

9.3 凸汎関数

数ベクトルの最小化問題を考えるときに凸関数が重要な役割を果たした。凸関数は次の性質を持つ;

$$\begin{aligned} f \text{ が凸関数である} &\Leftrightarrow \text{任意の } u, v \in \mathbb{R}^n \text{ について } f(v) \geq f(u) + \nabla f(u)(v - u) \\ &\Leftrightarrow \text{任意の } u \text{ についてヘッセ行列 } \nabla^2 f(u) \text{ が半正定値} \end{aligned}$$

以下では、汎関数についても凸性を考える。

定義. 汎関数 F が、任意の関数 y, \tilde{y} に対して

$$F(\tilde{y}) \geq F(y) + DF(y)(\tilde{y} - y)$$

を満たす時、汎関数 F を凸汎関数と呼ぶ。ここで、 $DF(y)$ は F の y におけるガトー微分を表す。($F(y+v) \geq F(y) + DF(y)(v)$) とも書ける。

上の不等式が集合 C 内の関数 y, \tilde{y} に対してのみ成り立つ場合は、 F は C 上で凸であるという。

命題 21. 汎関数 F を

$$F(y) = \int_a^b f[y(x)]dx = \int_a^b f(x, y(x), y'(x))dx$$

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 も凸になる。

補足. 「被積分関数 $f(x, y, z)$ が第 2, 第 3 変数に関して凸」とは、偏微分のように、 x を定数とみなして (y, z) の関数 $f(x, \cdot, \cdot)$ が凸、という意味である。

証明. $f(x, y, z)$ は第 2, 第 3 変数で凸なので、任意の $(v_1, v_2) \in \mathbb{R}^n$ に対して、

$$f(x, y + v_1, z + v_2) \geq f(x, y, z) + \langle (f_y(x, y, z), f_z(x, y, z)), (v_1, v_2) \rangle$$

が成り立つ。この式に $y = y(x), z = y'(x), v_1 = v(x), v_2 = v'(x)$ とすると、

$$\begin{aligned} &f(x, y(x) + v(x), y'(x) + v'(x)) \\ &\geq f(x, y(x), y'(x)) + \langle (f_y(x, y(x), y'(x)), f_z(x, y(x), y'(x))), (v(x), v'(x)) \rangle \end{aligned}$$

を得る。 $f[y(x)] = f(x, y(x), y'(x)), f_y[y(x)] = f_y(x, y(x), y'(x))$ などという略記を使うと、これは

$$f[y(x) + v(x)] \geq f[y(x)] + f_y[y(x)]v(x) + f_z[y(x)]v'(x)$$

となる。両辺を積分すると、

$$\int_a^b f[y(x) + v(x)]dx \geq \int_a^b f[y(x)]dx + \int_a^b \{f_y[y(x)]v(x) + f_z[y(x)]v'(x)\} dx$$

を得て、これはガトー微分の公式より $F(y+v) \geq F(y) + DF(y)(v)$ を表す。よって汎関数 F は凸になる。□

例 23. 凸汎関数の例

(i). $F(y) = \int_a^b \{x + y(x)^2 + y'(x)^2\} dx$

被積分関数は $f(x, y, z) = x + y^2 + z^2$.

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

被積分関数は $f(x, y, z) = -x^2 + y^2 + z^2$. x 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.

10 変分問題の最適性条件

以下の変分問題が基本的である:

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

数ベクトル上の最適化問題で得たように変分問題に対しても最適性条件を用いて最適解を絞り込むことができる.

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. しかし, 最適性条件を考えると, 結果的に制約なしの問題のような最適性条件が出てくる.

10.1 凸汎関数に対する最適性十分条件

目的関数の汎関数が凸の場合最適性十分条件が求まる.

定理 22. 最小化問題 (P) において, 目的関数 F が $y(a) = A, y(b) = B$ を満たす関数の集合上で凸とする. 関数 \bar{y} が

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

の解ならば, \bar{y} は (P) の大域最小解である.

補足. 定理内の式

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

を **オイラー・ラグランジュの方程式** と呼ぶ.

証明. 関数 \bar{y} を

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

の解とし, v を $v(a) = v(b) = 0$ を満たす任意の関数とする. ここで

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

という関係が成り立つことに注意する. すると,

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

となる. いま, 任意の $y(a) = A, y(b) = B$ を満たす y に対して, $(\bar{y} - y)(a) = (\bar{y} - y)(b) = \bar{y}(b) - \bar{y}(b) = 0$ となるので, $DF(\bar{y})(y - \bar{y}) = 0$ が成り立つ. ここで, 目的関数 F が凸なので,

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

となる. よって \bar{y} は (P) の大域最小解になる. \square

10.2 一般の汎関数に対する最適性必要条件

一般の汎関数に対しても数ベクトル空間のような主張が言える.

定理 23. \bar{y} を問題 (P) の局所最小解とする. すると, オイラー・ラグランジュの方程式を満たす. 言い換えると,

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

が成り立つ.

証明. 省略する.

「すべての $v(a) = v(b) = 0$ を満たす v に対して, $DF(\bar{y})(v) = 0$ 」 \Leftrightarrow 「 $\frac{d}{dx} f_z[\bar{y}(x)] = f_y[\bar{y}(x)]$ 」が成り立つ. \square

系 24. 最小化問題 (P) において, 目的関数 F を凸とする. すると, (P) の局所最小解はすべて大域最小解になり,

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

補足. 定理 23 を示すには, まず局所最小解をきちんと定義するところから始めなければならない.

関数の集合 $C = \{y \in C^1[a, b] \mid y(a) = A, y(b) = B\}$ とおく. 関数 $\bar{y} \in C$ に充分近い任意の $y \in C$ に対して

$$F(y) \geq F(\bar{y})$$

が成り立つとき、 \bar{y} を問題 (P) の局所最小解と呼ぶ。

ここで問題となってくるのは、「近い」という言葉の意味である。最適化する変数が数ベクトルの場合には二点の距離が小さいとき近いとした。変分問題の場合は変数が関数である。

よって、関数同士の「近さ」を図るために 2 つの関数の距離を定義しなければならない。こうやって話を進めると、 \bar{y} が局所最小解ならば、任意の $y \in C$ に対して、 $DF(\bar{y})(y - \bar{y}) = 0$ を得ることができる。

10.3 解法例

例 24.

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

$$\text{制約 } y(0) = 1, y(1) = 2$$

目的関数の被積分関数は $f(x, y, z) = y + z^2$ となり、これは第 2, 第 3 変数に関して凸である。定理 22 より、オイラー・ラグランジュの方程式 $\frac{d}{dx} f_z[y(x)] = f_y[y(x)]$ と $y(0) = 1, y(1) = 2$ 満たす関数が大域最小解になる。

いま、 $f_y = 1, f_z = 2z$ なので、オイラー・ラグランジュの方程式は

$$\frac{d}{dx} \{2y'(x)\} = 1$$

となる。これは微分方程式

$$2y'' = 1$$

を表す。よって、両辺を 2 回不定積分することにより、

$$y' = \frac{1}{2}x + c_1$$

$$y = \frac{1}{4}x^2 + c_1x + c_2$$

を得る。ここで、 $y(0) = 1, y(1) = 2$ より、 $c_2 = 1, 1/4 + c_1 + c_2 = 2$ を得る。よって、

$$\bar{y}(x) = \frac{1}{4}x^2 + \frac{3}{4}x + 1$$

が求める大域最小解になる。

実際、 $v(a) = v(b) = 0$ となる v に対して、 $DF(y)(v) = \int \{v(x) + 2y'(x)v'(x)\} dx$ に \bar{y} を代入すると、 $DF(\bar{y})(v) = 0$ となる。いま F は凸汎関数なので、

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

となる。