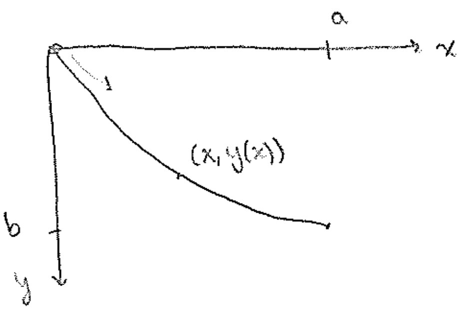
AMATH 456 - Calculus of Variations

2015 09 14

Optimization for unknown functions
Newton (1687) profile of minimum drag
Bernoulli (1696/7) Brachistochrone, shortest time curve



m = mass of the bead
initial point: (0,0)
final point: (a,b)
$a, b > 0$

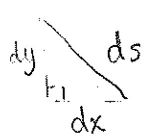
Conservation of energy

$$\frac{1}{2}mv^2 - mgy = \text{const.} = 0$$

start from rest, no friction

$$\rightarrow v^2 = 2gy \rightarrow v = \sqrt{2gy}$$

velocity: $v = \frac{ds}{dt}$, $ds = \sqrt{(dx)^2 + (dy)^2}$



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

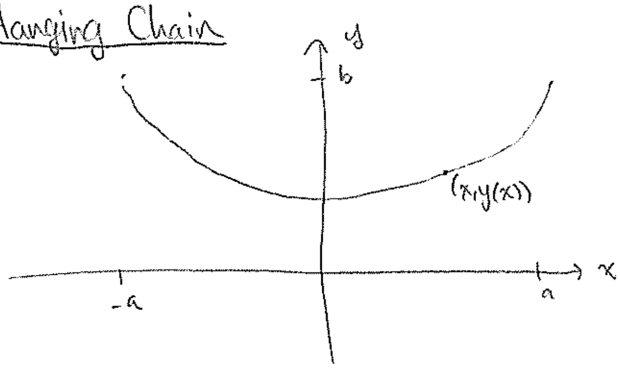
time of descent:

$$T[y] = \int dt = \int_0^a \frac{dt}{ds}\frac{ds}{dx}\,dx = \int_0^a \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{\sqrt{2gy}}\,dx$$

Problem: minimize $T[y] = \int_0^a \sqrt{\frac{1 + y'(x)^2}{2gy(x)}}\,dx$

subject to $y(0) = 0$, $y(a) = b$

Hanging Chain



ρ = density, mass per unit length

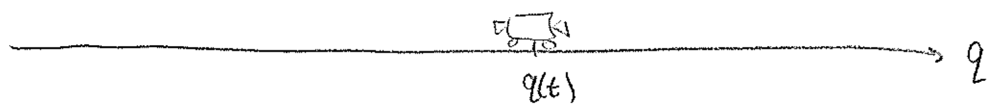
Minimize

$$E[y] = \int \rho g y \, ds = \int \rho g y(x) \sqrt{1 + y'(x)^2} \, dx$$

subject to $y(\pm a) = b, \quad a, b > 0$ $\qquad$ L = length of chain

$$\int ds = \int_{-a}^{a} \frac{ds}{dx} dx = \int_{-a}^{a} \sqrt{1 + y'(x)^2} \, dx = L. \qquad L \geq 2a$$

Rocket Car



m = mass of the car $\qquad$ assume frictionless, mass doesn't change

$m\ddot{q}(t) = u(t), \quad |u(t)| \leq M$

Let $p(t) = m\dot{q}(t)$,

$$\begin{cases} \dot{q} = \frac{1}{m} p \\ \dot{p} = u \end{cases}$$

$(q_0, p_0) = (q(0), p(0))$ is given

Target: $(q(t_1), p(t_1)) = (0, 0), \quad t_1 > 0$

(i) Minimize $\quad t_1 = \int_0^{t_1} dt$

or (ii) Minimize $\quad \int_0^{t_1} |u(t_1)| \, dt \qquad$ expenditure of fuel

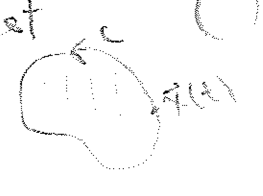
or (iii) Minimize $\quad \frac{1}{2m} \int_0^{t_1} p(t)^2 \, dt \qquad$ kinetic energy

→ $u(t) = \pm M \qquad$ bang-bang control

2015 M9 10

Isoperimetric Problem

Find the simplest closed curve C, $\vec{x}(t) = (x(t), y(t))$, $0 \leq t \leq 1$, $\vec{x}(0) = \vec{x}(1)$ of given length L which encloses the largest area A.



2015 09 16(

Answer: C is a circle of radius $\frac{L}{2\pi}$.
Hurwitz (1902) gave a proof based on Fourier series.

Green's Theorem.

$$\oint_C P\,dx + Q\,dy = \iint_D [Q_x - P_y]\,dA$$

$$\rightarrow \oint_C x\,dy - y\,dx = \iint_D [1-(-1)]\,dA = 2A$$

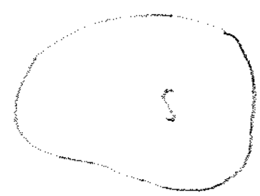
A = area of D

Minimize $\quad A[\vec{x}] = \frac{1}{2}\oint_C x\,dy - y\,dx = \frac{1}{2}\int_0^1 [x(t)y'(t) - y(t)x'(t)]\,dt$

Subject to $\quad \int_0^1 \sqrt{x'(t)^2 + y'(t)^2}\,dt = L, \quad \vec{x}(0) = \vec{x}(1)$

Optimization in Calculus

$$f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$



Ball of radius ε about x^*:

$$B_\varepsilon(x^*) = \{x \in \mathbb{R}^n ; \|x - x^*\| < \varepsilon\}$$

x^* is an interior point of S if $B_\varepsilon(x^*) \subseteq S$
x^* is a boundary point of S if $B_\varepsilon(x^*) \cap S \neq \emptyset$ & $B_\varepsilon(x^*) \cap S^c \neq \emptyset$ $\forall \varepsilon > 0$

$f \in C^1$ means $\frac{\partial f}{\partial x_i}$ is continuous on S $\forall i$
& f is continuous on S

S° is the set of interior points of S
∂S is the set of boundary points of S

Def] The variation of f in the direction of v $v \in \mathbb{R}^n$ at $x^* \in S^\circ$ is

$$\delta f(x^*; v) = \lim_{s \to 0} \frac{f(x^* + sv) - f(x^*)}{s}$$

This is the directional derivative if $\|v\| = 1$.

Theorem: If $f \in C^1$ then $\delta f(x^*; v) = \nabla f(x^*) \cdot v$.

Proof:
$$\delta f(x^*, v) = \frac{d}{ds} f(x^* + sv)\Big|_{s=0}$$
$$= \nabla f(x^* + sv) \cdot \frac{d}{ds}(x^* + sv)\Big|_{s=0}$$
$$= \nabla f(x^* + 0v) \cdot v$$
$$= \nabla f(x^*) \cdot v \qquad \blacksquare$$

Theorem: If $f \in C^1$ & f has a local max or min at $x^* \in S^\circ$ then

$$\frac{\partial f}{\partial x_i}(x^*) = 0, \quad 1 \le i \le n$$

(x^* is a critical point)

Proof: For $v \in \mathbb{R}^n$ consider
$$g(s) = f(x^* + sv), \quad g: \mathbb{R} \to \mathbb{R}$$
f has local max or min at x^* $\Rightarrow$ g has a local max or min at $s = 0$.
$$\Rightarrow g'(0) = 0 \Rightarrow \delta f(x^*; v) = \nabla f(x^*) \cdot v = 0$$
Let $v = e_i$ $\Rightarrow$
$$\frac{\partial f}{\partial x_i}(x^*) = 0, \quad 1 \le i \le n. \qquad \blacksquare$$

Consider
$$g(s) = f(x^* + sv), \quad 0 \le s \le 1$$
Taylor's Thm $\Rightarrow$
$$g(1) = g(0) + g'(0)(1-0) + \frac{1}{2} g''(c)(1-0)^2$$
for some c, $0 < c < 1$,
$$f(x^* + v) = f(x^*) + \nabla f(x^*) \cdot v + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^* + cv) v_i v_j$$

2015 09 16 (3)

$$g'(s) = \nabla f(x^*+sv)\cdot v = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^*+sv)\cdot v_i$$

$$g''(s) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*+sv)\, v_i v_j$$

2015 09 18

$$f(x^*+v) = f(x^*) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x^*)v_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*+cv)\, v_i v_j$$

$$\overset{(*)}{=} f(x^*) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*+cv)\, v_i v_j$$

Let

$$\nabla^2 f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^{n},$$

the Hessian matrix.

Theorem: Suppose $f \in C^{2?}$, $\nabla f(x^*) = 0$, $x^* \in S^\circ$.
(1) If $\nabla^2 f(x^*)$ is positive definite then f has a local min at x^*.
(2) If $\nabla^2 f(x^*)$ is negative definite then f has a local max at x^*.
(3) If $\nabla^2 f(x^*)$ indefinite then f has neither max nor min at x^*.

$n=1$: $\nabla^2 f(x^*) = f''(x^*)$

Theorem: Let $f \in C^2$, $\nabla f(x^*) = 0$, $x^* \in S^\circ$.
(1) If f has a local min at x^* then $\nabla^2 f(x^*)$ is positive semi-definite
(2) If f has a local max at x^* then $\nabla^2 f(x^*)$ is negative semi-definite.

Proof (of (1)): $g(s) = f(x^*+sv)$

$\longrightarrow g'(0) = 0$ & $g''(0) \geq 0$

$$\Rightarrow \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*)v_i v_j \geq 0$$

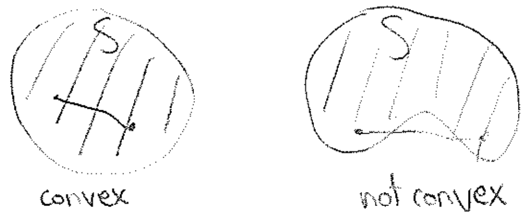
$$\Rightarrow v^T \nabla^2 f(x^*) v \geq 0$$

Convex Functions

$S \subseteq \mathbb{R}^n$ is <u>convex</u> if $\forall x, y \in S$, $x \neq y$, $\alpha x + (1-\alpha) y \in S$ for $0 < \alpha < 1$.

$\alpha x + (1-\alpha) y = y + \alpha(x - y)$

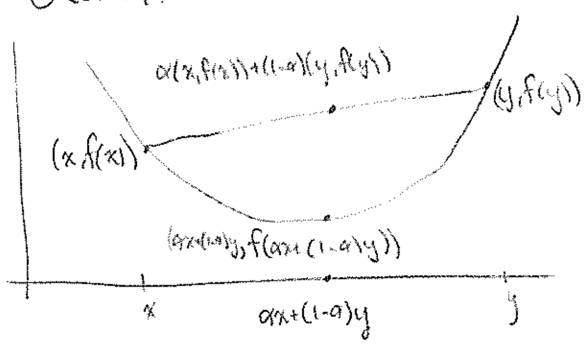
For $0 \le \alpha \le 1$, gives the line segment between x and y.



Let $f: S \subseteq \mathbb{R}^n \to \mathbb{R}$, S is convex, f is convex if $\forall x, y \in S$, $x \neq y$,

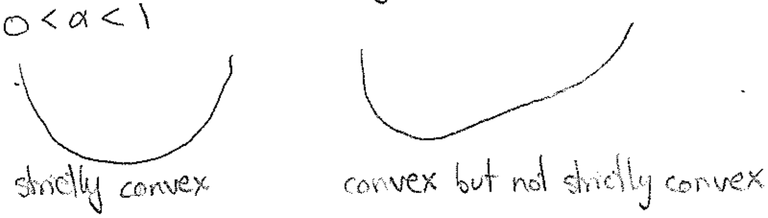
$$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y)$$

for all α, $0 < \alpha < 1$.



Any chord lies above the graph

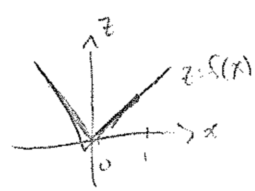
f is strictly convex if $\forall x, y \in S$, $x \neq y$, $f(\alpha x + (1-\alpha) y) < \alpha f(x) + (1-\alpha) f(y)$
$\forall \alpha$, $0 < \alpha < 1$



The epigraph of f is epi(f):

$$\{(x, z) ;\ z \ge f(x),\ x \in S\}$$

f is convex if and only if epi(f) is convex



Ex. $f(x) = |x|$, $f: \mathbb{R} \to \mathbb{R}$

f is convex. for $0 < \alpha < 1$

$f(\alpha x + (1-\alpha) y) = |\alpha x + (1-\alpha) y| \le |\alpha x| + |(1-\alpha) y| = \alpha |x| + (1-\alpha)|y| = \alpha f(x) + (1-\alpha) f(y)$

$x = 0, y = 1, \alpha = \frac{1}{2}$: $\alpha x + (1-\alpha) y = \frac{1}{2}x + \frac{1}{2}y = \frac{1}{2}$ $f(\frac{1}{2}) = \frac{1}{2}$ not strictly convex

2015 09 21

$f \in C^0$

$\forall x, y \in S,\ x \neq y,\ 0 < \alpha < 1$

f is convex: $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$

f is strictly convex: $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$

$f \in C^1$

f is convex if and only if $\forall x, x+v \in S, v \neq 0$

$$f(x+v) \geq f(x) + \delta f(x;v)$$

f is strictly convex if and only if $\forall x, x+v \in S, v \neq 0$

$$f(x+v) > f(x) + \delta f(x;v)$$

(the right hand side is linear approximation to f at x

$\nabla^2 f$ = Hessian matrix

$f \in C^2$

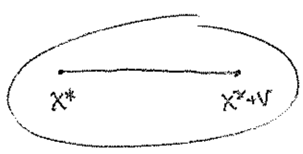
f is convex if and only if $\nabla^2 f$ is positive semi-definite

f is strictly convex if $\nabla^2 f$ is positive definite

Theorem: If $f \in C^1$ is strictly convex and $\nabla f(x^*) = 0$, $x^* \in S$, then x^* is the unique minimum point of f.

Proof: If $x^* + v \in S$, $v \neq 0$

$$f(x^*+v) > f(x^*) + \nabla f(x^*) \cdot v = f(x^*).$$



$f \in C^2$: let $g(s) = f(x+sv),\quad 0 \leq s \leq 1$

$$g'(s) = \nabla f(x+sv) \cdot v$$

$$g''(s) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j} v_i v_j = v^T \nabla^2 f(x+sv) v$$

By Taylor's theorem $\Longrightarrow$

$$g(1) = g(0) + g'(0)(1-0) + \tfrac{1}{2}g''(c)(1-0)^2$$
$$= g(0) + g'(0) + \tfrac{1}{2}g''(c) \quad \text{for some } c,\ 0<c<1$$
$$\geq g(0) + g'(0)$$
$$= f(x) + \nabla f(x)\cdot v$$
$$= f(x) + \delta f(x;v)$$

If $\nabla^2 f$ is pos. def. then $f(x) > f(x) + \delta f(x;v)$

Ex. $f:\mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = x^2 + 2xy + 2y^2 + 2x + 5$

$f_x = 2x + 2y + 2$ $\quad f_{xx} = 2,\ f_{xy} = 2,\ f_{yy} = 4$

$f_y = 2x + 4y$ $\quad f \in C^2$

$\nabla^2 f = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$ $\quad a_{11} = 2 > 0,\ a_{11}a_{22} - a_{12}^2 = 2\cdot 4 - 2^2 = 4 > 0$

$\therefore f$ is strictly convex on $\mathbb{R}^2$

Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$:

A is positive definite if and only if $a_{11} > 0$ and $a_{11}a_{22} - a_{12}^2 > 0$

A is positive semidefinite if and only if $a_{11} \geq 0$, $a_{11}a_{22} - a_{12}^2 \geq 0$, $a_{22} \geq 0$

Crit pt: $\begin{cases} f_x = 0 \to x + y + 1 = 0 \\ f_y = 0 \to x + 2y = 0 \to x = -2y \to -y + 1 = 0 \to y = 1,\ x = -2 \end{cases}$

$\therefore f(x,y) > f(-2,1) = 3$ for $(x,y) \neq (-2,1)$

$f \in C^1$ $\quad x, x+v \in S$, $v \neq 0$

let $g(s) = f(x+sv)$, $0 \leq s \leq 1$

f is convex $\Rightarrow g$ is convex $\Rightarrow g'(s)$ is increasing $\Rightarrow f(x+v) \geq f(x) + \delta f(x;v)$
$\Rightarrow f$ is convex

For $s_1 < s_2$,

$$\frac{g(s_1+\varepsilon_1) - g(s_1)}{\varepsilon_1} \leq \frac{g(s_2) - g(s_1)}{\varepsilon_2 - \varepsilon_1} \leq \frac{g(s_2) - g(s_2 - \varepsilon_2)}{\varepsilon_2}$$

let $\varepsilon_1 \to 0^+$, $\varepsilon_2 \to 0^+$: $g'(s_1) \leq g'(s_2)$

By Taylor's theorem:

$$g(1) = g(0) + g'(c) \quad \text{some } c,\ 0<c<1$$
$$\geq g(0) + g'(0)$$
$$\Rightarrow f(x+v) \geq f(x) + \delta f(x;v)$$

$x \neq y$, $x,y \in S$ Let $z = \alpha x + (1-\alpha)y$

$f(x) = f(z + (x-z)) \geq f(z) + \nabla f(z)\cdot(x-z)$

$f(y) = f(z + (y-z)) \geq f(z) + \nabla f(z)(y-z)$

$\Rightarrow \alpha f(x) + (1-\alpha) f(y) \geq \alpha f(z) + (1-\alpha) f(z) + \nabla f(z)(\alpha x + (1-\alpha)y - z)$

cont'd

2015 09 23

Theorem: Let $f: S \subseteq \mathbb{R}^n \to \mathbb{R}$, S convex, $x, x+v \in S$, $v \neq 0$, $g(s) = f(x+sv)$, $0 \le s \le 1$

repeats

(1) For $f \in C^1$, f is convex $\iff g'$ is increasing $\forall x, v$
$\iff f(x+v) \ge f(x) + \delta f(x;v)$ $\forall x, v$

(2) For $f \in C^2$, f is convex $\iff \nabla^2 f(x)$ is positive semidefinite $\forall x$

(3) For $f \in C^1$, f is strictly convex $\iff g'$ is strictly increasing
$\iff f(x+v) > f(x) + \delta f(x;v)$

(4) For $f \in C^2$, $\nabla^2 f(x)$ is positive definite $\forall x \Rightarrow f$ is strictly convex

Theorem: For $f: I \subseteq \mathbb{R} \to \mathbb{R}$, I an interval, $f \in C^2$

(1) f is convex if and only if f' is increasing if and only if $f'' \ge 0$

(2) f is strictly convex if and only if f' is strictly increasing.
& $f''(x) > 0$ except for a finite number of points $\Rightarrow f'$ strictly increasing & hence f is strictly convex

Ex (1) $f(x_1) = x_1^2$, $f: \mathbb{R} \to \mathbb{R}$
$f'(x_1) = 2x_1$, $f''(x_1) = 2 > 0$ $\forall x_1$ $\therefore f$ is strictly convex

(2) $f(x_1) = x_1^4$, $f: \mathbb{R} \to \mathbb{R}$
$f'(x_1) = 4x_1^3$, $f''(x_1) = 12x_1^2 > 0$ for $x_1 \neq 0$ $\therefore f$ is strictly convex

(3) $f(x_1) = \frac{1}{x_1^2}$, $x_1 \neq 0$
$f'(x_1) = -2/x_1^3$, $f''(x_1) = 6/x_1^4 > 0$ $\quad f$ is strictly convex on $(0, \infty)$
" $(-\infty, 0)$

Prop: Let $f(x_1)$ & $g(x_2)$ be strictly convex on $\mathbb{R}$ then $F(x_1, x_2) = f(x_1) + g(x_2)$ is strictly convex on $\mathbb{R}^2$.

Proof: Let $x = (x_1, x_2)$, $y = (y_1, y_2)$, $x \neq y$, $0 < \alpha < 1$

$F(\alpha x + (1-\alpha)y) = f(\alpha x_1 + (1-\alpha)y_1) + g(\alpha x_2 + (1-\alpha)y_2)$

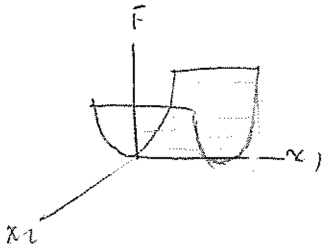
$x \neq y \Rightarrow$ $x_1 \neq y_1$ or $x_2 \neq y_2$ $\longrightarrow < \alpha f(x_1) + (1-\alpha)f(y_1) + \alpha g(x_2) + (1-\alpha)g(x_2)$
$= \alpha F(x_1, x_2) + (1-\alpha)F(y_1, y_2)$

$\therefore F$ is strictly convex ☒

Ex (1) $F(x_1, x_2) = x_1^4 + x_2^2$
$f(x_1) = x_1^4$ is strictly convex on $\mathbb{R}$
$g(x_2) = x_2^2$ is " "

Prop $\Rightarrow F(x_1, x_2)$ is strictly convex on $\mathbb{R}^2$

(2) $F(x_1, x_2) = x_1^2$ on $\mathbb{R}^2$



$\nabla^2 F = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ $\quad \nabla^2 F$ is pos. semi-def. $\forall x$

$\Rightarrow F$ is convex

However, F is not strictly convex on $\mathbb{R}^2$:

Let $x = (0,0)$, $y = (0,2)$, $\alpha = \frac{1}{2}$

$F(0,1) = 0 = \frac{1}{2}F(0,0) + \frac{1}{2}F(0,2)$

shows F is not strictly convex

<u>Functionals</u>

Let $\mathcal{Y}$ is a linear space (vector space) of functions.

If $y_1, y_2 \in \mathcal{Y}$ then $y_1 + y_2$, $cy_1 \in \mathcal{Y}$, $\forall c \in \mathbb{R}$

$C^m[a,b] = \{ f : [a,b] \to \mathbb{R} ;\ f, f', \ldots, f^{(m)}$ are continuous on $[a,b] \}$, $m \geq 0$

$(C^m[a,b])^d = \underbrace{C^m[a,b] \times \cdots \times C^m[a,b]}_{d \text{ factors}}$

$= \{ (f_1, \ldots, f_d) ;\ f_j \in C^m[a,b], 1 \leq j \leq d \}$

A functional is a real-valued function defined on $\mathcal{D}$ a subset of a linear space of functions $\mathcal{Y}$.

Ex 1 $\quad J(y) = \int_a^b \frac{1}{y(x)}\,dx \qquad \mathcal{Y} = C^0[a,b] = C[a,b]$

$\mathcal{D} = \{ y \in \mathcal{Y} ;\ y(x) > 0 \text{ on } [a,b] \}$

$\mathcal{D} \subset \mathcal{Y}$

Ex 2 $\quad J(y) = \int_a^b \sqrt{1 + y'(x)^2}\,dx \qquad J(y)$ is the length of the curve $y = y(x)$

$\mathcal{D} = \mathcal{Y} = C'[a,b]$

2015 09 25

Ex (2) $\quad J(\vec{x}) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\,dt$

$\forall \vec{x} = (x, y) \in C'[a,b] \times C'[a,b] = (C'[a,b])^2$