

# 最適化数学 第 12 回

## [今回の項目]

- ① 凸汎関数
- ② オイラー-ラグランジュ方程式

# 凸汎関数

数ベクトルの最小化問題を考えるとき、凸関数が重要な役割を果たした。 $\mathbb{R}^n$  の凸関数は次の性質を持っている；

$f$  が凸関数である

$$\Leftrightarrow \text{任意の } u, v \in \mathbb{R}^n \text{ について } f(v) \geq f(u) + \nabla f(u)(v - u)$$

$$\Leftrightarrow \text{任意の } u \text{ についてヘッセ行列 } \nabla^2 f(u) \text{ が半正定値}$$

ここでは、汎関数の凸性について考える。

# 凸汎関数の定義

## Definition

$F$  を汎関数とする. 任意の関数  $y(x), v(x)$  に対して

$$F(y + v) \geq F(y) + DF(y)(v)$$

が成り立つとき,  $F$  を凸汎関数と呼ぶ.

## Example

$F(y) = \int_0^1 y(x)^2 dx$  は凸汎関数である. 実際,

$$\begin{aligned} F(y + v) - F(y) &= 2 \int_0^1 v(x)y(x) dx + \int_0^1 v(x)^2 dx \\ &\geq 2 \int_0^1 v(x)y(x) dx = DF(y)(v) \text{ となる.} \end{aligned}$$

# 汎関数が凸になる条件

次に、一般的な凸性の判定法を紹介する。まず言葉を用意する。

## Definition

3変数関数  $f(x, y, z)$  に対して、 $x$  を定数と見なし、 $(y, z)$  の関数を

$$g(y, z) = f(x, y, z)$$

とおく。すべての  $x$  に対して  $g(y, z)$  が凸関数であるとき、 $f(x, y, z)$  は第 2, 第 3 変数に関して凸であるという。

## [命題]

3 変数関数  $f(x, y, z)$  に対して,

$f(x, y, z)$  が第 2, 第 3 変数に関して凸

$\iff$  任意の  $x, y, z$  に対して  $\begin{bmatrix} f_{yy}(x, y, z) & f_{yz}(x, y, z) \\ f_{zy}(x, y, z) & f_{zz}(x, y, z) \end{bmatrix}$  が半正定値

## [定理]

汎関数を  $F(y) = \int_a^b f(x, y(x), y'(x)) dx$  とする. 任意の  $x \in [a, b]$  に対して, 被積分関数  $f(x, y, z)$  が第 2, 第 3 変数に関して凸ならば, 汎関数  $F$  も凸である.

# 凸汎関数の例

## Example

$$\textcircled{1} \quad F(y) = \int_a^b \{x + y(x)^2 + y'(x)^2\} dx$$

被積分関数  $f(x, y, z) = x + y^2 + z^2$  は凸関数なので  $F$  は凸汎関数である。

$$\textcircled{2} \quad F(y) = \int_a^b \{-x^2 + y(x)^2 + y'(x)^2\} dx$$

被積分関数  $f(x, y, z) = -x^2 + y^2 + z^2$  は  $(x, y, z)$  に関しては凸ではないが,  $x$  を定数とみなすと, 第 2, 第 3 変数に関しては凸である。したがって,  $F$  は凸汎関数である。

# 変分問題の最適性条件

$$\text{最小化 } F(y) := \int_a^b f(x, y(x), y'(x)) dx$$

$$\text{制約 } y(a) = A, y(b) = B$$

この問題には最適解の候補となる関数  $y$  に対して、 $y(a) = A, y(b) = B$  という制約がついているので、固定端問題と呼ばれる。

以下、議論がやさしい順に

- ① 凸汎関数の大域最適解の十分条件
- ② 汎関数の局所最適解の必要条件

という順番で説明する。

# 方向微分の第 2 公式

## Lemma (方向微分の第 2 公式)

汎関数  $F(y) = \int_a^b f(x, y(x), y'(x)) dx$  に対して、方向微分は以下となる；

$$DF(y)(v) = \int_a^b \left[ f_y[y(x)] - \frac{d}{dx} \{ f_z[y(x)] \} \right] v(x) dx + \left[ f_z[y(x)] v(x) \right]_a^b.$$

【証明】 .

方向微分の公式に部分積分を用いると

$$\begin{aligned} DF(y)(v) &= \int_a^b \{ f_y[y(x)]v(x) + f_z[y(x)]v'(x) \} dx \\ &= \int_a^b f_y[y(x)]v(x) dx + \left[ f_z[y(x)]v(x) \right]_a^b - \int_a^b \frac{d}{dx} \{ f_z[y(x)] \} v(x) dx \\ &= \int_a^b \left[ f_y[y(x)] - \frac{d}{dx} \{ f_z[y(x)] \} \right] v(x) dx + \left[ f_z[y(x)]v(x) \right]_a^b. \end{aligned}$$



# 凸汎関数に対する最適性十分条件

[定理]

$$\text{最小化 } F(y) := \int_a^b f(x, y(x), y'(x)) dx$$

$$\text{制約 } y(a) = A, y(b) = B$$

(\*)

において、目的汎関数  $F$  が凸汎関数であるとする。関数  $\bar{y}(x)$  が

$$\begin{cases} \frac{d}{dx} f_z[y(x)] = f_y[y(x)] \\ y(a) = A, y(b) = B \end{cases}$$

の解ならば、 $\bar{y}(x)$  は問題 (\*) の大域最小解である。

# 定理の証明

$\bar{y}(x)$  が問題 (3) の大域最小解であることを示すには,

$$F(y) \geq F(\bar{y}) \quad (y(a) = A, y(b) = B \text{ を満たすすべての関数 } y(x))$$

を示せばよい. これは,  $\bar{y}(a) = A, \bar{y}(b) = B$  なので,  
 $v(x) = y(x) - \bar{y}(x)$  とおくことにより

$$F(\bar{y} + v) \geq F(\bar{y}) \quad (v(a) = v(b) = 0 \text{ を満たすすべての関数 } v(x))$$

と同値である. 以下で後者を示す.

## 定理の証明の続き

関数  $\bar{y}(x)$  を (\*)  $\begin{cases} \frac{d}{dx} f_z[y(x)] = f_y[y(x)] \\ y(a) = A, y(b) = B \end{cases}$  の解とし,  $v(x)$  を

$v(a) = v(b) = 0$  を満たす任意の関数とする. ここで, 方向微分の第 2 公式を用いると

$$\begin{aligned} DF(\bar{y})(v) &= \int_a^b \left[ f_y[\bar{y}(x)] - \frac{d}{dx} \{ f_z[\bar{y}(x)] \} \right] v(x) dx + \left[ f_z[\bar{y}(x)] v(x) \right]_a^b = 0 \end{aligned}$$

となる. いま, 目的関数  $F$  が凸なので, 定義より,

$$F(y + v) \geq F(\bar{y}) + DF(\bar{y})(v) = F(\bar{y})$$

が成り立つ. よって  $\bar{y}$  は  $(P)$  の大域最小解になる.

# オイラー-ラグランジュ方程式と停留関数

## Definition

$$\begin{cases} \frac{d}{dx} f_z[y(x)] = f_y[y(x)] \\ y(a) = A, y(b) = B \end{cases}$$

を満たす関数  $y(x)$  を, 停留関数 と呼ぶ. また, 上記の式

$$\frac{d}{dx} f_z[y(x)] = f_y[y(x)]$$

を オイラー-ラグランジュ方程式 と呼ぶ.

# 一般の汎関数に対する最適性必要条件

次に、一般の汎関数に対して局所最適解の必要条件を挙げる。

[定理]

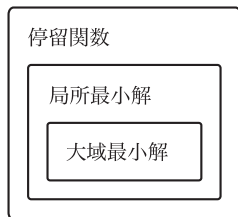
$$\begin{aligned} \text{最小化} \quad & F(y) := \int_a^b f(x, y(x), y'(x)) dx \\ \text{制約} \quad & y(a) = A, \quad y(b) = B \end{aligned}$$

に対して、 $\bar{y}(x)$  を局所最小解とする。このとき  $\bar{y}(x)$  は、以下を満たす：

$$(*) \quad \begin{cases} \frac{d}{dx} f_z[\bar{y}(x)] = f_y[\bar{y}(x)] \\ \bar{y}(a) = A, \quad \bar{y}(b) = B. \end{cases}$$

# 停留関数と最小解の関係

変分問題においても停留関数と最小解は下図のようになり、停留関数であっても最小解でない関数が存在する。



しかし、目的汎関数が凸のときはすべて一致する。

[系]

変分問題 (4) において、目的汎関数  $F$  を凸汎関数とする。すると、局所最小解はすべて大域最小解になり、

$$\bar{y}(x) \text{ が大域最小解} \iff \begin{cases} \frac{d}{dx} \{f_z[\bar{y}(x)]\} = f_y[\bar{y}(x)] \\ \bar{y}(a) = A, \bar{y}(b) = B \end{cases}$$

が成り立つ。

# 解法例

$$\text{最小化 } F(y) = \int_0^1 \{y(x) + y'(x)^2\} dx$$
$$y(0) = 1, y(1) = 2$$

板書

# 練習問題

## [練習問題]

変分問題の停留関数を求めよ.

$$(1) \text{ 最小化 } F(y) = \int_0^1 \{4e^x y(x) + y'(x)^2\} dx$$
$$y(0) = 0, \quad y(1) = 0$$

$$(2) \text{ 最小化 } F(y) := \int_0^1 \{(y'(x) - x)^2 + 2xy(x)\} dx$$
$$y(0) = 0, \quad y(1) = 5/3$$