Fundamental Theorems of Optimization

A brief overview of continuity, differentiability and optimization in normed spaces, introduction to Euler-Lagrange Equations and solving Brachistochrone curve problem as an application

Sandipan Samanta

MathematiX Club
NISER

Table of Contents

Introduction to Normed Spaces

Definition 1 (Normed Space)

A **normed space** is a vector space equipped with a norm.
A norm on a vector space X over field $K(=\mathbb{R}$ or $\mathbb{C})$ is a function $\|\cdot\| : X \to \mathbb{R}$ that satisfies the following properties for all vectors $x, y \in X$ and scalar $\alpha \in K$:

- 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 on Normed Spaces

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 .

Examples of Normed Spaces

Euclidean Spaces

$(\mathbb{R}^n, \|\cdot\|_\infty)$ and $(\mathbb{R}^n, \|\cdot\|_p)$ for real number $p \geq 1$ are normed spaces.
For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $p \in [1, \infty)$,

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \text{ and } \|x\|_\infty = \sup \left\{ |x_i| : i = 1, 2, \ldots, n \right\}$$

Sequence Spaces

Denoted as ℓ^p space or ℓ^∞ space depending on the norm:
For a sequence $x = (x_n)$ and $p \in [1, \infty)$,

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}} \text{ and } \|x\|_\infty = \sup \left\{ |x_i| : i \in \mathbb{N} \right\}$$

Examples of Normed Spaces

Function Spaces

Denoted as L^p space or L^∞ space depending on the norm:

p-norm: $\|f\|_p = \left(\int_{i=1}^{n} |f|^p d\mu \right)^{\frac{1}{p}}$

sup-norm: $\|f\|_\infty = \sup \left\{ |f(x)| : x \in \Omega \right\}$

$C[a, b]$

$C[a, b] = \left\{ f : [a, b] \to \mathbb{R} : \text{f is continuous on } [a, b] \right\}$

p-norm: $\|f\|_p = \left(\int_{i=1}^{n} |f|^p dx \right)^{\frac{1}{p}}$

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

Examples of Normed Spaces

Note

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.

Continuity in Normed Spaces

Definition 2 (Continuous function)

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 $x_0 \in X$

Remark

The map $D : (C^1[0,1], \|\cdot\|_\infty) \to (C[0,1], \|\cdot\|_\infty)$ by

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

is nowhere continuous on $C^1[0,1]$.

But if we change the norm of $C^1[0,1]$ to $\|\cdot\|_{1,\infty}$
then the map $D : (C^1[0,1], \|\cdot\|_{1,\infty}) \to (C[0,1], \|\cdot\|_\infty)$ by

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

is continuous on $C^1[0,1]$.

Continuous Linear Transformation

Theorem 3

Let X and Y be normed spaces over $\mathbb{R}$, and $T : X \to Y$ be a linear transformation. Then are equivalent:

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

Example

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

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

L is a continuous linear transformation.

Differentiation in Normed Spaces

Definition 4 (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

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

If f is differentiable at x_0, then this continuous linear transformation L is unique and is called the Frechet derivative of f at x_0, denoted by $Df(x_0)$.

Definition 5 (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

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

.

Differentiation in Normed Spaces

Definition 6 (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 $D_v f(a)$.

Differentiation in Normed Spaces

Theorem 7

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.

However, if $D_v f(x)$ exists and is continuous for all $v \in U$ in a neighbourhood of $a \in U$, then f is diiferentiable at a.

Mean Value Theorem

Theorem 8

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

$f : U \subset \mathbb{R}^n \to \mathbb{R}$ be a C^2 function.
For, simplicity, lets assume $0, x \in U$, U is open and star shaped at 0, then the taylor expansion of f at 0 is given by

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

This can also be written as

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

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

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

Convex Sets and Convex Functions

Definition 9 (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 10 (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
$\alpha \in (0,1)$, $f((1-\alpha) \cdot x_1 + \alpha \cdot x_2) \leq (1-\alpha) f(x_1) + \alpha f(x_2)$.

Convex Sets and Convex Functions

Theorem 11

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 $x \in X$.

Theorem 12

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

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

is positive semi-definite. Then f is convex.

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

Role of Vanishing Derivative in Optimization

Theorem 13

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 $Df(x_*) = \mathbf{0}$

Theorem 14

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_*.

Optimization: Euler-Lagrange Equation as a Special Case

Lemma 15

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* $\mathbf{k}(t) = c, \ \forall t \in [a,b]$*.*

Optimization: Euler-Lagrange Equation as a Special Case

Proof.

Take $c := \dfrac{1}{b-a}\int_a^b \mathbf{k}(t)dt$.
Define $\mathbf{h}_0 : [a,b] \to \mathbb{R}$ by $\mathbf{h}_0(t) = \int_a^t (\mathbf{k}(\tau) - c)d\tau$.
Then $\mathbf{h}_0 \in C^1[a,b]$ and $\mathbf{h}_0(a) = \mathbf{h}_0(b) = 0$. Thus $\int_a^b \mathbf{k}(t)\mathbf{h'}_0(t)dt = 0$.
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 $t \in [a,b]$ □

Euler-Lagrange Equation

Theorem 16

Suppose that

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

Then we have:

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

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

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

Euler-Lagrange Equation

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 $\widetilde{f}$.

Euler-Lagrange Equation

Proof.

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 $\theta \in (0, 1)$.

Euler-Lagrange Equation

Proof.

Using this for each $t \in [a, b]$, we obtain

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

$$+ \int_a^b \frac{1}{2}\begin{bmatrix}\mathbf{h}(t) & \mathbf{h}'(t) & 0\end{bmatrix} H_F\left(\mathbf{P}(t)\right)\begin{bmatrix}\mathbf{h}(t)\\ \mathbf{h}'(t)\\ 0\end{bmatrix} dt$$

where $\mathbf{\Theta} : [a, b] \to (0, 1)$

$$\mathbf{A}(t) = \frac{\partial F}{\partial \xi}(\mathbf{x}_*(t), \mathbf{x}'_*(t), t)$$

$$\mathbf{B}(t) = \frac{\partial F}{\partial \eta}(\mathbf{x}_*(t), \mathbf{x}'_*(t), t)$$

$$\mathbf{P}(t) = \big(\mathbf{x}_*(t) + \mathbf{\Theta}(t)\mathbf{h}(t), \mathbf{x}'_*(t) + \mathbf{\Theta}(t)\mathbf{h}'(t), t\big)$$

Euler-Lagrange Equation

Proof.

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

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

We know that L is a continuous linear transformation. It can be shown that for $\mathbf{h} \in X$,

$$|\widetilde{f}(\mathbf{h}) - \widetilde{f}(\mathbf{0} - L(\mathbf{h} - \mathbf{0})| \leq M\|\mathbf{h}\|_{1,\infty}^2$$

where

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

Euler-Lagrange Equation

Proof.

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

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

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

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

Euler-Lagrange Equation

Proof.

Step 3. Utilizing $D\widetilde{f}(\mathbf{0}) = \mathbf{0}$.
By using integration by parts,

$$\int_a^b \mathbf{A}(t)\mathbf{h}(t)dt = \mathbf{h}(t)\int_a^t \mathbf{A}(\tau)d\tau\Big|_a^b - \int_a^b \left(\mathbf{h}'(t)\int_a^t \mathbf{A}(\tau)d\tau\right)dt$$

$$= -\int_a^b \left(\mathbf{h}'(t)\int_a^t \mathbf{A}(\tau)d\tau\right)dt$$

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

$$L(\mathbf{h}) = \int_a^b \left(\mathbf{A}(t)\mathbf{h}(t) + \mathbf{B}(t)\mathbf{h}'(t)\right)dt$$

$$= \int_a^b \left(-\int_a^t \mathbf{A}(\tau)d\tau + \mathbf{B}(t)\right)\mathbf{h}'(t)dt$$

Euler-Lagrange Equation

Proof.

Now, as $\mathbf{0}$ is a minimizer for $\widetilde{f}$, by Theorem 13, $D\widetilde{f}(\mathbf{0}) = L = \mathbf{0}$. 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).

Euler-Lagrange Equation

Proof.

(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 implies $\widetilde{f}$ is convex. So, by Theorem 14, $\widetilde{f}$ 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. □

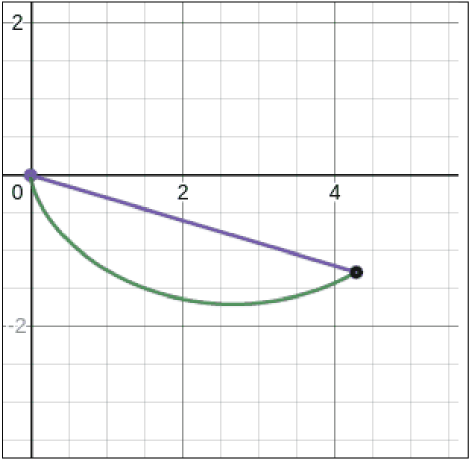
Brachistochrone Curve

The Problem Statement

Johann Bernoulli posed the problem of the brachistochrone to the readers of Acta Eruditorum in June, 1696. Bernoulli wrote the problem statement as:

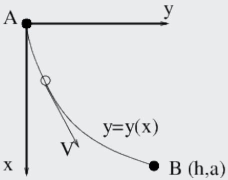
> *Given two points A and B in a vertical plane, what is the curve traced out by a point acted on only by gravity, which starts at A and reaches B in the shortest time.*

Brachistochrone Curve



▸ Open in Desmos

Solution



We will here use the illustrated coordinate system.
Let the particle starts from $A(0,0)$ and reaches $B(h,a)$, $h,a > 0$ and moves along the curve $y = y(x)$ such that $y(0) = 0$ and $y(h) = a$.
Using the conservation of energy, we have

$$\frac{1}{2}mv^2 = mgx \implies v = \sqrt{2gx}$$

Solution

Also, the arc length or distance along the curve s satisfies

$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

and

$$v = \frac{ds}{dt}$$

. So the total time required by the particle to descend along the curve $y = y(x)$ is given by

$$T(y) = \int_A^B dt = \int_A^B \frac{ds}{v} = \int_0^h \frac{\sqrt{1 + (y'(x))^2}}{\sqrt{2gx}} dx$$

Brachistochrone Curve

Solution

Now, to find the minimizer of T we will apply Theorem 16 i.e., $F(y(x), y'(x), x) = \frac{\sqrt{1+(y'(x))^2}}{\sqrt{2gx}}$ need to satisfy the Euler-Lagrange equation. We get

$$0 - \frac{d}{dx}\left(\frac{y'(x)}{\sqrt{2gx(1+(y'(x))^2}}\right) = 0$$

Integrating this, we get

$$\frac{y'(x)}{\sqrt{2gx(1+(y'(x))^2}} = c$$

where c is a constant. Rearranging this, we get

$$y'(x) = \sqrt{\frac{x}{\alpha - x}}, \text{where } \alpha = \frac{1}{2gc^2}$$

Brachistochrone Curve

Solution

It turns out the solution to this (in parametric form) is given by

$$x(\theta) = \frac{\alpha}{2}(1 - \cos\theta)$$

,

$$y(\theta) = \frac{\alpha}{2}(\theta - \sin\theta)$$

This is exactly the cycloid passing though $(0, 0)$ and (h, a)!

Thank You!