Summer Internship Project
Report

Fundamental Theorems of Optimization

Submitted by
Sandipan Samanta
Second Year Integrated MSc.
National Institue of Science Education and Research,
Bhubaneswar

Under the guidance of
Dr. Anupam Pal Choudhury

School of Mathematical Sciences
NATIONAL INSTITUTE OF SCIENCE EDUCATION AND RESEARCH,
BHUBANESWAR

Summer Internship 2023

Acknowledgement

I would like to express my heartfelt gratitude to all those who have supported and guided me throughout the journey of completing this project report.

First and foremost, I am deeply indebted to my project guide, Dr. Anupam Pal Choudhury, for his unwavering support, valuable insights, and guidance. His expertise and encouragement have been instrumental in shaping the direction of this project.

I would also like to extend my appreciation to my friends and seniors who have been a constant source of inspiration and assistance. Their willingness to share their knowledge and experiences has been invaluable in navigating through the complexities of this project.

A special note of thanks goes to Mr. Harihara Maharma from IISER Trivandrum. Our collaborative efforts and countless study sessions under the mentorship of Dr. Anupam Pal Choudhury have significantly enriched my understanding of the subject matter.

Last but not least, I would like to acknowledge the collective support of my family, whose encouragement and understanding have been pivotal in ensuring the successful completion of this project.

Thank You
Sandipan Samanta

Contents

Introduction **3**

1 Continuity and Differentiability in Normed Spaces **4**
1.1 Introduction to Normed Spaces 4
1.2 Continuity in Normed Spaces 5
1.3 Differentiation in Normed Spaces 8

2 Condition for Optimization of Functions **12**
2.1 Convex Sets and Convex Functions 12
2.2 Role of Vanishing Derivative in Optimization 12
2.3 Sufficiency of Vanishing Derivative in Optimizing Convex Functions 13

3 Optimization: Euler-Lagrange Equation as a Special Case **15**
3.1 Euler-Lagrange Equation: Fixed End Points 15

References **19**

Introduction

This report is all about understanding how to make things work best using a kind of math called "optimization." We're going to explore a special kind of math space where these ideas work, and it's called "normed spaces."

Step by step, we're going to learn the main ideas of how this optimization stuff works. We'll start by understanding how things change and behave smoothly in these special math spaces. Then, we'll dive into the rules that help us find the best answers for certain math problems.

A big focus of this report is something called the "Euler-Lagrange equation." It's like a super tool that helps us find the very best solutions to problems. We'll break down how this equation works when things are described in these normed spaces.

It's worth noting that this report is more about the ideas and theories behind how things can be made better. But these ideas are like the building blocks that can help solve problems in the real world. By diving into this report, you'll be getting a solid understanding of how to tackle tough math challenges.

1 Continuity and Differentiability in Normed Spaces

In the realm of optimization, the study of functions defined on normed spaces provides a rich framework for understanding and solving complex problems. The concept of continuity and differentiability is well-established for functions from $\mathbb{R}$ to $\mathbb{R}$, so the primary objective of this section is to extend the fundamental concepts of continuity and differentiability in normed spaces. We begin by visiting the notion of normed spaces, which encompass a wide range of function spaces and provide a metric structure for measuring distances and defining convergence.

1.1 Introduction to Normed Spaces

Definition 1. *A* ***normed space*** *is a vector space equipped with a norm, which is a function that assigns a non-negative real number to each vector in the space. Formally, a norm on a vector space X is a function $\|\cdot\| : X \to \mathbb{R}$ that satisfies the following properties for all vectors $x, y \in X$ and scalar $\alpha \in \mathbb{R}$:*

- *Non-negativity:* $\|x\| \geq 0$, *and* $\|x\| = 0$ *if and only if* $x = \mathbf{0}$
- *Homogeneity:* $\|\alpha x\| = |\alpha|\|x\|$
- *Triangle inequality:* $\|x + y\| \leq \|x\| + \|y\|$

Metric structure of normed spaces: The norm induces a metric, or a distance function, on the normed space. This metric enables the measurement of distances between vectors and provides a notion of convergence and continuity. In normed space $(X, \|\cdot\|)$, if we define $d : X \times X \to \mathbb{R}$ by $d(x, y) = \|x - y\|$ for $x, y \in X$, then (X, d) is a metric space with the metric/distance function d.

Examples of normed spaces:

- Euclidean Spaces: In familiar spaces $\mathbb{R}^n$, if we define $\|x\|_p = \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$ for $p \in \mathbb{N}$ and $\|x\|_\infty = \sup\{|x_i| : i = 1, 2, \ldots, n\}$ where $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, then $(\mathbb{R}^n, \|\cdot\|_\infty)$ and $(\mathbb{R}^n, \|\cdot\|_p)$ for $p \in \mathbb{N}$ are different normed spaces, though most common being the $\|\cdot\|_2$ on $\mathbb{R}^n$

- Sequence spaces: Sequence spaces, denoted as ℓ^p, consist of sequences of real or complex numbers that possess certain convergence properties. The p-norm, or ℓ^p norm, is commonly associated with sequence spaces and provides a measure of the ”size” of a sequence. The p-norm is defined as:

  $\|x\|_p = \left(\sum\limits_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}$ where $x = (x_n)$ represents a sequence.

  Depending on the value of p, different sequence spaces are obtained. Here are a few examples:

  (i) ℓ^1 space (Manhattan norm): $\|x\|_1 = |x_1| + |x_2| + \cdots$

(ii) ℓ^2 space (Euclidean norm): $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots}$

(iii) ℓ^∞ space (Supremum norm): $\|x\|_\infty = \sup\{|x_i| : i \in \mathbb{N}\}$

- Function Spaces (L^p): Function spaces encompass a broad class of spaces consisting of functions with specific properties. Different function spaces are associated with different norms, reflecting characteristics such as integrability or smoothness. These spaces can be thought of continuous analogue of the sequence spaces. Here are a few examples:

  (i) L^p space: For a real valued function f defined on a measurable, the L^p norm is defined as:

  $\|f\|_p = \left(\int |f(x)|^p dx \right)^{\frac{1}{p}}$

  (ii) $C[a,b]$ space: It is defined as the space of continuous functions from $[a,b]$ to $\mathbb{R}$. The L^p norms can be defined here as well, $\|f\|_p = \left(\int_a^b |f(x)|^p dx \right)^{\frac{1}{p}}$, but the most common is the supremum or uniform norm, defined as:

  $\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$

  Spaces like $C^1[a,b]$ which represents the space of continuously differentiable real valued function on $[a,b]$ can also be equipped with the supremum norm. However, it turns out that in applications, this is not a good choice (we will discuss this later on in the context of continuity). So, we'll use a different norm on $C^1[a,b]$ given below:

  $\|f\|_{1,\infty} = \|f\|_\infty + \|f'\|_\infty$, $f \in C^1[a,b]$, where f' refers to the derivative of f.

1.2 Continuity in Normed Spaces

Definition 2. *A function $f : X \to Y$, where X and Y are normed spaces, is said to be continuous at a point $x_0 \in X$ if, for any $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $x \in X$ satisfies $\|x - x_0\| < \delta$, it follows that $\|f(x) - f(x_0)\| < \epsilon$.*

The map $f : X \to Y$ is called continuous if f is continuous at x_0 for all $x_0 \in X$

Example 3. *Define $D : C^1[0,1] \to C[0,1]$ by*

$$(D\mathbf{x})(t) = \mathbf{x}'(t), \quad t \in [0,1], \quad \mathbf{x} \in C^1[0,1].$$

Examine the continuity of f when:

(i) both $C^1[0,1]$ and $C[0,1]$ are equipped with $\|\cdot\|_\infty$

(ii) $C^1[0,1]$ is equipped with $\|\cdot\|_{1,\infty}$ and $C[0,1]$ with $\|\cdot\|_\infty$

Note. Here, the D is "differentiation pointwise"

Solution.

(i) Let $\mathbf{x_0} \in C^1[0,1]$ and $\epsilon > 0$.
For any $\delta > 0$, by Archimedean property, there exists $n \in \mathbb{N}$ such that $n > \dfrac{2\epsilon}{\delta}$
Consider $\mathbf{x} \in C^1[0,1]$ given by $\mathbf{x}(t) = \mathbf{x_0}(t) + \dfrac{\delta}{2}t^n$.Then,

$$\|\mathbf{x} - \mathbf{x_0}\|_\infty = \left\|\frac{\delta}{2}t^n\right\|_\infty = \left|\frac{\delta}{2}\right| \|t^n\|_\infty = \frac{\delta}{2} < \delta$$

But,

$$\|D\mathbf{x} - D\mathbf{x_0}\|_\infty = \|\mathbf{x}' - \mathbf{x_0}'\|_\infty = \left\|\frac{n\delta}{2}t^{n-1}\right\|_\infty = \left|\frac{n\delta}{2}\right| \|t^{n-1}\|_\infty = \frac{n\delta}{2} > \epsilon$$

As $\mathbf{x_0}$ and δ are arbitrary, so D is nowhere conitnuous on $C^1[0,1]$ equipped with the norm $\|\cdot\|_\infty$

(ii) Let $\mathbf{x_0} \in C^1[0,1]$ and $\epsilon > 0$. Set $\delta = \epsilon$. Then $\delta > 0$ and whenever $\mathbf{x} \in C^1[0,1]$ satisfies $\|\mathbf{x} - \mathbf{x_0}\|_{1,\infty} < \delta$, we have

$$\|D\mathbf{x} - D\mathbf{x_0}\|_\infty = \|\mathbf{x}' - \mathbf{x_0}'\|_\infty \leq \|\mathbf{x} - \mathbf{x_0}\|_{1,\infty} < \delta = \epsilon$$

As $\mathbf{x_0} \in C^1[0,1]$ is arbitrary, it follows that D is continuous on $C^1[0,1]$ equipped with $\|\cdot\|_{1,\infty}$

Definition 4 (Linear Transformation)**.** *Let X and Y be vector spaces over $\mathbb{R}$. A map $T : X \to Y$ is said to be a linear transormation if it satisfies the following:*

- *For all $x_1, x_2 \in X$, $\quad T(x_1 + x_2) = T(x_1) + T(x_2)$.*
- *For all $x \in X$ and all $\alpha \in \mathbb{R}$, $\quad T(\alpha \cdot x) = \alpha \cdot T(x)$*

Theorem 5 (Continuous Linear transformation)**.** *Let x and Y be normed spaces over $\mathbb{R}$, and $T : X \to Y$ be a linear transformation. Then are equivalent:*

(i) *T is continuous.*
(ii) *T is continuous at 0.*
(iii) *There exists an $M > 0$ such that for all $x \in X$, $\quad \|T(x)\| \leq M\|x\|$*

Proof. We will show the three implications (i) $\implies$ (ii), (ii) $\implies$ (iii) and (iii) $\implies$ (i) for proving all the three equivalences.

(i) $\implies$ (ii). This holds by the definition of continuity on X, i.e., if T is continuous on X, then T is continuous on x_0 for all $x_0 \in X$, so T is continuous at 0.

(ii) $\implies$ (iii). Take $\epsilon = 1 > 0$. Then there exists a $\delta > 0$ such that $\|x - \mathbf{0}\| = \|x\| < \delta \implies \|T(x) - T(\mathbf{0})\| = \|T(x)\| < 1$

Claim: $\|T(x)\| \leq \dfrac{2}{\delta}\|x\|$ for all $x \in X$.

Clearly, when $x = 0$, LHS = RHS.

Now, if $x \in X,\ x \neq \mathbf{0}$. Set $y = \dfrac{\delta}{2\|x\|} \cdot x$.Then

$$\|y\| = \left\| \frac{\delta}{2\|x\|} \cdot x \right\| = \frac{\delta}{2\|x\|}\|x\| = \frac{\delta}{2} < \delta$$

, so $\|T(y)\| < 1$, i.e.,

$$\left\| T\left(\frac{\delta}{2\|x\|} \cdot x\right) \right\| = \left\| \frac{\delta}{2\|x\|} \cdot T(x) \right\| = \frac{\delta}{2\|x\|}\|T(x)\| < 1 \implies \|T(x)\| \leq \frac{2}{\delta}\|x\|$$

So (iii) holds with $M = \dfrac{2}{\delta}$.

(iii) $\implies$ (i). Let $M > 0$ be such that for all $x \in X,\ \|T(x)\| \leq M\|x\|$. Let $x_0 \in X$ and $\epsilon > 0$ be arbitrary. Take $\delta = \dfrac{\epsilon}{M} > 0$. Then whenever $\|x - x_0\| < \delta$, we have

$$\|T(x) - T(x_0)\| = \|T(x - x_0)\| \leq M\|x - x_0\| < M\delta = \epsilon$$

So T is continuous on X $\square$

Example 6. *Let $S := \{\mathbf{h} \in C^1[a, b] : \mathbf{h}(a) = \mathbf{h}(b) = 0\}$. Let $\mathbf{A}, \mathbf{B} \in C[a, b]$ be two fixed functions. Consider the map $L : S \to \mathbb{R}$ given by*

$$L(\mathbf{h}) = \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big)dt, \quad \mathbf{h} \in S.$$

Show that L is a continuous linear transformation.

Solution. Let us first check that L is a linear transformation. We have:

(i) For all $\mathbf{h_1}, \mathbf{h_2} \in S$,

$$\begin{aligned}
&L(\mathbf{h_1} + \mathbf{h_2}) \\
&\quad = \int_a^b \Big(\mathbf{A}(t)(\mathbf{h_1} + \mathbf{h_2})(t) + \mathbf{B}(t)(\mathbf{h_1'} + \mathbf{h_2'})(t)\Big)dt \\
&\quad = \int_a^b \Big(\mathbf{A}(t)(\mathbf{h_1}(t) + \mathbf{h_2}(t)) + \mathbf{B}(t)(\mathbf{h_1'}(t) + \mathbf{h_2'}(t))\Big)dt \\
&\quad = \int_a^b \Big(\mathbf{A}(t)\mathbf{h_1}(t) + \mathbf{B}(t)\mathbf{h_1'}(t) + \mathbf{A}(t)\mathbf{h_2}(t) + \mathbf{B}(t)\mathbf{h_2'}(t)\Big)dt \\
&\quad = \int_a^b \Big(\mathbf{A}(t)\mathbf{h_1}(t) + \mathbf{B}(t)\mathbf{h_1'}(t)\Big)dt + \int_a^b \Big(\mathbf{A}(t)\mathbf{h}_2(t) + \mathbf{B}(t)\mathbf{h_2'}(t)\Big)dt \\
&\quad = L(\mathbf{h_1}) + L(\mathbf{h_2})
\end{aligned}$$

(ii) For all $\mathbf{h} \in S$ and $\alpha \in \mathbb{R}$,

$$\begin{aligned}
L(\alpha \cdot \mathbf{h}) &= \int_a^b \Big(\mathbf{A}(t)(\alpha \cdot \mathbf{h}()t) + \mathbf{B}(t)(\alpha \cdot \mathbf{h}')(t)\Big) dt \\
&= \int_a^b \Big(\mathbf{A}(t)\alpha\mathbf{h}(t) + \mathbf{B}(t)\alpha\mathbf{h}'(t)\Big) dt \\
&= \alpha \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big) dt \\
&= \alpha L(\mathbf{h})
\end{aligned}$$

So, L is a linear transformation. Next, we will show L is continuous using Theorem 5. We have for $\mathbf{h} \in S$ that

$$\begin{aligned}
|L(\mathbf{h})| &= \left| \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big) dt \right| \\
&\leq \int_a^b \Big|\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big| dt \\
&\leq \int_a^b \Big(\big|\mathbf{A}(t)\big|\big|\mathbf{h}(t)\big| + \big|\mathbf{B}(t)\big|\big|\mathbf{h}'(t)\big|\Big) dt \\
&\leq \int_a^b \Big(\big|\mathbf{A}(t)\big|\|\mathbf{h}\|_\infty + \big|\mathbf{B}(t)\big|\|\mathbf{h}'\|_\infty\Big) dt \\
&\leq \int_a^b \Big(\big|\mathbf{A}(t)\big|\|\mathbf{h}\|_{1,\infty} + \big|\mathbf{B}(t)\big|\|\mathbf{h}\|_{1,\infty}\Big) dt \\
&\leq \left(\int_a^b \Big(\big|\mathbf{A}(t)\big| + \big|\mathbf{B}(t)\big|\Big) dt \right) \|\mathbf{h}\|_{1,\infty} = M\|\mathbf{h}\|_{1,\infty},
\end{aligned}$$

where $M = \displaystyle\int_a^b \Big(\big|\mathbf{A}(t)\big| + \big|\mathbf{B}(t)\big|\Big) dt$. Thus L is continuous using Theorem 5.

1.3 Differentiation in Normed Spaces

Definition 7 (Frechet derivative)**.** *Let X, Y be normed spaces, $f : X \to Y$ be a map, and $x_0 \in X$. f is said to be differentiable at x_0 if there exists a continuous linear transformation $L : X \to Y$ such that for every $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $x \in X$ satisfies $0 < \|x - x_0\| < \delta$, we have*

$$\frac{\|f(x) - f(x_0) - L(x - x_0)\|}{\|x - x_0\|} < \epsilon$$

In other words, f is differentiable at x_0 if there exists a continuous linear transformation $L : X \to Y$ such that

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|}{\|x - x_0\|} = 0$$

It can be shown that if f is differentiable at x_0, then this continuous linear transformation L is unique and is called the Frechet derivative (or simply derivative) of f at x_0.

Here, we will denote the derivative of f at x_0 as $Df(x_0)$. If f is differentiable at every point $x \in X$, then f is said to be differentiable.

Definition 8 (Gradient)**.** *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ be differentiable at $a \in U$ with $Df(a) = A$. As A is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$, we know that there exists a unique vector $\alpha \in \mathbb{R}^n$ such that $Ah = \alpha \cdot h = \sum_{i=1}^n \alpha_i h_i$ if $h = (h_1, \ldots, h_n)$. This unique vector α is called the gradient of f at a. It is denoted by* $\mathrm{grad} f(a)$ *or $\nabla f(a)$. So, we have*

$$\mathrm{grad} f(a) = \big(Df(a)(e_1), \ldots, Df(a)(e_n)\big) \quad \textit{and} \quad Df(a)(h) = \nabla f(a) \cdot h$$

.

Definition 9 (Directional Derivative)**.** *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}$ be any function. Fix a vector $v \in \mathbb{R}^n$. We say that f has directional derivative at a in the direction of v if the limit $\lim_{t \to 0} \frac{f(a+tv)-f(a)}{t}$ exists. Note that since U is open, as observed earlier, $a + tv \in U$ for all t in a sufficiently small interval around 0. The limit, if exists, is denoted by $D_v f(a)$.*

Theorem 10. *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $a \in U$. Then $D_v f(a)$ exists for all $v \in U$ and we have*

$$D_v f(a) = Df(a)(v)$$

Remark. There exists functions $f : U \to \mathbb{R}$ such that $D_v f(a)$ exists for all $v \in U$ but f is not differentiable.

Considering $\{e_i : i = 1, 2, \ldots, n\}$ to be the standard basis of $\mathbb{R}^n$, the direction derivative $D_{e_i} f(a)$, if exists, is called the i-th partial derivative of f at a and is usually denoted $\frac{\partial f}{\partial x_i}(a)$ or at times by $D_i f(a)$.

Also, if $T : \mathbb{R}^n \to \mathbb{R}$ is a linear transformation, then $T(v) = (Te_i, \ldots, Te_n) \cdot (v_1, \ldots, v_n)$. Hence, we have

$$\begin{aligned} Df(a)(h) =& \big(Df(a)(e_1), \ldots, Df(a)(e_n)\big) \cdot (h_1, \ldots, h_n) \\ =& \big(D_{e_1} f(a), \ldots, D_{e_1} f(a)\big) \cdot (h_1, \ldots, h_n) \\ =& \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right) \cdot (h_1, \ldots, h_n) \end{aligned}$$

Thus,

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)$$

Theorem 11 (Mean Value Theorem)**.** *Let $f : U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable on U and $x, y \in U$. Assume that the line segment $[x, y] := \{(1-t)x + ty : t \in [0, 1]\} \subset U$. Then there exists $t_0 \in (0, 1)$ such that if we set $z := (1 - t_0)x + t_0 y$, then*

$$f(y) - f(x) = Df(z)(y - x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(z)(y_i - x_i)$$

Taylor's formula for functions $f : U \subset \mathbb{R}^n \to \mathbb{R}$.

We shall here restrict ourselves to C^2 functions. A function f is said to be C^2 if all partial derivatives of the form

$$\frac{\partial^{\alpha_1+\cdots+\alpha_n} f}{\partial x_1^{\alpha_1}\cdots\partial x_1^{\alpha_n}}, \qquad \text{where } \alpha_j \in \mathbb{Z} \text{ and } \alpha_1+\cdots+\alpha_n \leq 2$$

exist and are continuous. For, simplicity, lets assume $0, x \in U$, U is open and star shaped at 0 and we wish to find a taylor expansion of f at 0.

Now, consider the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(t) = f(tx)$. We will try to differentiate g in the usual calculus sense and compute its derivative.

$$\begin{aligned} g'(t) &= \lim_{h\to 0} \frac{g(t+h)-g(t)}{h} \\ &= \lim_{h\to 0} \frac{f(tx+hx)-f(tx)}{h} \\ &= D_x f(tx) \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(tx)x_i \end{aligned}$$

In particular,

$$g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(0)x_i$$

Let $g_i(t) = \dfrac{\partial f}{\partial x_i}(tx)$. Then, proceeding as above, we have

$$g_i'(t) = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(tx)x_j = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(tx)x_j$$

So,

$$g"(t) = \sum_{i=1}^{n} g_i'(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(tx)x_j x_i$$

Note that the above calculation shows that g is twice continuously differentiable and so we can apply Taylor's theorem for one-variable calculus to g. We get

$$g(t) = g(0) + g'(0)t + \frac{g''(\theta)}{2}t^2, \qquad \text{where } \theta \in (0,t)$$

Taking $t = 1$ and writing g in terms of f, we get

$$f(x) = f(0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(0)x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\theta x)x_j x_i, \qquad \text{where } \theta \in (0,1)$$

This can also be written as

$$f(x) = f(0) + \nabla f(0) \cdot x + x^T H_f(\theta x)x,$$

$$\text{where } H_f(x) = D^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix} \text{ and } \theta \in (0,1)$$

$H_f(x)$ is known as the Hessian of f at x.

2 Condition for Optimization of Functions

Optimization is the process of finding the maximum or minimum value of a function subject to certain constraints. In the context of functions defined on normed spaces, the study of optimization involves identifying critical points and understanding the behavior of functions near these points. This subsection explores the conditions required for optimizing functions in normed spaces, focusing on convex sets, convex functions, and the use of the Hessian matrix to classify critical points.

2.1 Convex Sets and Convex Functions

Definition 12 (Convex Sets)**.** *Let X be a normed space. A subset $C \subset X$ is said to be convex set if for every $x_1, x_2 \in C$, and all $\alpha \in (0,1)$, $(1-\alpha)\cdot x_1 + \alpha \cdot x_2 \in C$.*

Definition 13 (Convex Functions)**.** *Let X be a normed space and $C \subset X$ be convex. A map $f : C \to \mathbb{R}$ is said to be convex function if for every $x_1, x_2 \in C$, and all $\alpha \in (0,1)$, $f((1-\alpha)\cdot x_1 + \alpha \cdot x_2) \leq (1-\alpha)f(x_1) + \alpha f(x_2)$.*

When the inequality in the definition of convex functions is reversed, it is known as concave functions.

Theorem 14. *Let $f : \mathbb{R} \to \mathbb{R}$ be a twice continuously differentiable function. Then f is convex if and only if $f''(x) \geq 0$ for every $x \in X$.*

For the sake of generality, we will focus on the study of minimization problems here. The principles applicable to minimization extend to maximization, with the distinction that convexity for minimization corresponds to concavity for maximization.

2.2 Role of Vanishing Derivative in Optimization

Theorem 15. *Let X be a normed space, and let $f : X \to \mathbb{R}$ be a function that is differentiable at $x_* \in X$. If f has a minimum at x_*, then $Df(x_*) = \mathbf{0}$*

Proof. On the contrary, let us assume $Df(x_*) \neq \mathbf{0}$. Then there exists $h_0 \in X$ such that $df(x_*)(h_0) \neq 0$. Clearly $h_0 \neq 0$.

Now, $Df(x_*)$ is continuous, let $\epsilon > 0$, then there exists $\delta > 0$ such that whenever $x \in X$ satisfies $0 < \|x - x_*\| < \delta$, we have

$$\frac{|f(x) - f(x_*) - Df(x_*)(x - x_*)|}{\|x - x_*\|} < \epsilon$$

.

Thus whenever $0 < \|x - x_*\| < \delta$, we have

$$\begin{aligned} -\frac{Df(x_*)(x-x_*)}{\|x-x_*\|} &\le \frac{f(x) - f(x_*) - Df(x_*)(x-x_*)}{\|x-x_*\|} \\ &\le \frac{|f(x) - f(x_*) - Df(x_*)(x-x_*)|}{\|x-x_*\|} < \epsilon \end{aligned}$$

Consider $x := x_* - \left(\frac{\delta}{2} \cdot \frac{Df(x_*)(h_0)}{|Df(x_*)(h_0)|} \cdot \frac{1}{\|h_0\|}\right) \cdot h_0$.

Then $x \neq x_*$ and $\|x - x_*\| = \frac{\delta}{2} < \delta$.

So, we get

$$-\frac{Df(x_*)(x-x_*)}{\|x-x_*\|} = \frac{\frac{\delta}{2} \cdot \frac{(Df(x_*)(h_0))^2}{|Df(x_*)(h_0)|} \cdot \frac{1}{\|h_0\|}}{\frac{\delta}{2}} < \epsilon$$

. Thus, $|Df(x_*)(h_0)| < \epsilon\|h_0\|$. As $\epsilon > 0$ is arbitrary, hence $|Df(x_*)(h_0)| = 0$. So we obtain a contradiction. □

2.3 Sufficiency of Vanishing Derivative in Optimizing Convex Functions

Theorem 16. *Let C be an open convex set, $f : C \to \mathbb{R}$ be a C^2 function such that for all $x \in C$,*

$$H_f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix}$$

is positive semi-definite. Then f is convex.

Note. A matrix M is positive semi-definite if $v^T M v \ge 0$ for all $v \in \mathbb{R}^n$

Proof. Let $x, y \in C$ and $d = y - x$. Then by Taylor's formula, we have

$$f(y) = f(x) + \nabla f(x) \cdot d + \frac{1}{2} d^T H_f(x + \theta d) d, \qquad \text{for some } \theta \in (0, 1).$$

As $H_f(x)$ is positive semi-definite for all $x \in C$, so $f(y) \ge f(x) + \nabla f(x) \cdot d$

Now, let $u, v \in C$ and $\alpha \in (0,1)$. Fix $x = (1-\alpha)u + \alpha v$.

If we take $y = u$, we get $f(u) \geq f(x) + \alpha \nabla f(x) \cdot (u - v)$.

And if $y = v$, we get $f(v) \geq f(x) + (1-\alpha)\nabla f(x) \cdot (u - v)$.

Using the above two inequalities, we get

$$(1-\alpha)f(u) + \alpha f(v) \geq (1-\alpha)f(x) + \alpha f(x) = f(x) = f((1-\alpha)u + \alpha v)$$

So, f is convex. □

Theorem 17. *Let X be normed space and $f : X \to \mathbb{R}$ be convex and differentiable. If $x_* \in X$ is such that $Df(x_*) = \mathbf{0}$, then f has a minimum at x_*.*

Proof. Let us assume that x_* is not a minimizer of f, then there exists $x_0 \in X$ such that $f(x_0) < f(x_*)$. Define $\varphi : \mathbb{R} \to \mathbb{R}$ by $\varphi(t) = f(tx_0 + (1-t)x_*)$, $t \in \mathbb{R}$.

Now, for $\alpha \in (0,1)$ and $t_1, t_2 \in R$,

$$\begin{aligned}
\varphi((1-\alpha)t_1 + \alpha t_2) &= f\Big(\big((1-\alpha)t_1 + \alpha t_2\big)x_0 + \big(1 - (1-\alpha)t_1 - \alpha t_2\big)x_*\Big) \\
&= f\Big((1-\alpha)(t_1 x_0 + (1-t_1)x_*) + \alpha(t_2 x_0 + (1-t_2)x_*)\Big) \\
&\leq (1-\alpha)f\big(t_1 x_0 + (1-t_1)x_*\big) + \alpha f\big(t_2 x_0 + (1-t_2)x_*\big) \\
&= (1-\alpha)\varphi(t_1) + \alpha\varphi(t_2)
\end{aligned}$$

So, φ is convex. Using the concept of directional derivative, we have

$$\varphi'(0) = D_{x_0 - x_*}f(x_*) = Df(x_*)(x_0 - x_*) = \mathbf{0}(x_0 - x_*) = 0$$

Since $\varphi(1) = f(x_0) < f(x_*) = \varphi(0)$, it follows from the mean value theorem that there exists a $\theta \in (0,1)$ such that $\varphi'(\theta) = \dfrac{\varphi(1) - \varphi(0)}{1 - 0} < 0 = \varphi'(0)$.

But, φ is convex, by theorem 14, $\varphi''(t) \geq 0$, so φ' is increasing. So, this is a contradiction. Thus f has a minimum at x_* □

3 Optimization: Euler-Lagrange Equation as a Special Case

In this section, we will outline a crucial requisite for the minimization of a constrained function of the form

$$f(x) = \int_a^b F(\mathbf{x}(t), \mathbf{x}'(t), t)dt.$$

The necessary condition takes the shape of a differential equation, which the minimizer must adhere to. This equation is commonly referred to as the Euler-Lagrange equation.

First we will start with a lemma which we will use to prove the next theorem.

Lemma 18. *If* $\mathbf{k} \in C[a, b]$ *such that for all* $\mathbf{h} \in C^1[a, b]$ *with* $\mathbf{h}(a) = \mathbf{h}(b) = 0$*, we have*

$$\int_a^b \mathbf{k}(t)\mathbf{h}'(t)dt = 0,$$

then there exists a constant $c \in \mathbb{R}$ *such that* $\mathbf{k}(t) = c$, $\forall t \in [a, b]$.

Proof. Take $c := \dfrac{1}{b-a}\displaystyle\int_a^b \mathbf{k}(t)dt$.

Define $\mathbf{h}_0 : [a, b] \to \mathbb{R}$ by $\mathbf{h}_0(t) = \displaystyle\int_a^t (\mathbf{k}(\tau) - c)d\tau$.

Then $\mathbf{h}_0 \in C^1[a, b]$ and $\mathbf{h}_0(a) = \mathbf{h}_0(b) = 0$. Thus $\displaystyle\int_a^b \mathbf{k}(t)\mathbf{h}'_0(t)dt = 0$. Since $\mathbf{h}'_0(t) = \mathbf{k}(t) - c$, we get

$$\begin{aligned}\int_a^b (\mathbf{k}(t) - c)^2 dt = \int_a^b (\mathbf{k}(t) - c)\mathbf{h}'_0(t)dt &= \int_a^b \mathbf{k}(t)\mathbf{h}'_0(t)dt - c\int_a^b \mathbf{h}'_0(t)dt \\ &= 0 - c(\mathbf{h}_0(b) - \mathbf{h}_0(a)) = 0\end{aligned}$$

Thus $\mathbf{k}(t) - c = 0$ for all $t \in [a, b]$ □

3.1 Euler-Lagrange Equation: Fixed End Points

Theorem 19. *Suppose that*

- $S = \{\mathbf{x} \in C^1[a, b] : \mathbf{x}(a) = y_a, \mathbf{x}(b) = y_b\}$
- $F : \mathbb{R}^3 \to \mathbb{R}, (\xi, \eta, \tau) \overset{F}{\mapsto} F(\xi, \eta, \tau)$, *has continuous partial derivatives of order 2*
- $f : S \to \mathbb{R}$ *is given by* $f(\mathbf{x}) = \displaystyle\int_a^b F(\mathbf{x}(t), \mathbf{x}'(t), t)dt$, $\mathbf{x} \in S$

Then we have:

(i) *If x_* is a minimizer of f, then it satisfies the Euler-Lagrange equation:*

$$\frac{\partial F}{\partial \xi}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \eta}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)\right) = 0 \quad \textit{for all } t \in [a,b]$$

(ii) *If f is convex and $\mathbf{x}_* \in S$ satisfies the Euler-Lagrange equation, then $\mathbf{x}_*$ is a minimizer of f.*

Proof.

(i) The proof is long, so we divide it into multiple steps.

Step 1. The set S is not a vector space (unless $y_a = y_b = 0$). So, we introduce a new vector space $X = \{\mathbf{h} \in C^1[a,b] : \mathbf{h}(a) = \mathbf{h}(b) = 0\}$ with norm $\|\cdot\|_{1,\infty}$.

Note that $\mathbf{h} \in X \iff \mathbf{x}_* + \mathbf{h} \in S$

Define $\widetilde{f} : X \to \mathbb{R}$ given by $\widetilde{f}(\mathbf{h}) = f(\mathbf{x}_* + \mathbf{h}),\ \mathbf{h} \in X$.

Now, $\widetilde{f}(\mathbf{h}) = f(\mathbf{x}_* + \mathbf{h}) \geq f(\mathbf{x}_*) = \widetilde{f}(\mathbf{0})$, for all $\mathbf{h} \in X$. So, $\mathbf{0}$ is a minimizer of $\widetilde{f}$.

Step 2. Calculating $D\widetilde{f}(\mathbf{0})$.

By applying Taylor's formula on F, we have

$$\begin{aligned}F(\xi_0 + p, \eta_0 + q, \tau_0 + r) - F(\xi_0, \eta_0, \tau_0) \\ &= p\frac{\partial F}{\partial \xi}(\xi_0, \eta_0, \tau_0) + q\frac{\partial F}{\partial \eta}(\xi_0, \eta_0, \tau_0) + r\frac{\partial F}{\partial \tau}(\xi_0, \eta_0, \tau_0) \\ &\quad + \frac{1}{2}\begin{bmatrix} p & q & r \end{bmatrix} H_F(\xi_0 + \theta p, \eta_0 + \theta q, \tau_0 + \theta r)\begin{bmatrix} p \\ q \\ r \end{bmatrix}\end{aligned}$$

for some $\theta \in (0,1)$. Using this for each $t \in [a,b]$, we obtain

$$\begin{aligned}&\widetilde{f}(\mathbf{h}) - \widetilde{f}(\mathbf{0}) \\ &= \int_a^b \Big(F(\mathbf{x}_*(t) + \mathbf{h}(t), \mathbf{x'}_*(t) + \mathbf{h}'(t), t) - F(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)\Big)dt \\ &= \int_a^b \Bigg(\mathbf{h}(t)\frac{\partial F}{\partial \xi}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t) + \mathbf{h}'(t)\frac{\partial F}{\partial \eta}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t) \\ &\quad + \frac{1}{2}\begin{bmatrix} \mathbf{h}(t) & \mathbf{h}'(t) & 0 \end{bmatrix} H_F\big(\mathbf{x}_*(t) + \mathbf{\Theta}(t)\mathbf{h}(t), \mathbf{x'}_*(t) + \mathbf{\Theta}(t)\mathbf{h}'(t), t\big)\begin{bmatrix} \mathbf{h}(t) \\ \mathbf{h}'(t) \\ 0 \end{bmatrix}\Bigg)dt \\ &= \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big)dt + \int_a^b \frac{1}{2}\begin{bmatrix} \mathbf{h}(t) & \mathbf{h}'(t) & 0 \end{bmatrix} H_F\big(\mathbf{P}(t)\big)\begin{bmatrix} \mathbf{h}(t) \\ \mathbf{h}'(t) \\ 0 \end{bmatrix} dt\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Theta} :&[a,b] \to (0,1)\\ \mathbf{A}(t) =&\frac{\partial F}{\partial \xi}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)\\ \mathbf{B}(t) =&\frac{\partial F}{\partial \eta}(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)\\ \mathbf{P}(t) =&\big(\mathbf{x}_*(t) + \boldsymbol{\Theta}(t)\mathbf{h}(t), \mathbf{x'}_*(t) + \boldsymbol{\Theta}(t)\mathbf{h}'(t), t\big)\end{aligned}$$

and $H_F(\cdot)$ denotes the Hessian of F.

Define $L : X \to \mathbb{R}$ by

$$L(\mathbf{h}) = \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big)dt, \ \mathbf{h} \in X.$$

From Example 6, we know that L is a continuous linear transformation. We will now show that L is the derivative of $\widetilde{f}$ at $\mathbf{0}$. For $\mathbf{h} \in X$,

$$\begin{aligned}&|\widetilde{f}(\mathbf{h}) - \widetilde{f}(\mathbf{0} - L(\mathbf{h} - \mathbf{0})|\\ &= \left|\frac{1}{2}\int_a^b \Big((\mathbf{h}(t))^2\frac{\partial^2 F}{\partial \xi^2}(\mathbf{P}(t)) + 2\mathbf{h}(t)\mathbf{h}'(t)\frac{\partial^2 F}{\partial \xi \partial \eta}(\mathbf{P}(t)) + (\mathbf{h}'(t))^2\frac{\partial^2 F}{\partial \eta^2}(\mathbf{P}(t))\Big)dt\right|\\ &\le \frac{1}{2}\int_a^b \left(|\mathbf{h}(t)|^2\left|\frac{\partial^2 F}{\partial \xi^2}(\mathbf{P}(t))\right| + 2|\mathbf{h}(t)||\mathbf{h}'(t)|\left|\frac{\partial^2 F}{\partial \xi \partial \eta}(\mathbf{P}(t))\right| + |\mathbf{h}'(t)|^2\left|\frac{\partial^2 F}{\partial \eta^2}(\mathbf{P}(t))\right|\right)dt\\ &\le \frac{1}{2}\int_a^b \left(\|\mathbf{h}\|_\infty^2\left|\frac{\partial^2 F}{\partial \xi^2}(\mathbf{P}(t))\right| + 2\|\mathbf{h}\|_\infty\|\mathbf{h}'\|_\infty\left|\frac{\partial^2 F}{\partial \xi \partial \eta}(\mathbf{P}(t))\right| + \|\mathbf{h}'\|_\infty^2\left|\frac{\partial^2 F}{\partial \eta^2}(\mathbf{P}(t))\right|\right)dt\\ &\le \frac{1}{2}\int_a^b \|\mathbf{h}\|_{1,\infty}^2\left(\left|\frac{\partial^2 F}{\partial \xi^2}(\mathbf{P}(t))\right| + 2\left|\frac{\partial^2 F}{\partial \xi \partial \eta}(\mathbf{P}(t))\right| + \left|\frac{\partial^2 F}{\partial \eta^2}(\mathbf{P}(t))\right|\right)dt\\ &= M\|\mathbf{h}\|_{1,\infty}^2\end{aligned}$$

where

$$M = \frac{1}{2}\int_a^b \left(\left|\frac{\partial^2 F}{\partial \xi^2}(\mathbf{P}(t))\right| + 2\left|\frac{\partial^2 F}{\partial \xi \partial \eta}(\mathbf{P}(t))\right| + \left|\frac{\partial^2 F}{\partial \eta^2}(\mathbf{P}(t))\right|\right)dt$$

Note that for each $t \in [a,b]$, $\mathbf{P}(t) \in \mathbb{R}^3$ lies inside a ball with center $(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)$ and radius $\|\mathbf{h}\|_{1,\infty}$. Also, $\mathbf{x}_*, \mathbf{x'}_*$ are continuous, so $(\mathbf{x}_*(t), \mathbf{x'}_*(t), t)$, for all $t \in [a,b]$, lies inside a compact set in $\mathbb{R}^3$. Thus, there exists a compact set $K \in \mathbb{R}^3$ such that $\mathbf{P}(t) \in K$ for all $t \in [a,b]$. Since the second order partial derivatives of F are continuous, it follows that their absolute values are bounded on K. Hence, M is finite.

Let $\epsilon > 0$. Set $\delta = \dfrac{\epsilon}{M}$. Then, whenever $\mathbf{h} \in X$ satisfies $0 < \|\mathbf{h} - \mathbf{0}\|_{1,\infty} < \delta$, we have

$$\frac{|\widetilde{f}(\mathbf{h}) - \widetilde{f}(\mathbf{0}) - L(\mathbf{h} - \mathbf{0})|}{\|\mathbf{h}\|_{1,\infty}} \le \frac{M\|\mathbf{h}\|_{1,\infty}^2}{\|\mathbf{h}\|_{1,\infty}} = M\|\mathbf{h}\|_{1,\infty} < M\delta = \epsilon.$$

Thus, $D\widetilde{f}(\mathbf{0}) = L$, i.e.,

$$D\widetilde{f}(\mathbf{0})(\mathbf{h}) = L(\mathbf{h}) = \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big)dt, \ \mathbf{h} \in X$$

Step 3. Utilizing $D\widetilde{f}(\mathbf{0}) = \mathbf{0}$.

By using integration by parts,

$$\begin{aligned}\int_a^b \mathbf{A}(t)\mathbf{h}(t)dt =& \mathbf{h}(t)\int_a^t \mathbf{A}(\tau)d\tau\Big|_a^b - \int_a^b \Big(\mathbf{h}'(t)\int_a^t \mathbf{A}(\tau)d\tau\Big)dt \\ =& -\int_a^b \Big(\mathbf{h}'(t)\int_a^t \mathbf{A}(\tau)d\tau\Big)dt\end{aligned}$$

because $\mathbf{h}(a) = \mathbf{h}(b) = 0$. So,

$$L(\mathbf{h}) = \int_a^b \Big(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\Big)dt = \int_a^b \Big(-\int_a^t \mathbf{A}(\tau)d\tau + \mathbf{B}(t)\Big)\mathbf{h}'(t)dt,\ \mathbf{h} \in X$$

Now, as $\mathbf{0}$ is a minimizer for $\widetilde{f}$, by Theorem 15, $D\widetilde{f}(\mathbf{0}) = L = \mathbf{0}$. This means $L(\mathbf{h}) = 0$ for all $\mathbf{h} \in X$. Using Lemma 18, we obtain

$$-\int_a^t \mathbf{A}(\tau)d\tau + \mathbf{B}(t) = c,\ \forall\, t \in [a, b]$$

for some constant c. By differentiating this with respect to t, we obtain

$$\mathbf{A}(t) - \frac{d}{dt}\big(\mathbf{B}(t)\big) = 0,\ \forall\, t \in [a, b]$$

which is same as

$$\frac{\partial F}{\partial \xi}(\mathbf{x}_*(t), \mathbf{x}'_*(t), t) - \frac{d}{dt}\left(\frac{\partial F}{\partial \eta}(\mathbf{x}_*(t), \mathbf{x}'_*(t), t)\right) = 0,\ \forall\, t \in [a, b]$$

This completes the proof of (i).

(ii) Now, let f is convex and $\mathbf{x}_* \in S$ satisfies the Euler-Lagrange equation. Define X and f in the same manner as **Step 1**. By retracing the steps of **Step 3** above, we see that $D\widetilde{f}(\mathbf{0}) = \mathbf{0}$. Also, f is convex. So, if $\mathbf{h}_1, \mathbf{h}_2 \in X$ and $\alpha \in (0, 1)$, then

$$\begin{aligned}\widetilde{f}((1-\alpha)\mathbf{h}_1 + \alpha\mathbf{h}_2) =& f(\mathbf{x}_* + (1-\alpha)\mathbf{h}_1 + \alpha\mathbf{h}_2) \\ =& f((1-\alpha)(\mathbf{x}_* + \mathbf{h}_1) + \alpha(\mathbf{x}_* + \mathbf{h}_2)) \\ \leq& (1-\alpha)f(\mathbf{x}_* + \mathbf{h}_1) + \alpha f(\mathbf{x}_* + \mathbf{h}_2) \\ =& (1-\alpha)\widetilde{f}(\mathbf{h}_1) + \alpha\widetilde{f}(\mathbf{h}_2)\end{aligned}$$

Thus, $\widetilde{f}$ is convex and $D\widetilde{f}(\mathbf{0}) = \mathbf{0}$. So, by Theorem 17, $\widetilde{f}$ has a minimum at $\mathbf{0}$.

For any $\mathbf{x} \in S$, we have

$$f(\mathbf{x}) = f(\mathbf{x}_* + (\mathbf{x} - \mathbf{x}_*)) = \widetilde{f}(\mathbf{x} - \mathbf{x}_*) \geq \widetilde{f}(\mathbf{0}) = f(\mathbf{x}_*).$$

Hence, $\mathbf{x}_*$ is a minimizer of f.

This completes the proof.

□

References

1. Sasane, Amol. Optimization in function spaces: Dover Publications (2016).

2. Shifrin, Theodore. Multivariable mathematics: Linear Algebra, Multivariable Calculus, and Manifolds: Wiley (2005).

3. S. Kumaresan. A Conceptual Introduction to Several Variable Calculus: Mathematics Newsletter, University of Hyderabad (2003).