)\;\forall u \in U$ implies

$$\mathrm{J}(x_0 + tu) \;-\; \mathrm{J}(x_0) \;\geq\; 0 \;\forall u \in U$$

$$\frac{\mathrm{J}(x_0 + tu) \;-\; \mathrm{J}(x_0)}{t} \;\geq\; 0$$

$$\lim_{t \to 0} \frac{\mathrm{J}(x_0 + tu) \;-\; \mathrm{J}(x_0)}{t} \;\geq\; 0$$

$$\mathrm{J}(x_0, u) \geq 0 \quad ------ (*)$$

Since U is a subspace of $C^1[a, b]$, then $-u \in U$

$$\mathrm{J}(x_0, u) \geq 0 \text{ and}$$

$$\mathrm{J}(x_0, -u) \geq 0 \Leftrightarrow -\mathrm{J}(x_0, u) \geq 0$$

$$\Rightarrow \mathrm{J}(x_0, u) \leq 0 \quad ------ (**)$$

Form $(*)$ and $(**)$ we have get

$$\mathrm{J}(x_0, u) \;=\; 0, \;\forall u \in U$$

□

2.2 The Euler Lagrange Differential Equation

If $x_0 \in S$ is a solution of a VP fixed end points,then

$$\mathrm{J}'(x_0, u) = 0, \;\forall u \in U$$

This implies,

$$\int_a^b (\mathrm{L}_x(x, \dot{x}, t)u \;+\; L_{\dot{x}}(x, \dot{x}, t)\dot{u})dt \;=\; 0$$

$$\int_a^b (\mathrm{L}_x(x_0, \dot{x_0}, t) \;-\; \frac{d}{dt}\mathrm{L}_{\dot{x}}(x, \dot{x}, t))udt \;+\; \mathrm{L}_{\dot{x}}(x, \dot{x}, t)u|_a^b \;=\; 0$$

Since $u(a) = 0$ and $u(b) = 0$ we have

$$\int_a^b (\mathrm{L}_x(x, \dot{x}, t) \;-\; \frac{d}{dt}\mathrm{L}_{\dot{x}}(x, \dot{x}, t))udt \;=\; 0$$

$$(\mathrm{L}_x(x, \dot{x}, t) \;-\; \frac{d}{dt}\mathrm{L}_{\dot{x}}(x, \dot{x}, t)) \;=\; 0.$$

Let $u(t) = h(t)(t-a)(b-t) > 0$ where $h \in C^1[a,b]$

$$(\mathrm{L}_x(x,\dot{x},t) \ - \ \frac{d}{dt}\mathrm{L}_{\dot{x}}(x,\dot{x},t)) \ = \ 0$$

is the Euler Lagrange differential equation(ELDE)

Theorem 2.2.1. (Necessary condition on optimal)
Let $a,\ b,\ \alpha,\ \beta \in \mathbb{R}$. Assume, $a < b$ and $\mathrm{L} : \mathbb{R}\times\mathbb{R}\times[a,b] \to R$ be continuous and continuously partially differentiable with respect the first two components.
Assume the VP:

$$\begin{aligned}\min_x \ \mathrm{J}(x) \ &= \ \int_a^b (\mathrm{L}(x(t),\dot{x}(t),t))dt \\ S \ &= \ \{ \ x \in C^1[a,b] \mid x(a) = \alpha, \ x(b) = \beta \ \}\end{aligned}$$

suppose x_0 be an optimal solution the VP.
The necessary condition for x_0 to be extremal solution of the variational problem is that x_0 should satisfy the $ELDE$ given by

$$\frac{d}{dt}\mathrm{L}_{\dot{x}}(x_0(t),\dot{x}_0(t),t) \ = \ \mathrm{L}_x(x_0(t),\dot{x}_0(t),t), \forall t \in [a,b] \quad (ELDE)$$

Definition 2.2.1. (local minimizer points) Suppose $S = x_0 + u$ and $x_0 \in S$. A point x_0 is called algebraically local minimizer point of J on region S if and only if $\forall u \in U$ there exists an $\varepsilon > 0$ such that x_0 is a minimizer point of J on $[x_0 - \varepsilon u, x_0 + \varepsilon u]$. Then we have get the following theorem

Theorem 2.2.2. Let $a,\ b,\ \alpha,\ \beta \in \mathbb{R}$. Assume, $a < b$ and $\mathrm{L} : \mathbb{R}\times\mathbb{R}\times[a,b] \to R$ be continuous and continuously partially differentiable with respect the first two components.

$$\begin{aligned}\min_x \ \mathrm{J}(x) \ &= \ \int_a^b (\mathrm{L}(x(t),\dot{x}(t),t))dt \\ s.t. \ S \ &= \ \{ \ x \in C^1[a,b] \mid x(a) = \alpha, \ x(b) = \beta \ \}\end{aligned}$$

If x_0 is algebraically a local minimizer point J on the region S then $_0$ satisfies

$$\frac{d}{dt}\mathrm{L}_{\dot{x}}(x_0(t),\dot{x}_0(t),t) \ = \ \mathrm{L}_x(x_0(t),\dot{x}_0(t),t), \forall t \in [a,b] \quad (ELDE)$$

2.3 Convexity of VP

For x_0 to be a minimizer point in the case of convexity we can able to apply a sufficient condition instead of necessary condition. Suppose
$\mathrm{L} : \mathbb{R} \times \mathbb{R} \times [a,b] \to \mathbb{R}$, be convex $\forall t \in [a,b]$. We need to say L be convex function with regard to the first two components. We need to show J be convex VP function. New take $x, y \in S$ and the value with respect $x,\ y$ is given as follow

$$\begin{aligned} \mathrm{J}(x) &= \int_a^b (\mathrm{L}(x(t), \dot{x}(t), t))dt \\ \mathrm{J}(y) &= \int_a^b (\mathrm{L}(y(t), \dot{y}(t), t))dt \end{aligned}$$

since L is convex on the first two component, we have

$$\mathrm{L}(\lambda x \ + \ (1-\lambda)y, \lambda\dot{x} \ + \ (1-\lambda)\dot{y}, t) \ \leq \ \lambda\mathrm{L}(x, \dot{x}, t) \ + \ (1-\lambda)\mathrm{L}(y, \dot{y}, t)$$

take integral to both sides and we get

$$\int_a^b \mathrm{L}(\lambda x \ + \ (1-\lambda)y, \lambda\dot{x} \ + \ (1-\lambda)\dot{y}, t) \ \leq \ \int_a^b \lambda\mathrm{L}(x, \dot{x}, t) \ + \ (1-\lambda)\mathrm{L}(y, \dot{y}, t)$$

this implies J is convex.

Proposition 2.1. (Necessary and sufficient optimal conditions)
$\mathrm{L} : \mathbb{R} \times \mathbb{R} \times [a,b] \to R$ be continuous and continuously partially differentiable with respect the first two components and x_0 be an optimal solution of the (VP)

$$\begin{aligned} \min_x \ \mathrm{J}(x) &= \int_a^b (\mathrm{L}(x(t), \dot{x}(t), t))dt \\ s.t. \ S &= \{ \ x \in C^1[a,b] \mid x(a) = \alpha, \ x(b) = \beta \ \} \end{aligned}$$

i If the $ELDE$ condition holds

ii L is convex with regard to the first two components $forall t \in [a,b]$(i.e if the Heissiean matrix of L with regard to the first two components is positive semi-definite $\forall t \in [a,b]$).

then x_0 is the optimal solution of the VP.

Theorem 2.3.1. Let L be convex with regard to the first two components $\forall t \in [a,b]$. Then $x_0 \in S$ is a minimizer point of J on region S if and only if x_0 is a solution of $ELDE$.

Proof. i Let x_0 be a minimizer of J on space S. Then by the Thorem 2.2.1 $ELDE$ holds

ii Let now the $ELDE$ be true for x_0 by (i). Let; $\mathrm{L}_x = \mathrm{L}_x(x_0, \dot{x}_0, t)$ and $\mathrm{L}_{\dot{x}} = \mathrm{L}_{\dot{x}}(x_0, \dot{x}_0, t)$, $t \in [a, b]$, then from the validty of $ELDE$ we have

$$\mathrm{L}_x(x, \dot{x}, t) - \frac{d}{dt}\mathrm{L}_x(x, \dot{x}, t) \;=\; 0$$

$$For\ u \in U = \;\{u \in C^1[a, b] \mid u(a) = 0, u(b) = 0\}$$

$$\int_a^b (\mathrm{L}_x(x, \dot{x}, t) - \frac{d}{dt}\mathrm{L}_x(x, \dot{x}, t))udt \;=\; 0$$

$$\int_a^b (\mathrm{L}_x(x, \dot{x}, t) - \frac{d}{dt}\mathrm{L}_x(x, \dot{x}, t))udt \;+\; \mathrm{L}_{\dot{x}}u|_a^b = \;0$$

$$\int_a^b (\mathrm{L}_x u \;+\; \mathrm{L}_{\dot{x}}\dot{u})dt = \;0$$

$$\mathrm{J}'(x_0, u) \;=\; 0$$

This implies $\mathrm{J}'(x_0, u) \;=\; 0 \;\geq 0,\ \forall u \in U$. If we set $x = x_0 \,+\, uthenu = x \;-\; y_0$ there fore we get $\mathrm{J}'(x_0, x - x_0)\forall x \in S$ By the characterizing theorem of convex optimization ,we have that x_0 is the minimizer point of J on space S.

□

2.4 Variational problems with free end points

Let L be continuous and continuously partially differentiable with regard to the first two components. Then we assuming the following VP With one free end point:

$$\min_x \;\mathrm{J(x)} \;=\; \int_a^b \mathrm{L(x(t), \dot{x}(t), t)}$$

$$s.t\ D = \{x \in C^1[a, b] \mid x(a) = \alpha\}$$

With value of left boundary point is known, but the value of the right boundary point is not known(free). Simply D is a linear manifold in the vector space $C^1[a, b]$ In order to place a necessary conditions for the solution of the problem, we assume that we have already found a solution $x_0(t)$ of (PV)

$$Let S = \{x \in D \;:\; x(b) = x_0(b)\} \subset D$$

that is, $D = \{x \in C^1[a,b] \mid x(a) = \alpha, x_0(b) = x(b)\}$ If x_0 is a minimum point of J on the region S (as in the VP with fixed end points) $ELDE$ holds, that is $\frac{d}{dt}L_{\dot{x}}(x_0, \dot{x}_0, t) \;=\; L_x(x_0, \dot{x}_0, t)$ is a necessary condition for x_0 to be a minimum point of J on the region S. More over the continuity of L_x implies the continuity of $\frac{d}{dt}L_{\dot{x}}$. We consider the subspace

$$V \;=\; \{v \in C(1)[a,b] \mid v(a) = 0\}$$

As in section (2.1) for x_0 to be a minimizer point of J on the region D then following equation must hold

$$0 = J(x_0, w) \;=\; \int_a^b (\mathrm{L}_x(x,\dot{x},t)v - \frac{d}{dt}\mathrm{L}_x(x,\dot{x},t)\dot{v})dt \;=\; \int_a^b (\mathrm{L}_x(x,\dot{x},t) - \frac{d}{dt}\mathrm{L}_{\dot{x}}(x,\dot{x},t))vdt + L_{\dot{x}}v|_a^b$$

because of

$$L_x - \frac{d}{dt}L_{\dot{y}} \;=\; 0$$

, this implies

$$L_{\dot{x}}v|_a^b \;=\; 0$$
$$\Rightarrow L_{\dot{x}}(x_0(b), x_0(b), b)v(b)) - L_{\dot{x}}(x_0(a), \dot{x}_0(a), a)v(a)) = 0$$
$$L_{\dot{x}}(x_0(b), \dot{x}_0(b), b)v(b) = 0,\; sincev(a) = 0$$
$$\forall v \in V\; then,\; L_{\dot{x}}(x_0(b), \dot{x}_0(b), b) = 0,\; Sincev(b) > 0$$

We call this condition transversality condition(TR).

Theorem 2.4.1. (Necessary and sufficient optimality conditions) Let L be continuous and continuously partially differentiable with regard to the first two components and x_0 be a solution of the VP.

$$\min_x \;\mathrm{J(x)} \;=\; \int_a^b \mathrm{L}(x(t), \dot{x}(t), t)dt$$
$$s.t\; D = \{x \in C^1[a,b] \mid x(a) = \alpha\}$$

i If the $ELDE$ and TR conditions hold and

ii L is convex with regard to the first two components $forallt \in [a,b]$(i.e If the Heissiean Matrix of L with regard to the first two components is positive semi definite $forallt \in [a,b]$). Then x_0 is optimal solution of the problem.

Then x_0 is optimal solution of the problem.

2.5 Bolza problem

The following class of variational problem is called special Bolza Problem.
$J(x) = g(x(a), a) + h(x(b), b) + \int_a^b L(x, \dot{x}, t)$

$$\begin{array}{cc} \min & J(x) \\ s.t & x \in S = \{x \in C^1[a,b] | x(a) = x_a,\ x(b) = x_b\} \end{array}$$

Theorem 2.1. *Let L be continuous and continuously partially differentiable with regard to the first two components. if x_0 is minimizer of $J(x)$ over S then the following holds:*

$$\frac{d}{dt} L_{\dot{x}}(x_0, \dot{x}_0, t) - L_x(x_0, \dot{x}_0, t) = 0 \quad (ELDE)$$
$$L_{\dot{x}}(x_0(a), \dot{x}_0(a), a) = g'(x_0(a)),\ L_{\dot{x}}(x_0(b), \dot{x}_0(b), a) = -h'(x_0(b)) \quad (TR)$$

Theorem 2.2. *(Necessary and sufficient condition) Let g and h be convex, K be convex set,L be convex with regard to the first two components for each $t \in [a, b]$.If for each $x_0 \in K$ the conditions*

$$\frac{d}{dt} L_{\dot{x}}(x_0, \dot{x}_0, t) - L_x(x_0, \dot{x}_0, t) = 0 \quad (ELDE)$$
$$L_{\dot{x}}(x_0(a), \dot{x}_0(a), a) = g'(x_0(a)),\ L_{\dot{x}}(x_0(b), \dot{x}_0(b), a) = -h'(x_0(b)) \quad (TR)$$

Then x_0 is the minimum point of $J(x)$ on K.

Chapter 3

Optimal Control Theory

3.1 Introduction

Optimal Control theory is an area of applied mathematics that study about the principles,laws, and desire of dynamic systems.The subject of Optimal Control is concerned with dynamic systems and their optimization over time. System is entity which respond to relevant inputs according to certain laws such as physical laws. Dynamic system is a system that evolves over time such as moving object, production process, population growth, etc.

Example 3.1.1.

- Consider the moving object :- the trajectories are $x(t)$ = position at a time t and $x_(t)$ = velocity at a time t and the input(control) is $u(t)$ force at a time t is Dynamic system .

- Consider the Manufacturing Firm :-the trajectories are $x(t)$= inventory level at a time t and $x_(t)$ = rate of change of inventory level at a time t,and the control is $u(t)$ = production level at a time t is a dynamic system.

Controlled Dynamics is when a suitable external force (input), called control, is applied to influence (or regulate) the dynamics to achieve a desired goal. Let the system is given by

$$\begin{aligned} \dot{x}(t) = x(t) \qquad & t > 0 \\ x(0) = 1 \qquad & \Rightarrow x(t) = e^t, t \geq 0 \end{aligned}$$

Can we influence the dynamics/ motion of this system so that $x(1) = 0$.For instance, if a control u is included to the above dynamics as:

$$\begin{aligned} \dot{x}(t) = x(t) + u \qquad & 0 < t < 1 \\ x(0) = 1 \qquad & \Rightarrow x(t) = e^t, t \geq 0 \end{aligned}$$

we find a constant u so that $x(1) = 0$ as $u(t) = \frac{e}{1-e}$ and $x(t) = \frac{e^t - e}{1-e}$

Letting $x(t)$ to be the quantity of interest at time t, a continuous time dynamics is modeled by a differential equation of the form:

$$\begin{aligned} \dot{x}(t) &= \psi(x(t), u, t) \quad t > 0 \quad (ODE) \\ x(0) &= x_0 \qquad\qquad\qquad (IC) \end{aligned}$$

$x(t)$ is called state of the system.

In Optimal control theory we need to obtain a control $u(t)$ which forces the state $x(t)$ from initial state to the desired terminal state where and $u(t)$ = control variable which are related by a differential equation. $\dot{x} = \psi(x, u, t), t \in [a, b]$ where $u(t) \in \mathcal{A}$, for some non empty set $\mathcal{A}$ called the set of admissible control. A optimal control problem is used to obtain a control $u(t)$ and the corresponding trajectory $x(t)$ over $[a, b]$ such that $(x(t), u(t))$ is a solution of a differential equation

Definition 3.1.1. A Optimal Control problem is used to characterize a control function $u(t)$ and the corresponding trajectory function $x(t)$ over the interval time horizon $[0, T]$ such that $(x(t), u(t))$ is a solution of a differential equation

$$\begin{array}{ccc} \dot{x}(t) = \psi(x, u, t) & 0 < t < T & (ODE) \\ x(0) = x_0, & x(T) = x_T & (IC) \\ u(t) \in \mathcal{A} & 0 < t < T & \end{array}$$

where $x(.)$ is a piecewise smooth function and $u(.)$ is a piecewise continuous function

One often interested not merely in finding arbitrary control $u(t)$ and the corresponding trajectory $x(t)$ that solves the control problem but wishes to do so in a "best" possible manner say minimize the cost associated with the operation

In general, the subject of Control Theory is by itself a wide area in applied mathematics which deals with the principles(laws), analysis, and design of control of a dynamic system. There are various issues that need investigation in the subject of control theory. Controllability is among these issues. The controllability question is: Given a dynamic system with a state equation $\dot{x} = \psi(x, u, t)$. can we bring the state of the system from its current state x_0 to an arbitrary prescribed target state x_T by applying some admissible control $\mathcal{A}$ over a finite time? i.e., does there exist an admissible control $u(t)$ that steers the system from initial state x_0 to a given terminal state x_T by some time T? If an admissible u exists, the system is said to be controllable and the problem is to find such $u(t)$ or characterize it. In Optimal Control we assume that the system is controllable (with various alternative admissible controls) and our concern is to find the best control among those admissible controls.

3.2 Optimal Control Problems [OCP]

Solution of a control problem depends on the choice of the control u. consider the CP

$$\begin{aligned} \dot{x} &= x + u \qquad 0 < t < 1 \\ x(0) &= 1, x(1) = 0 \\ u(.) &\in C[0,1] \end{aligned}$$

If u is constant,the solution of this CP is $u(t) = \frac{e}{1-e}$ and $x(t) = \frac{e^t - e}{1-e}$ If we choose an exponential u, then the solution is $u(t) = e^{-t}$ and $x(t) = (1-t)e^t, 0 < t < 1$.Other types of the control u can lead to other possible solutions. So, a natural question is what is the best control? In deed, one often interested not simply in finding certain control $u(t)$ that solves a control problem, but wishes to do so in a best possible manner, say, minimize the total cost of operation.

Suppose the cost of exerting a control $u(t)$ while the system is in state $x(t)$ at a time t be $f(x,u,t)$, then the total cost over [0,T] is given by $\int_0^T f(x,u,t)dt$. Thus the OCP that solves an optimal control $u(t)$ and the corresponding trajectory $x(t)$ is given as:

$$\begin{aligned} min \quad & \int_0^T f(x,u,t)dt \\ s.t \quad & \dot{x} = \psi(x,u,t) \qquad t \in [0,T] \\ & x(0) = x_0,\ x(T) = x_T \\ & u(t) \in \mathcal{A} \qquad t \in [0,T] \end{aligned}$$

This is called the OCP with fixed endpoints. If the end condition $x(T) = x_T$ is not prescribed, then it is called the OCP with free endpoint or free at right side..

Remark 4. A variational problem can be converted into OCP consider the following variational problem and setting $u(t) = \dot{x}(t)$:

$$\begin{aligned} min \quad & \int_0^T L(x,\dot{x},t)dt \\ s.t \quad & x(0) = x_0,\ x(T) = x_T \end{aligned}$$

Thus, the corresponding OCP is :

$$\begin{aligned} min \quad & \int_0^T L(x,u,t)dt \\ s.t \quad & \dot{x} = u, t \in [0,T] \\ & x(0) = x_0,\ x(T) = x_T \end{aligned}$$

3.3 Optimality Conditions for OCP

In this section we study a method to characterize and solve OCP of the following type:

$$\begin{array}{rll} min & \varphi(x(0), x(T)) + \int_0^T f(x,u,t)dt & \\ s.t & \dot{x} = \psi(x,u,t) & t \in [0,T] \\ & x(.) \subseteq PWS[o,T] & \\ & u(.) \subseteq PWC[0,T] & \end{array}$$

3.3.1 Lagrange Approach for OCP

To characterize the optimal control u^* and corresponding optimal trajectory x^* that solves the OCP with free endpoint(fE) of the form:

$$(\mathbf{OCP}_{fE}) \quad \begin{array}{rll} \min & \int_0^T f(x,u,t)dt & \\ s.t & \dot{x} = \psi(x,u,t) & (ODE) \\ & x(0) = x_0 & (IC) \\ & u(t) \in \mathcal{A} & 0 < t < T \end{array}$$

and with Fixed endpoint(FE) of the form:

$$(\mathbf{OCP}_{FE}) \quad \begin{array}{rll} \min & \int_0^T f(x,u,t)dt & \\ s.t & \dot{x} = \psi(x,u,t) & (ODE) \\ & x(0) = x_0, x(T) = x_T & (BC) \\ & u(t) \in \mathcal{A} & 0 < t < T \end{array}$$

In Lagrange method or approach, corresponding to the (ODE) constraint, we take a piecewise smooth function $\lambda : [0,T] \to R^n$, called costate function, and define the following Lagrange function: $L(x, \dot{x}, u, \lambda, t) = f(x,u,t) - \lambda\psi(x,u,t) + \lambda\dot{x}$, Then, the corresponding Lagrange functional is

$$\mathfrak{F}_\lambda(x,u) = \int_0^T L(x, \dot{x}, u, \lambda, t)dt$$

That is, the Auxiliary Lagrange Problem for the given $(\mathbf{OCP}_{fE})$ is

$$(\mathbf{ALP}) \quad \begin{array}{rll} \min & \mathfrak{F}_\lambda(x,u) & \\ s.t & x(0) = x_0 & (IC) \\ & u(t) \in \mathcal{A}, & 0 < t < T \end{array}$$

The Auxiliary Lagrange Problem for the given $(\mathbf{OCP}_{FE})$ is

$$(\mathbf{ALP}) \quad \begin{array}{rll} \min & \mathfrak{F}_\lambda(x,u) & \\ s.t & x(0) = x_0, x(T) = x_T & (BC) \\ & u(t) \in \mathcal{A}, & 0 < t < T \end{array}$$

The relation between the ($\mathbf{OCP}_{fE}$) and ($\mathbf{OCP}_{FE}$) to (**ALP**) is given in the following Lemma.

Lemma 3.3.1. *(Lagrange Lemma) If (u^*, x^*) are optimal solutions of the* ***ALP****, for some λ, and satisfy the ODE, then u^* is an optimal control and x^* is the corresponding optimal trajectory (optimal state) for the* ***OCP****.*

So, to characterize the optimal control u^* and corresponding x^* for the OCP, it suffices to characterize optimal solutions of the ALP that satisfy the ODE.In order to characterize the solutions u^* and x^* of the OCP, the Lagrange approach looks for the following three conditions:

1. First take x and λ (say $x = x^*$, $\lambda = \lambda^*$) as fixed parameters, then find the conditions on u^* that minimizes the Lagrange

$$\mathfrak{F}_\lambda(x^*, u) = \int_0^T L(x^*, \dot{x}, u, \lambda^*, t)dt$$

over $\mathcal{A}$. This will be handled by Pontryagin min (max) Principle.

2. take u and λ as fixed parameters, (say, $u = u^*$ and $\lambda = \lambda^*$), find the conditions on $x^*(.)$ that solve the problem:

$$\begin{array}{lll} \min & L(x, \dot{x}, u^*, \lambda^*, t)dt & \\ s.t & x(0) = x_0 & (IC) \\ & u(t) \in \mathcal{A}, & 0 < t < T \end{array}$$

This is VP with free Endpoint. So, (ELDE) and (TR) will be applicable to it.

$$\begin{array}{lll} \min & L(x, \dot{x}, u^*, \lambda^*, t)dt & \\ s.t & x(0) = x_0,\ x(T) = x_T & (BC) \\ & u(t) \in \mathcal{A}, & 0 < t < T \end{array}$$

This is VP with fixed Endpoint. So, (ELDE) and (BC) will be applicable to it.

3. u^* and corresponding x^* that result from (1) and (2) should satisfy the ODE: $\dot{x} = \psi(x, u, t)$ and IC: $x(0) = x_0$.

Theorem 3.1. *Let x and λ be fixed. $u^*(.) \in \mathcal{A}$ is a minimizer of $\mathfrak{F}_\lambda(x, u) = \int_0^T L(x, \dot{x}, u, \lambda, t)dt$ if and only if $u^*(.)$ is minimizer of $L(x, \dot{x}, u, \lambda, t)$ over $\mathcal{A}$.*

Proof. ($\Rightarrow$)By contrary, suppose there is some $\hat{u} \in \mathcal{A}$ such that $L(x, \dot{x}, \hat{u}, \lambda, t) < L(x, \dot{x}, u^*, \lambda, t)$. $\int_0^T L(x, \dot{x}, \hat{u}, \lambda, t) < \int_0^T L(x, \dot{x}, u^*, \lambda, t)$. $\mathfrak{F}_\lambda(x, \hat{u}) < \mathfrak{F}_\lambda(x, \dot{u})$ which is contradiction to u^* is optimal point for $\mathfrak{F}_\lambda(x, u)$.

($\Leftarrow$) suppose u^* be optimal point for $\mathfrak{F}_\lambda(x, u)$ over $\mathcal{A}$. then $\mathfrak{F}_\lambda(x, u^*) < \mathfrak{F}_\lambda(x, u)$ for all $u \in \mathcal{A}$. hence,$L(x, \dot{x}, u^*, \lambda, t) < L(x, \dot{x}, u, \lambda, t)$ for all $u \in \mathcal{A}$. $\square$

Recall that $L(x, \dot{x}, u, \lambda, t) = f(x, u, t) - \psi(x, u, t) + \lambda \dot{x}$. This implies that u^* $mathcalA$ minimizes $L(x, \dot{x}, u, \lambda, t)$ over $\mathcal{A}$ if and only if u^* minimizes $f(x, u, t) - \psi(x, u, t)$ over $\mathcal{A}$. To simplify the description of optimality conditions, we define the following function corresponding to the OCP:

$$H(x, u, \lambda, t) = f(x, u, t) - \psi(x, u, t)$$

called Hamiltonian function for the OCP.

Definition 3.3.1. Given an Optimal Control problem OCP

$$(\mathbf{OCP}) \quad \begin{array}{ccc} \min & \int_0^T f(x, u, t)dt & \\ s.t & \dot{x} = \psi(x, u, t) & (ODE) \\ & x(.) \in \mathcal{S},\ u(.) \in \mathcal{A} & (IC) \end{array}$$

we define the Hamiltonian of the OCP by

$$H(x, u, \lambda, t) = f(x, u, t) - \psi(x, u, t)$$

where $x(.) \in \mathcal{S}$, $u(.) \in \mathcal{A}$, $\lambda(.)$ the costate function, $t \in [0, T]$.

Note that from Theorem (3.1), the definition of H and the above definition, u^* is a minimizer of $\mathfrak{F}_\lambda(x, \dot{u})$ over $\mathcal{A}$ iff u^* is a minimizer of $H(x, u, \lambda, t)$ over $\mathcal{A}$. Therefore, we have the following Theorem.

Theorem 3.2. *(Pontryagin Minimum Principle) By considering x and λ as fixed parameters, if $u^* \in \mathcal{A}$ is a minimizer of $\mathfrak{F}_\lambda(x, u)$ over $\mathcal{A}$ iff u^* is a minimizer of $H(x, u, \lambda, t)$ over $\mathcal{A}$. That is,* $H(x, u^*, \lambda, t) = \begin{array}{cc} \min & H(x, u, \lambda, t) \\ s.t & u \in \mathcal{A} \end{array}$

Proof. □

Assuming we have optimal control u^*, we look for the conditions on corresponding optimal, trajectory x^* Situation (2) and (3)

Theorem 3.3. *(Hamiltonian Dynamics) Let u^* be an optimal control and $x^*(.)$ is the corresponding optimal trajectory for the above OCP, then there is a co-state $\lambda^*(.)$ such that $x^*(.)$ and $\lambda^*(.)$ are solutions of the following ODE*

$$\begin{array}{lcr} \dot{\lambda} = H_x(x, u^*, \lambda, t) & \ldots & (ADJ) \\ \lambda(T) = 0 & \ldots & (TR) \\ \dot{x} = \psi(x, u^*, t)) & \ldots & (ODE) \\ x(0) = x_0 & \ldots & (IC) \end{array}$$

Proof. As stated in description of Situation (2) above, x^* is a solution of the VP :

$$\begin{array}{cc} \min & \int_0^T L(x, \dot{x}, u^*, \lambda^*, t)dt \\ s.t & x(0) = x_0 \end{array}$$

$L(x, \dot{x}, u, \lambda, t) = H(x, u, \lambda, t) + \lambda\dot{x} \Rightarrow L_x = H_x$ and $L_{\dot{x}} = \dot{\lambda}$. By (ELDE) gives that $\dot{\lambda} = H_x(x, u^*, \lambda^*, t)$ and from (TR) of VP with free end, gives that $L_{\dot{x}}(T) = \dot{\lambda}(T) = 0$. □

Thus, collecting the above results together the necessary optimality conditions for the OCP with Free Endpoints:

$$(\mathbf{OCP}) \qquad \begin{array}{rll} \min & \int_0^T f(x, u, t)dt & \\ s.t & \dot{x} = \psi(x, u, t) & (ODE) \\ & x(0) = x_0 & (IC) \\ & u(t) \in \mathcal{A} & 0 < t < T \end{array}$$

where the Hamiltonian is $H(x, u, \lambda, t) = f(x, u, t) - \lambda\psi(x, u, t)$.

Theorem 3.4. *(Necc. Opt. cond. for the ($\boldsymbol{OCP}_{fE}$)) If $u^*(.)$ is an optimal control and $x^*(.)$ is corresponding optimal trajectory of the above ($\boldsymbol{OCP}_{fE}$), then there is a costate $\lambda^*(.)$ such that the followings hold:*

A. *(PMP) $u^*(.)$ is the minimizer of $H(x^*, u, \lambda^*, t)$ over $\mathcal{A}$.*

B. *$x^*(.)$ and $\lambda^*(.)$ are solutions of :* $$\begin{array}{lll} \dot{\lambda} = H_x(x, u^*, \lambda, t) & ... \rightarrow & (ADJ) \\ \lambda(T) = 0 & ... \rightarrow & (TR) \\ \dot{x} = \psi(x, u^*, t)) & ... \rightarrow & (ODE) \\ x(0) = x_0 & ... \rightarrow & (IC) \end{array}$$

Theorem 3.5. *(Necc. Opt. cond. for the ($\boldsymbol{OCP}_{FE}$)) If $u^*(.)$ is an optimal control and $x^*(.)$ is corresponding optimal trajectory of the above ($\boldsymbol{OCP}_{FE}$), then there is a costate $\lambda^*(.)$ such that the followings hold:*

A. *(PMP) $u^*(.)$ is the minimizer of $H(x^*, u, \lambda^*, t)$ over $\mathcal{A}$.*

B. *$x^*(.)$ and $\lambda^*(.)$ are solutions of :* $$\begin{array}{lll} \dot{\lambda} = H_x(x, u^*, \lambda, t) & ... \rightarrow & (ADJ) \\ \dot{x} = \psi(x, u^*, t)) & ... \rightarrow & (ODE) \\ x(0) = x_0, x(T) = x_T & ... \rightarrow & (BC) \end{array}$$

3.3.2 Bolza Optimal Control Problem

To characterize the optimal control u^* and optimal trajectory x^* that solves the Bolza OCP with Initial and terminal cost of fixed endpoint.

$$(\mathbf{BOCP}_{FE}) \quad \begin{array}{cc} \min & \varphi(x(0), x(T)) + \int_0^T f(x,u,t)dt \\ s.t & \dot{x} = \psi(x,u,t),\ t \in [0,T] \\ & x(0) = x_0,\ x(T) = x_T \\ & u(.) \in \mathcal{A} \end{array}$$

To characterize the optimal control u^* and optimal trajectory x^* that solves the Bolza OCP with Initial and terminal cost of free endpoint.

$$(\mathbf{BOCP}_{fE}) \quad \begin{array}{cc} \min & \varphi(x(0), x(T)) + \int_0^T f(x,u,t)dt \\ s.t & \dot{x} = \psi(x,u,t),\ t \in [0,T] \\ & x(0) = x_0 \\ & u(.) \in \mathcal{A} \end{array}$$

Using Lagrange method, exactly as described in Section (3.3.1), with $L(x, \dot{x}, u, \lambda, t) = f(x,u,t) - \lambda\psi(x,u,t) + \lambda\dot{x}$, and the Lagrange functional

$$\mathfrak{F}_\lambda(x,u) = \varphi(x(0), x(T)) + \int_0^T L(x, \dot{x}, u, \lambda, t)dt$$

Let $\varphi(x(0), x(T))$=g(x(0),0)+ h(x(T),T) Then

$$\mathfrak{F}_\lambda(x,u) = g(x(0),0) + h(x(T),T) + \int_0^T L(x, \dot{x}, u, \lambda, t)dt$$

all previous results hold , except now the Situation 2 is take u and λ as fixed parameters, (say, $u = u^*$ and $\lambda = \lambda^*$), find the conditions on $x^*(.)$ that solve the problem:

$$\begin{array}{lccr} \min & g(x(0),0) + h(x(T),T) + L(x, \dot{x}, u^*, \lambda^*, t)dt & & \\ s.t & x(0) = x_0 & & (IC) \\ & u(t) \in \mathcal{A}, & & 0 < t < T \end{array}$$

This is Bolza VP with free Endpoint. So, (ELDE) and (TR) will be applicable to it.

$$\begin{array}{lccr} \min & g(x(0),0) + h(x(T),T) + L(x, \dot{x}, u^*, \lambda^*, t)dt & & \\ s.t & x(0) = x_0,\ x(T) = x_T & & (BC) \\ & u(t) \in \mathcal{A}, & & 0 < t < T \end{array}$$

This is Bolza VP with fixed Endpoint. So, (ELDE) and (BC) will be applicable to it. Thus, given BOCP with initial and terminal cost of the form :

$$(\mathbf{BOCP}_{fE}) \quad \begin{array}{ll} \min & \varphi(x(0), x(T)) + \int_0^T f(x,u,t)dt \\ s.t & \dot{x} = \psi(x,u,t),\ t \in [0,T] \\ & x(0) = x_0 \\ & u(.) \in \mathcal{A} \end{array}$$

and letting the Hamiltonian: $H(x,u,\lambda,t) = f(x,u,t) - \lambda\psi(x,u,t)$, we have the following necessary optimality conditions:

Theorem 3.6. *If $u^*(.)$ is an optimal control and $x^*(.)$ is corresponding optimal trajectory of the above ($\boldsymbol{BOCP}_{fE}$), then there is a costate $\lambda^*(.)$ such that the followings hold:*

A. (PMP) $u^*(.)$ is the minimizer of $H(x^*, u, \lambda^*, t)$ over $\mathcal{A}$.

B. $x^*(.)$ and $\lambda^*(.)$ are solutions of :
$$\begin{array}{lll} \dot{\lambda} = H_x(x,u^*,\lambda,t) & ... \rightarrow & (ADJ) \\ \lambda(T) = -h_x(x(T)) & ... \rightarrow & (TR) \\ \dot{x} = \psi(x,u^*,t)) & ... \rightarrow & (ODE) \\ x(0) = x_0 & ... \rightarrow & (IC) \end{array}$$

3.3.3 Sufficient conditions for OCP

Theorem 3.7. *(Necessary and sufficient optimality conditions) Let L be continuous and continuously partially differentiable with regard to the first two components and x^* be a solution of the (VP)with Fixed endpoint*

$$\begin{array}{ll} \min & \int_a^b L(x,\dot{x},t)dt \\ s.t & x \in \mathcal{S} = \{x \in C^1[a,b] : x(a) = x_a,\ x(b) = x_b\} \end{array}$$

- *If the ELDE condition holds and*
- *L is convex with regard to the first two components for all $t \in [a,b]$ (i.e If the Hessian matrix of L with regard to the first two components is positive semi definite for all $t \in [a,b]$.*

Then x^* is optimal solution of the problem.

Theorem 3.8. *Let L be convex with regard to the first two components for all $t \in [a,b]$. Then $x^* \in \mathcal{S} = \{x \in C^1[a,b] : x(a) = x_a, x(b) = x_b\}$ is a minimum point of L on $\mathcal{S} = \{x \in C^1[a,b] : x(a) = x_a,\ x(b) = x_b\}$ if and only if x^* is a solution of ELDE.*

Theorem 3.9. *Let L be continuous and continuously partially differentiable with regard to the first two components and x^* be a solution of the (VP)with free endpoint*

$$\begin{array}{ll} \min & \int_a^b L(x,\dot{x},t)dt \\ s.t & x \in \mathcal{S} = \{x \in C^1[a,b] : x(a) = x_a\} \end{array}$$

- *If the (ELDE) and (TR) condition holds and*
- *L is convex with regard to the first two components for all $t \in [a,b]$ (i.e If the Hessian matrix of L with regard to the first two components is positive semi definite for all $t \in [a,b]$.*

Then x^* is optimal solution of the problem.

Theorem 3.10. *Let $\varphi(x(0), x(T))$=g(x(0),0)+ h(x(T),T) Then $\mathfrak{F}_\lambda(x) = g(x(0),0) + h(x(T),T) + \int_0^T L(x,\dot{x},\lambda,t)dt$ on $\mathcal{S} = \{x \in C^1[a,b] : x(a) = x_a,\ x(b) = x_b\}$. Let g and h be convex, $\mathcal{S}$ be convex set,L be differentiable and convex with regard to the first two components for each $t \in [a,b]$ (i.e If the Hessian matrix of L with regard to the first two components is positive semi definite for all $t \in [a,b]$) .If for each $x^* \in \mathcal{S}$ the conditions*

- *If the (ELDE) condition holds and*
- *(TR) condition holds i.e $L_{\dot{x}}(x(a),\dot{x}(a),a) = g_x(x(a))$ and $L_{\dot{x}}(x(b),\dot{x}(b),b) = -h_x(x(b))$.*

Then x^* is optimal solution of the problem on $\mathcal{S}$.

For convex(or partially convex)problems the Characterizing theorem of convex optimization can be used in order to get sufficient conditions.The following special case is particularly easy and many practical OCP have this structure.Let $L(x,\dot{x},u,\lambda t) = G(x,\dot{x},\lambda t)+W(u,\lambda,t)$. We will denote this types of problems as separated problems. Now consider the following optimization problem :

$$\mathfrak{F}(x,u) = \int_a^b f(x,u,t)dt$$

$$(\mathbf{LOCP}_{fE}) \quad \begin{array}{ll} \min & \mathfrak{F}(x,u) \\ s.t & \dot{x} = \psi(x,u,t) \\ & x(a) = x_a \qquad u(.) \in \mathcal{A} \end{array}$$

And the corresponding Lagrange problem is:

$$\mathfrak{F}_\lambda(x,u) = \int_a^b (f(x,u,t) + \lambda\dot{x} - \lambda\psi(x,u,t))dt =$$

So we have the following types of optimization problems:

- $$\begin{array}{lll} \min & \mathfrak{F}_\lambda(x,u) = \int_a^b G(x,\dot{x},\lambda t) + W(u,\lambda,t)dt & \\ s.t & x(a) = x_a & u(.) \in \mathcal{A} \end{array}$$

  (Separated Problem)

- $$\begin{array}{lll} \min & g(x) = \int_a^b G(x,\dot{x},\lambda t)dt & \\ s.t & x(a) = x_a & u(.) \in \mathcal{A} \end{array}$$

  (Variational problem)

- $$\begin{array}{lll} \min & w(u) = \int_a^b W(u,\lambda,t)dt & \\ s.t & x(a) = x_a & u(.) \in \mathcal{A} \end{array}$$

  (Normal optimization problem)

Theorem 3.11. *Let $\mathfrak{F}_\lambda(x,u) = \int_a^b G(x,\dot{x},\lambda t) + W(u,\lambda,t)dt \rightarrow min$ is separable control problem. If x^* is a solution of Variational problem $g(x)$ And u^* is a solution of Normal optimization problem $w(u)$ then (x^*,u^*) is a solution of Separated Problem $\mathfrak{F}_\lambda(x,u)$.*

Proof. □

3.4 Linear optimal control problems

In this section we consider optimal control problems, where the functions φ and f are linear with respect to x and u,and the differential equation is also linear. That is optimal control problem of the following type:

$$\mathfrak{F}(x,u) = \Upsilon^T x(b) + \int_a^b (c^T x + d^T u)dt$$

$$(\mathbf{LOCP}_{fE}) \quad \begin{array}{lll} \min & \mathfrak{F}(x,u) & \\ s.t & \dot{x} = Ax + Bu, t \in [a,b] & \\ & x(a) = x_a & u(.) \in \mathcal{A} = [-1,1] \end{array}$$

This type of problem is called Linear Optimal Control problem. Then we get the lagrange function

$L(x,\dot{x},u,\lambda,t) = c^T x + d^T u + \lambda^T(\dot{x} - Ax - Bu) = c^T x + \lambda^T \dot{x} - \lambda^T Ax + d^T u - \lambda^T Bu$

So,we have $G(x,\dot{x},\lambda,t) = c^T x + \lambda^T \dot{x} - \lambda^T Ax$ And,also, $W(u,\lambda,t) = d^T u - \lambda^T Bu$. Obviously G is convex(since it is linear) with respect to the pair of variables $(x,\dot{x})$ and W is also convex (since it is linear) with respect to u.

Theorem 3.12. *If $u^*(.)$ is an optimal control and $x^*(.)$ is corresponding optimal trajectory of the above ($\boldsymbol{LOCP}_{fE}$), then there is a costate $\lambda^*(.)$ such that the followings hold:*

A. (PMP) $u^*(.)$ is the minimizer of $H(x^*, u, \lambda^*, t)$ over $\mathcal{A}$.

B. $x^*(.)$ and $\lambda^*(.)$ are solutions of :
$$\begin{array}{lll} \dot{\lambda} = -A^T\lambda + c & ... \rightarrow & (ADJ) \\ \lambda(b) = -\Upsilon & ... \rightarrow & (TR) \\ \dot{x} = Ax + Bu & ... \rightarrow & (ODE) \\ x(a) = x_a & ... \rightarrow & (IC) \end{array}$$

3.5 Quadratic optimal control problems

In this section we consider optimal control problems ,where the functions φ and f are quadratic with respect to x and u,and the differential equation is linear. That is problem of the following type: $\mathfrak{F}(x, u) = \frac{1}{2}x^T(b) \bigwedge x(b) + \Upsilon^T x(b) + \int_a^b (\frac{1}{2}x^T Cx + c^T x + \frac{1}{2}u^T Dx + d^T u)dt$

$$(\mathbf{QOCP}_{fE}) \quad \begin{array}{lll} \min & \mathfrak{F}(x, u) & \\ s.t & \dot{x} = Ax + Bu, t \in [a, b] & \\ & x(a) = x_a & u(.) \in \mathcal{A} \end{array}$$

This type of problem is called Quadratic optimal control problem.Then its Lagrange function is given by

$$\begin{aligned} L(x, \dot{x}, u, \lambda, t) = & \ \tfrac{1}{2}x^T Cx + c^T x + \tfrac{1}{2}u^T Dx + d^T u + \lambda^T(\dot{x} - Ax - Bu) \\ = & \ \tfrac{1}{2}x^T Cx + c^T x + \lambda^T(\dot{x} - Ax) + \tfrac{1}{2}u^T Dx + d^T u - \lambda^T Bu \end{aligned}$$

So, $G(x, \dot{x}, \lambda, t) = \frac{1}{2}x^T Cx + c^T x + \lambda^T(\dot{x} - Ax)$ $W(u, \lambda, t) = \frac{1}{2}u^T Dx + d^T u - \lambda^T Bu$ Obviously G is convex with regard to x and $\dot{x}$, W is convex with regard to u. So we have a sufficient condition for (x^*, u^*) to be a solution of our optimal control problem.

Theorem 3.13. *If $u^*(.)$ is an optimal control and $x^*(.)$ is corresponding optimal trajectory of the above ($\boldsymbol{QOCP}_{fE}$), then there is a costate $\lambda^*(.)$ such that the followings hold:*

A. (PMP) $u^* = D^{-1}B^T\lambda - D^{-1}d$, over $\mathcal{A}$.

B. $x^*(.)$ and $\lambda^*(.)$ are solutions of :
$$\begin{array}{lll} \dot{\lambda} = Cx - A^T\lambda + c & ... \rightarrow & (ADJ) \\ \lambda(b) = -\bigwedge^T x - \Upsilon & ... \rightarrow & (TR) \\ \dot{x} = Ax + Bu^* & ... \rightarrow & (ODE) \\ x(a) = x_a & ... \rightarrow & (IC) \end{array}$$

So From Theorem (3.13) we have the following system of differential equation

$$\begin{cases} \dot{x} = Ax + BD^{-1}B^T\lambda - BD^{-1}d, \\ \dot{\lambda} = Cx - A^T\lambda + c, \\ x(a) = x_a, \ and \ \lambda(b) = -\bigwedge^T x(b) - \Upsilon \end{cases}$$

$$\begin{cases} \begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} A & BD^{-1}B^T \\ C & -A^T \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} + \begin{pmatrix} -BD^{-1}d \\ c \end{pmatrix} \\ x(a) = x_a, \ and \ \lambda(b) = -\bigwedge^T x(b) - \Upsilon \end{cases}$$

The above system of differential equation is differential equation with boundary value problem If we have the solution (x^*, λ^*) of this system of differential equations,then we can calculate the control function u^* by substitute λ^* in the equation $u^* = D^{-1}B^T\lambda^* - D^{-1}d$.
If $\bigwedge = 0$,we have optimal control problem of the form: $\mathfrak{F}(x,u) = \Upsilon^T x(b) + \int_a^b (\frac{1}{2}x^T Cx + c^T x + \frac{1}{2}u^T Dx + d^T u)dt$

$$\begin{array}{llll} & \min & \mathfrak{F}(x,u) & \\ (\mathbf{QOCP}_{fE}) & s.t & \dot{x} = Ax + Bu, t \in [a,b] & \\ & & x(a) = x_a & u(.) \in \mathcal{A} \end{array}$$

So From Theorem (3.13) we have the following system of differential equation

$$\begin{cases} \dot{x} = Ax + BD^{-1}B^T\lambda - BD^{-1}d, \\ \dot{\lambda} = Cx - A^T\lambda + c, \\ x(a) = x_a, \ and \ \lambda(b) = -\Upsilon \end{cases}$$

$$\begin{cases} \begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} A & BD^{-1}B^T \\ C & -A^T \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} + \begin{pmatrix} -BD^{-1}d \\ c \end{pmatrix} \\ x(a) = x_a, \ and \ \lambda(b) = -\Upsilon \end{cases}$$

The above system of differential equation is differential equation with boundary value problem. If we have the solution (x^*, λ^*) of this system of differential equations,then we can calculate the control function u^* by substitute λ^* in the equation $u^* = D^{-1}B^T\lambda^* - D^{-1}d$.
From the quadratic control problem,if $\Upsilon = 0$,$c = 0$,and $d = 0$ then we get $\mathfrak{F}(x,u) = \frac{1}{2}x^T(b) \bigwedge x(b) + \int_a^b (\frac{1}{2}x^T Cx + \frac{1}{2}u^T Dx)dt$

$$\begin{array}{llll} & \min & \mathfrak{F}(x,u) & \\ (\mathbf{LOCP}_{fE}) & s.t & \dot{x} = Ax + Bu, t \in [a,b] & \\ & & x(a) = x_a & u(.) \in \mathcal{A} = [-1,1] \end{array}$$

which is called Linear-Quadratic optimal control problem.

So From Theorem (3.13) we have the following system of differential equation

$$\begin{cases} \dot{x} = Ax + BD^{-1}B^T\lambda, \\ \dot{\lambda} = Cx - A^T\lambda, \\ x(a) = x_a, \ and \ \lambda(b) = -\Upsilon \end{cases}$$

$$\begin{cases} \begin{pmatrix} \dot{x} \\ \dot{\lambda} \end{pmatrix} = \begin{pmatrix} A & BD^{-1}B^T \\ C & -A^T \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} \\ x(a) = x_a, \ and \ \lambda(b) = 0 \end{cases}$$

The above system of differential equation is differential equation with boundary value problem. If we have the solution (x^*, λ^*) of this system of differential equations,then we can calculate the control function u^* by substitute λ^* in the equation $u^* = D^{-1}B^T\lambda^*$.

Chapter 4

Numerical Implementation for QOCP

Using Pontryagin's minimum principle, optimal control problems are converted to two point BVP.

$$y' = F(t, y, \lambda), y(0) = A$$
$$\lambda' = G(t, y, \lambda), \lambda(T) = B$$

There are many ways of solving two point BVP but for this project we will be using the shooting method technique. Here the above two point BVP convert to an IVP.

$$y' = F(t, y, \lambda), y(0) = A$$
$$\lambda' = G(t, y, \lambda), \lambda(0) = \alpha$$

Since guessing is not very efficient we will use an ODE solver to solve our IVP at each guess. We can then implement a root finding algorithm to approach the correct solution after the first initial guess on the IVP. We can relate this process to root finding by defining the shoot function. Note $\lambda(\alpha, t)$ is the solution to the IVP at time t. Then the shooting function to be defined as $f(\alpha) = \lambda(\alpha, T) - B$. For each root finding algorithm method there will be a corresponding shooting method if we combine the root finding method and compute the shooting function.

Example 4.0.1. Find the optimal state function and the optimal control function of the following quadratic optimal control problem.

$$\mathfrak{F}(x, u) = \Upsilon^T x(b) + \int_0^1 (\frac{1}{2}x^T Cx + c^T x + \frac{1}{2}u^T Dx)dt$$

$$\begin{array}{lll} \min & \mathfrak{F}(x,u) & \\ s.t & \dot{x} = Ax + Bu, t \in [0,1] & \\ & x(0) = (\frac{-1}{3}, \frac{-1}{3})T & u(.) \in \mathcal{A} = [-1,1] \end{array} \quad \text{where } A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, B =$$

$$\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, C = D \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, c = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \Upsilon = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$$

Solution 4.0.1. Since the given optimal control is quadratic optimal control problem with $d = 0$ from the Minimum condition, we have

$$u = D^{-1}B^T\lambda = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

(I). From Validity of ODE:

$$\dot{x} = Ax+Bu = Ax+BD^{-1}B^T\lambda = Ax+\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

(II.) From Validity of (ElDE) we have :

$$\dot{\lambda} = Cx - A^T\lambda + c = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

(III.) From TR and (IC) Condition we have

$$\lambda(1) = -\Upsilon = \begin{pmatrix} \frac{-2}{3} \\ \frac{-1}{3} \end{pmatrix} \, and \; x(0) = (\frac{-1}{3}, \frac{-1}{3})T$$

So,we let's solve the following system of differential equations

$$\begin{cases} \dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}, \\ \dot{\lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ x(0) = \begin{pmatrix} \frac{-1}{3} \\ \frac{-1}{3} \end{pmatrix}, \; and \; \lambda(1) = \begin{pmatrix} \frac{-2}{3} \\ \frac{-1}{3} \end{pmatrix} \end{cases}$$

This system of differential equations can be write as

$$\begin{cases} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{\lambda}_1 \\ \dot{\lambda}_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} \frac{-1}{3} \\ \frac{-1}{3} \end{pmatrix}, \; and \; \begin{pmatrix} \lambda_1(1) \\ \lambda_2(1) \end{pmatrix} = \begin{pmatrix} \frac{-2}{3} \\ \frac{-1}{3} \end{pmatrix} \end{cases}$$

Now,First obtain the solution of the homogeneous system part:

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = \begin{pmatrix}-1 & 0 & 1 & -1\\ 0 & 1 & -1 & 1\\ 1 & 0 & 1 & 0\\ 0 & 1 & 0 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \lambda_1\\ \lambda_2\end{pmatrix}$$

To find the solution of the homogeneous system part first compute the eigenvalues the matrix for the homogeneous system par of the system. let $||A-\alpha I_4|| = 0$,which is

$$\begin{Vmatrix}-1-\alpha & 0 & 1 & -1\\ 0 & 1-\alpha & -1 & 1\\ 1 & 0 & 1-\alpha & 0\\ 0 & 1 & 0 & -1-\alpha\end{Vmatrix} = 0 \Rightarrow (\alpha-1)(\alpha^3+\alpha^2-3\alpha-3) = (\alpha-1)(\alpha+1)((\alpha^2-3)) = 0$$

Hence $\alpha_1 = 1$, $\alpha_2 = -1$, $\alpha_3 = \sqrt{3}$ and $\alpha_3 = -\sqrt{3}$ are the eigenvalues the matrix for the homogeneous system par of the system.we have to solve the following

$$\begin{pmatrix}-1-\alpha & 0 & 1 & -1\\ 0 & 1-\alpha & -1 & 1\\ 1 & 0 & 1-\alpha & 0\\ 0 & 1 & 0 & -1-\alpha\end{pmatrix}\begin{pmatrix}v_1\\ v_2\\ v_3\\ v_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

For $\alpha = -1$,we have

$$\begin{pmatrix}0 & 0 & 1 & -1\\ 0 & 2 & -1 & 1\\ 1 & 0 & 2 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}v_1\\ v_2\\ v_3\\ v_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix} \Rightarrow K_1 = \begin{pmatrix}-2\\0\\1\\1\end{pmatrix}e^{-t}$$

For $\alpha = 1$,we have

$$\begin{pmatrix}-2 & 0 & 1 & -1\\ 0 & 0 & -1 & 1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & -2\end{pmatrix}\begin{pmatrix}v_1\\ v_2\\ v_3\\ v_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix} \Rightarrow K_2 = \begin{pmatrix}0\\2\\1\\1\end{pmatrix}e^{t}$$

For $\alpha = \sqrt{3}$,we have

$$\begin{pmatrix}-1-\sqrt{3} & 0 & 1 & -1\\ 0 & 1-\sqrt{3} & -1 & 1\\ 1 & 0 & 1-\sqrt{3} & 0\\ 0 & 1 & 0 & -1-\sqrt{3}\end{pmatrix}\begin{pmatrix}v_1\\ v_2\\ v_3\\ v_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix} \Rightarrow K_3 = \begin{pmatrix}1-\sqrt{3}\\1+\sqrt{3}\\-1\\1\end{pmatrix}e^{\sqrt{3}t}$$

For $\alpha = -\sqrt{3}$,we have

$$\begin{pmatrix} -1+\sqrt{3} & 0 & 1 & -1 \\ 0 & 1+\sqrt{3} & -1 & 1 \\ 1 & 0 & 1+\sqrt{3} & 0 \\ 0 & 1 & 0 & -1+\sqrt{3} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow K_4 = \begin{pmatrix} 1+\sqrt{3} \\ 1-\sqrt{3} \\ -1 \\ 1 \end{pmatrix} e^{-\sqrt{3}t}$$

So the general solution of the homogeneous system par of the system is given by:

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \lambda_1(t) \\ \lambda_2(t) \end{pmatrix} = a_1 \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} e^{-t} + a_2 \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix} e^{t} + a_3 \begin{pmatrix} 1-\sqrt{3} \\ 1+\sqrt{3} \\ -1 \\ 1 \end{pmatrix} e^{\sqrt{3}t} + a_4 \begin{pmatrix} 1+\sqrt{3} \\ 1-\sqrt{3} \\ -1 \\ 1 \end{pmatrix} e^{-\sqrt{3}t}$$

To obtain the particular solution of the system suppose $\begin{pmatrix} x_1 \\ x_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix}$ This implies that

$$\begin{pmatrix} \dot{k}_1 \\ \dot{k}_2 \\ \dot{k}_3 \\ \dot{k}_4 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_1 \\ k_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ k_1 \\ k_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -k_1+k_3-k_4 \\ k_2-k_3+k_4 \\ k_1+k_3 \\ k_2+k_4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} -k_1+k_3-k_4 \\ k_2-k_3+k_4 \\ k_1+k_3 \\ k_2+k_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

Hence, $k_1 = 0$, $k_2 = 0$, $k_3 = -1$ and $k_4 = 0$,

$$\begin{pmatrix} x_1 \\ x_2 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

So, The general solutions for the system of differential equation

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ \lambda_1(t) \\ \lambda_2(t) \end{pmatrix} = a_1 \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} e^{-t} + a_2 \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix} e^{t} + a_3 \begin{pmatrix} 1-\sqrt{3} \\ 1+\sqrt{3} \\ -1 \\ 1 \end{pmatrix} e^{\sqrt{3}t} + a_4 \begin{pmatrix} 1+\sqrt{3} \\ 1-\sqrt{3} \\ -1 \\ 1 \end{pmatrix} e^{-\sqrt{3}t} + \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}$$

$$\begin{cases} x_1(t) = -2a_1e^{-1} + a_3(1-\sqrt{3})e^{\sqrt{3}t} + a_4(1+\sqrt{3})e^{-\sqrt{3}t}, \\ x_2(t) = 2a_2e^{t} + a_3(1+\sqrt{3})e^{\sqrt{3}t} + a_4(1-\sqrt{3})e^{-\sqrt{3}t}, \\ x_1(0) = -\frac{1}{3}, x_2(0) = -\frac{1}{3} \end{cases}$$

and

$$\begin{cases} \lambda_1(t) = a_1e^{-1} + a_2e^{1} - a_3e^{\sqrt{3}t} - a_4e^{-\sqrt{3}t} - 1, \\ \lambda_2(t) = a_1e^{-1} + a_2e^{1} + a_3e^{\sqrt{3}t} + a_4e^{-\sqrt{3}t}, \\ \lambda_1(1) = \frac{-2}{3}, \lambda_2(1) = \frac{-2}{3} \end{cases}$$

When we solve the above system of linear equations,we obtain the following results $a_1 = 1.4$, $a_2 = 12.3$, $a_3 = -0.2$,and $a_4 = -0.4$ Hence,

$$\begin{cases} \lambda_1^*(t) = 1.4e^{-1} + 12.3e^{1} + 0.2e^{\sqrt{3}t} + 0.4e^{-\sqrt{3}t} - 1, \\ \lambda_2^*(t) = 1.4e^{-1} + 12.3e^{1} - 0.2e^{\sqrt{3}t} - 0.4e^{-\sqrt{3}t}, \end{cases}$$

But from Minimum condition we have, $u^* = D^{-1}B^T\lambda^*$

$$\begin{pmatrix} u_1^* \\ u_2^* \end{pmatrix} = \begin{pmatrix} \lambda_1^* - \lambda_2^* \\ 0 \end{pmatrix}$$

so, $\begin{cases} u_1^*(t) = -24.3e^{-\sqrt{3}t} + 0.8e^{\sqrt{3}t} - 1, \\ u_2^*(t) = 0, \end{cases}$

Hence The optimal control function and the optimal state function of the problem are:

$$\begin{cases} x_1^*(t) = -2.8e^{-1} - 0.2(1-\sqrt{3})e^{\sqrt{3}t} - 0.4(1+\sqrt{3})e^{-\sqrt{3}t}, \\ x_2^*(t) = 24.6e^{t} - 0.2(1+\sqrt{3})e^{\sqrt{3}t} - 0.4(1-\sqrt{3})e^{-\sqrt{3}t}, \\ \\ u_1^*(t) = -24.3e^{-\sqrt{3}t} + 0.8e^{\sqrt{3}t} - 1, \\ u_2^*(t) = 0, \end{cases}$$

Example 14. We consider the optimal control problem:

$$\begin{cases} min J(u); J(u) = \frac{1}{2}\int_0^{t_f} u^2(t)dt \\ s.t\dot{x}(t) = -x(t) + u(t), u(t) \in \mathbb{R}, t \in [0, t_f] \\ x(0) = x_o, x(t_f) = x_f \end{cases}$$

with $t_f = 1, x_o = -1, x_f = 0$ and $\forall(t) \in [0, t_f], x(t) \in \mathbb{R}$.

SOLUTION First application of Pontryagin minimum principle, we have the Hamiltonian $H(x,p,u) = p(-x+u) + p^o\frac{1}{2}u^2$ and the Pontryagin minimum principle $\begin{cases} \dot{x(t)} = \nabla_p H[t] = -x(t) + u(t), \\ \dot{p(t)} = -\nabla_x H[t] = p(t), \\ 0 = \nabla_u H[t] = p(t) + p^o u(t), \end{cases}$ where $[t] = (x(t), p(t), u(t))$.

If $p^o = 0$, then $p = 0$ by the third equation and so $(p, p^o) = (0, 0)$ which is not. Hence any extremal x, p, p^o, u given by the Pontryagin minimum principle is said to be normal, that is $p^o = -1$.. We do not consider the transversal condition when the target x_f is fixed. We can just retrieve the Lagrange multiplier by the relation $p(t_f) = \lambda$. By concavity $\nabla_u H[t] = 0$ since we have convex objective function and the control satisfies $u(t) = u[x(t), p(t)] = p(t)$ where we have introduced the smooth function on $\mathbb{R} \times \mathbb{R} : u[x, p] = p$. Second we have the control in feedback form, we introduce the following smooth two points BVP:

$$\begin{cases} \dot{x(t)} = -x(t) + u[x(t), p(t)] = -x(t) + p(t), \\ \dot{p(t)} = p(t), \\ x(0) = x_o, s.t x(t_f) = x_f \end{cases}$$

The unknown of this BVP is the initial covector $p(0)$. Indeed, fixing $p_o = p(0)$, then according to the Cauchy-Lipschits theorem, there exists a unique solution denoted $z(.x_o, p_o) = (x(., x_o, p_o), p(., x_o, p_o))$ satisfying the dynamic $\dot{z(t)} = (-x(t) + p(t), p(t)$ together with the initial condition $z(0) = (x_o, p_o)$. Here our goal is thus to find the right initial covector p_o such that $x(t_f, x_o, p_o) = x_f$. Thirdly solving the shooting equation; from $\dot{p(t)} = p(t)$, we get

$$p(t, x_o, p_o) = e^t p_o$$

which leads to

$$x(t, x_o, p_o) = p_o sinh(t) + x_o e^{-t}.$$

Solving $x(t_f, x_o, p_o) = x_f$, we obtain

$$p_o^* = \frac{x_f - x_o e^{-t_f}}{sinh(t_f)} = \frac{2}{e^2 - 1} \approx 0.313$$

. To compute p_o^*, we have solved the linear shooting equation

$$S(p_o) = \pi_x(z(t_f, x_o, p_o)) - x_f = p_o sinh(t_f) + x_o e^{-t_f} - x_f$$

with $\pi_x(x, p) = x$. Solving $S(p_o) = 0$ is what we call the shooting method.

Conclusion

This paper tries to give optimal solutions of the variational problems and OCP by using optimality conditions with shooting method. The optimal solutions are summarized as follows.
a)consider variational problem:

$$\begin{array}{ll} \min & \int_a^b L(x, \dot{x}, t)dt \\ s.t & (x, \dot{x}) \in S \end{array}$$

Let L is continuous and continuously partially differentiable with respect to the first two arguments and let y_0 is a solution of the variational problem. For y_0 to be a solution of the variational problem then (ELDE) and (TR) condition(for the problem of free right end point) hold and if L is convex with regard to the first two components for all t in $[a, b]$ then y_0 is the optimal solution of the problem.
From the given practical examples of minimizing length of a curve and minimizing surface area of solid revolution; variational calculus gives us economical advantage interims of time and material resources.
Consider an optimal control problem:

$$\begin{array}{ll} \min & \int_a^b L(x, u, t)dt \\ s.t & (x, u) \in R \end{array}$$

The necessary optimality conditions for $x^*(t)$ and $u^*(t)$ to be extremal functions of the problem are the validity of

- Pontryagin minimum principle
- Euler Lagrange Differential Equation(ELDE) and Transversalis condition(TR)(for the problem of free right end point).
- The ordinary differential equation(ODE)

And a sufficient optimality condition for the extremals $x^*(t)$ and $u^*(t)$ to be the optimal solutions of the problem is the convexity of the separated

problems.i.e the convexity of g with regard to x and $\dot{x}$, and the convexity of W with regard to u. If these necessary and sufficient conditions are satisfied then the extremals $x^*(t)$ and $u^*(t)$ are the optimal solution of he problem.
In the case of Quadratic optimal control problems; the separated problems g and W are convex with respect to $(x, \dot{x})$ and u respectively(i.e a sufficient condition holds). If the three necessary conditions are satisfied then the extremals $x^*(t)$ and $u^*(t)$ are the optimal solutions of the problem.
Using Pontryagin's minimum principle, optimal control problems are converted ton two point BVP. There are many ways of solving two point BVP but for this project we have discussed the shooting method technique. The idea of the shooting method is to convert the two point BVP into an IVP. It is convenient to solve this IVP since it has one condition on one hand of the interval and another condition on the other end of the interval; where we can guess the condition at the beginning of the interval.

Bibliography

[1] David C.Lay . Steven R.Lay . Jude J. McDonald: Linear Algebra and its applications (5^{th} Edition), PEARSON.

[2] Robert Ellis - Denny Gulick: Calculus, University of Maryland at College Park, 6^{th} edition.

[3] Optimization - Dr. Birhanu Guta: Optimal control, Addis Ababa University, 2020/21 lecture notes.

[4] H.L.Royden and P.M.Fitzpatrick: Real Analysis, American Mathematical Society, 4^{th} edition.

[5] MOKHTAR S. BAZARAA, HANIF D. SHERALI, C. M. SHETTY: NONLINEAR PROGRAMMING Theory and Algorithms, 3^{rd} edition.

[6] Dr. Birhanu Guta: Lecture Notes on Optimization and Theory of Approximation (Nonlinear Programming), Math 652, Addis Ababa University, 2021.

[7] E.A. Coddington: *Theory of ODE*, McGraw-Hill Book Company, New York, 1984.

[8] Kibremarkos Aklilu: Quadratic Optimal Control, Addis Ababa University, 2017.

[9] Antole K. & Boris H: Introduction to *Modern Theory of Dynamical Systems*.

[10] Michael B/. & Garrett S: *Introduction to Dynamical Systems*.

[11] Ross & Weinberger: *Introduction to Differential Equations*, Blaisdell Publishing Company, a division of Ginn and Company.

[12] Ritger-Rose & Dennmeyer: *Differential Equations*, McGraw-Hill, USA, 1968.

[13] Murray R. Spiegel $ProfessorofMathematicsatRensselaerPolytechnicInstitute$: Applied Differential Equations, Instituto Del Libro, 1968.

[14] Richard L. Burden and J. Douglas Faires: Numerical Analysis, Youngstown State University, Seventh edition 2001.

[15] Philip E.Gill Walter Murray and Margaret H.Wright: Practical Optimization, Stanford University, California, USA 1981

[16] The MATH WORKS Inc. The Student Edition of MATLAB, The

Ultimate Computing Environment for Technical Education, Version 4 User's Guid 1995, ISBN 0-13-184979-4

[17] Shepley L.Ross (University of New Hampshire) and Shepley L.Ross, II (Bates College): Introduction to Ordinary Differential Equations, 1989.

[18] Charles E.Roberts, Jr Textbooks in mathematics Ordinary Differential Equations, Applications, Models, and Computing: Indiana State University, Terre Haute, USA 2010

[19] William E. Boyce and Richard C. DiPrima: Elementary Differential Equations and Boundary Value Problems, Fifth edition, Rensselaer Polytechnic Institute, 1992

[20] M. K. Jain, S. R. K. Iyengar, and R. K. Jain: "Numerical Methods for Scientific and Engineering Computation", New Age International.

[21] S. Arumugam, A. Thanga Paandi Isaac, and A. Somasundaram: "Numerical Methods", SciTech Publications, Chennai, India.

[22] B. S. Grawel: "Numerical Methods in Engineering and Science", Khanna Publishers, India.

[23] Singiresu S. Rao, (2002). Applied Numerical Methods for Engineers and Scientists University of Miami, Coral Gables, Florida, Prentice Hall, Upper Saddle River, NJ 07458, ISPN 013089480X

[24] Department of Mathematics, Guru Nanak Dev University Amritsar-143005, Punjab, INDIA.

[25] https://www.researchgate.net/publication/2339223_Hilbert_Space_Methods_in_Partial_Differential_Equations

[26] https://mathworld.wolfram.com/InnerProduct.html

[27] https://www.jstor.org/stable/1995875

[28] https://pastebin.com/TXE4UuT7

[29] https://imgur.com/a/dQkQDZM

[30] https://imgur.com/a/K7dwlja