

最適化数学 第2回

[今回の項目]

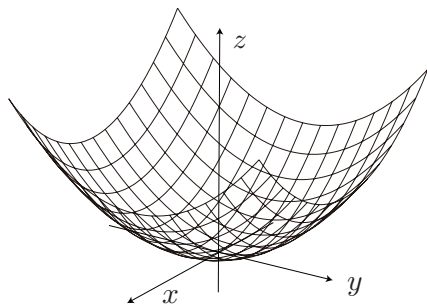
- ① 数学的準備：多変数関数
- ② 凸関数の性質
- ③ 凸関数の判定（ヘッセ行列）
- ④ 行列の正値性

多変数関数

$$f(x, y) = x^2 + y^2$$

いくつかの (x, y) において関数がどのような値をとるか調べると、右のようになる。

$x \setminus y$	-2	-1	0	1	2
-2	8	5	4	5	8
-1	5	2	1	2	5
0	4	1	0	1	4
1	5	2	1	2	5
2	8	5	4	5	8



次に、関数の値を幾何的にとらえるために

$$z = x^2 + y^2$$

とおく。この関数のグラフを xyz 空間に描くと、左図のようになる。

等高線

$f(x, y) = x^2 + y^2$ で、同じ値をとる点 (x, y) からなる曲線を等高線と呼ぶ。

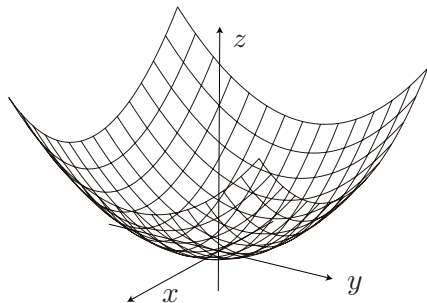


Figure: $z = f(x, y)$ のグラフ

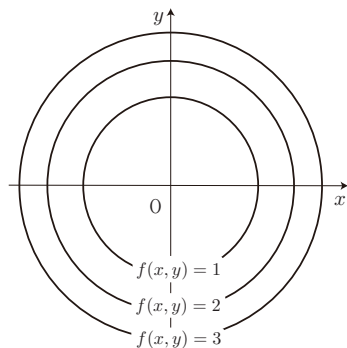


Figure: $z = x^2 + y^2$ の等高線

等高線

参考に $f(x, y) = -x^3 - 3xy^2 + y^3 + 3x$ のグラフと等高線も挙げておこう.

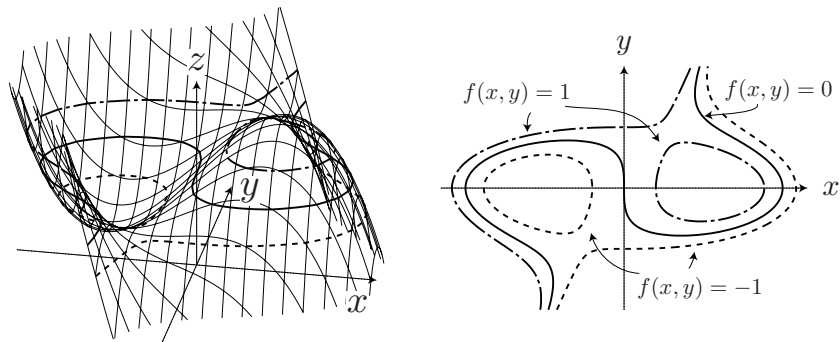


Figure: $f(x, y) = -x^3 - 3xy^2 + y^3 + 3x$ のグラフと等高線

凸関数

一般には 2 次関数のように下に凸な関数は、最小値を求めることが比較的楽である。下に凸な関数は、最適化問題においては単に凸関数と呼ぶ。

最適化の世界では、「凸である」か「凸でない」かが問題の分かれ目になる。凸関数は扱いやすいだけでなく、応用上も頻繁に現れる関数である。

凸関数の定義

[定義]

f を n 変数関数とする. $0 < \lambda < 1$ を満たすすべての λ とすべての $u, v \in \mathbb{R}^n$ に対して,

$$\text{凸関数} : f((1 - \lambda)u + \lambda v) \leq (1 - \lambda)f(u) + \lambda f(v)$$

$$\text{狭義凸関数} : f((1 - \lambda)u + \lambda v) < (1 - \lambda)f(u) + \lambda f(v) \quad (u \neq v)$$

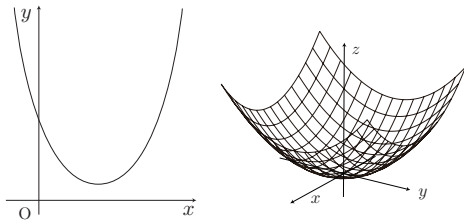


Figure: 凸関数の例 左が 1 変数凸関数, 右が 2 変数凸関数

定義の解説：狭義凸関数

説明がやさしい順に、はじめに狭義凸関数について解説する。
定義式

$$\begin{aligned} & f((1 - \lambda)u + \lambda v) \\ & < (1 - \lambda)f(u) + \lambda f(v) \\ & (u \neq v) \end{aligned}$$

は、

「 f が狭義凸関数」



「点 $(u, f(u))$ と $(v, f(v))$
を結んだ線分が、

f グラフより上部にある」

ということを表す。

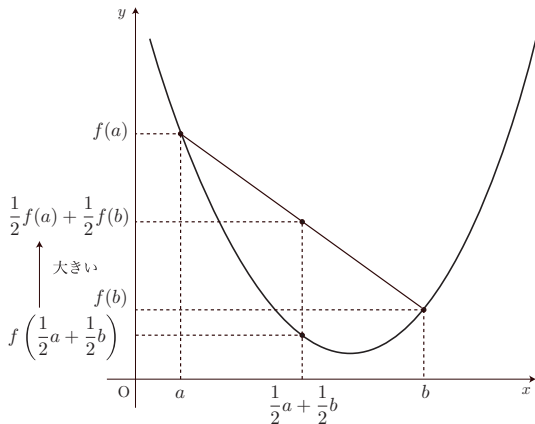


Figure: 狭義凸関数に対して、定義式で $\lambda = \frac{1}{2}$ とした場合

定義の解説：凸関数

次に、凸関数について考えよう。
狭義凸関数と違い凸関数では

$$\begin{aligned} f((1-\lambda)u + \lambda v) \\ \leq (1-\lambda)f(u) + \lambda f(v) \end{aligned}$$

と等号も許されている。これは

「 f が凸関数」



「点 $(u, f(u))$ と $(v, f(v))$ を結んだ
線分が、 f のグラフ上、またはグラフより上部にある」

ということを表している。

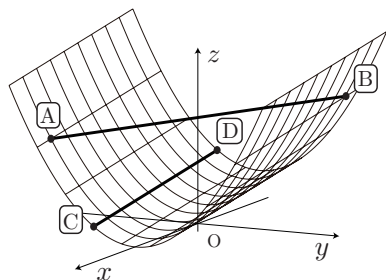


Figure: 狭義凸関数でない凸関数の例

2 変数凸関数と接平面

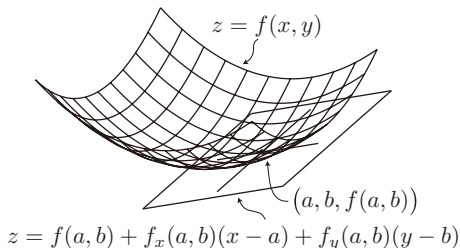
2 変数関数 f の点 $(a, b, f(a, b))$ における接平面：

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

2 変数関数 f が凸ならば,

$$f(x, y) \geq f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

が成り立つ。



準備：勾配ベクトル

[定義] (勾配ベクトル)

$$2 \text{ 変数関数 } f \text{ の場合： } \nabla f(x, y) = \begin{bmatrix} f_x(x, y) \\ f_y(x, y) \end{bmatrix}$$

$$n \text{ 変数関数 } f \text{ の場合： } \nabla f(u) = \begin{bmatrix} f_{x_1}(u) \\ \vdots \\ f_{x_n}(u) \end{bmatrix}$$

Example

$f(x, y) = x^2 + 3y^2$ とすると、点 $(5, 3)$ における勾配ベクトルは、 $f_x = 2x$, $f_y = 6y$ より、

$$\nabla f(5, 3) = \begin{bmatrix} f_x(5, 3) \\ f_y(5, 3) \end{bmatrix} = \begin{bmatrix} 2 \cdot 5 \\ 6 \cdot 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 18 \end{bmatrix}$$

となる。

n 変数凸関数と接平面

[命題]

n 変数関数 f に対して、以下が成り立つ：

f が凸関数

$$\iff f(v) \geq f(u) + \nabla f(u) \cdot (v - u) \quad (\text{すべての } u, v \in \mathbb{R}^n)$$

Proof.

テイラーの定理を用いる。詳しくは教科書。 □

- 勾配ベクトル： $\nabla f(u) = \begin{bmatrix} f_{x_1}(u) \\ \vdots \\ f_{x_n}(u) \end{bmatrix}$
- $\nabla f(u) \cdot (v - u)$ はベクトル $\nabla f(u)$ と $(v - u)$ の内積を表す。

凸関数の判定

関数が凸かどうかを，関数の微分を用いて判定することができる。

凸関数の判定 (1 変数関数)

関数 $f(x) = x^2$ はグラフより, 凸関数. それでは, 次の関数は?

$$f(x) = \sqrt{1 + x^2}$$

[定理]

1 変数関数 h に対して, 以下が成り立つ:

$$h \text{ が凸関数} \Leftrightarrow \text{すべての } t \in \mathbb{R} \text{ に対して } h''(t) \geq 0$$

Proof.

テイラーの定理より. (直感的には, 増減表より) □

$f'(x) = \frac{x}{\sqrt{1+x^2}}$, $f''(x) = \frac{1}{(1+x^2)^{\frac{3}{2}}}$ より, すべての x に対して $f''(x) > 0$ となるので, 定理より, $f(x)$ は凸関数である.

多変数の場合：ヘッセ行列

次に，2変数関数

$$g(x, y) = x^2 + 2xy + 2y^2$$

は凸関数か？多変数関数の2階微分は，ヘッセ行列に対応する．

[定義]

2変数関数に対して，ヘッセ行列 $\nabla^2 f(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix}$

Example

$f(x, y) = x^3 + 2xy + 3y^2$ とすると,

$$\text{勾配ベクトル } \nabla f(x, y) = \begin{bmatrix} 3x^2 + 2y \\ 2x + 6y \end{bmatrix},$$

$$\text{ヘッセ行列 } \nabla^2 f(x, y) = \begin{bmatrix} 6x & 2 \\ 2 & 6 \end{bmatrix}$$

一般の n 変数関数 f と $u = (x_1, \dots, x_n)$ に対しては,

$$\text{ヘッセ行列 } \nabla^2 f(u) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(u) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(u) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(u) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(u) & \ddots & & \vdots \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(u) & \cdots & & \frac{\partial^2 f}{\partial x_n \partial x_n}(u) \end{bmatrix}$$

ヘッセ行列による凸性の判定

[定理]

① f が凸関数 \iff 各点 $a \in \mathbb{R}^n$ で,

$${}^t u \nabla^2 f(a) u \geq 0 \quad (\text{すべてのベクトル } u \in \mathbb{R}^n).$$

② f が狭義凸関数 \iff 各点 $a \in \mathbb{R}^n$ で,

$${}^t u \nabla^2 f(a) u > 0 \quad (u \neq 0 \text{ となるすべてのベクトル } u \in \mathbb{R}^n).$$

Proof.

$b, c \in \mathbb{R}^n$ に対して, $h(t) = f((1-t)b + tc)$, $a = (1-t)b + tc$ とおく. すると

f が凸 \iff すべての $b, c \in \mathbb{R}^n$ に対して $h(t)$ が凸

$$\iff \text{すべての } b, c \in \mathbb{R}^n, t \in \mathbb{R} \text{ に対して } \frac{d^2}{dt^2} h(t) \geq 0$$

が成り立つ. ここで, 合成関数の微分より, $\frac{d^2}{dt^2} f(t) = {}^t u \nabla^2 f(a) u$ となる. \square

Example

$f(x, y) = x^2 - 2xy + 2y^2$ とすると, 勾配ベクトルとヘッセ行列は

$$\nabla f(x, y) = \begin{bmatrix} 2x - 2y \\ -2x + 4y \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

となる. 定理を用いて f が凸関数かどうか調べよう. $u = {}^t(u_1, u_2)$ とすると, 各点 (x, y) で,

$${}^t u \nabla f(x, y) u = [u_1 \ u_2] \begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 2u_1^2 - 4u_1u_2 + 4u_2^2$$

となる. さらに, すべてのベクトル u に対して

$$2u_1^2 - 4u_1u_2 + 4u_2^2 = 2(u_1^2 - 2u_1u_2 + 2u_2^2) = 2\{(u_1 - u_2)^2 + u_2^2\} \geq 0$$

となるので, 定理より f は凸関数である.

線形代数を用いた凸関数の判定

線形代数を用いると，計算により関数の凸性を判定することが出来る．

線形代数：行列と2次形式

Definition

行列 A とベクトル $u = (u_1, u_2, \dots, u_n)$ に対して, u_1, \dots, u_n を変数に持つ多項式

$$p(u) = {}^t u A u$$

を2次形式と呼ぶ.

Example

行列 $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ に対して $u = (x, y)$ とすると,

$$(1) \quad {}^t u A u = [x \ y] \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x \ y] \begin{bmatrix} 2x \\ y \end{bmatrix} = 2x^2 + y^2$$

$$(2) \quad {}^t u B u = [x \ y] \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [x \ y] \begin{bmatrix} x + y \\ x - y \end{bmatrix} \\ = x^2 + xy + yx - y^2 = x^2 + 2xy - y^2$$

は2次形式である.

線形代数：行列の正値性

[定義]

行列 A を対称行列とし，2次形式 ${}^t u A u$ を考える．

- A は 半正定値 $\iff {}^t u A u \geq 0$ (すべての $u \in \mathbb{R}^n$)
また，逆の不等号が成り立つとき， A を 半負定値 という．
- A は 正定値 $\iff {}^t u A u > 0$ (すべての $u \neq \mathbf{0}$)
逆の不等号が成り立つとき， A を 負定値 という．
- A は 不定 $\iff u \in \mathbb{R}^n$ によって ${}^t u A u$ の符号が正にも負にもなる．

Example

行列 $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ の正値性を調べてみよう. $u = (x, y)$ とすると, 2次形式は

$${}^t u A v = 2x^2 + y^2$$

となる. ここで, $2x^2 + y^2 > 0$ ($u \neq 0$) となるので, A は正定値である.

一方, 行列 $B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ に対して, 2次形式は

$${}^t u B u = x^2 + 2xy - y^2$$

となる. ここで,

- ① $u = (1, 0)$ とすると ${}^t u B u = 1 > 0$ となるが,
- ② $u = (0, 1)$ とすると ${}^t u B u = -1 < 0$ となるので,

B は不定である.

[定理] (前出の定理の言い換え)

関数 f に対して以下が成り立つ：

- f が凸関数
 \iff 各点 $a \in \mathbb{R}^n$ において、ヘッセ行列 $\nabla^2 f(a)$ が半正定値である.
- f が狭義凸関数
 \iff 各点 $a \in \mathbb{R}^n$ において、ヘッセ行列 $\nabla^2 f(a)$ が正定値である.

以下で説明するように、固有値または行列式を調べることで、行列の正值性を判定することができる！

正定値行列と固有値

[定理]

A を対称行列とする。以下の主張が成り立つ。

- ① A が半正定値 $\iff A$ の固有値がすべて 0 以上
- ② A が正定値 $\iff A$ の固有値がすべて正
- ③ A が不定 $\iff A$ が正と負の固有値を持つ

Proof.

対称行列は、直交行列で対角化できるという性質を用いる。 □

Example

$f(x, y) = 3x^2 - 2xy + 3y^2$ とすると,

$\nabla f(x, y) = \begin{bmatrix} 6x - 2y \\ 2x - 6y \end{bmatrix}$, $\nabla^2 f(x, y) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$ となる。ヘッセ行列は定数行列であり、この固有値は 4, 8 になる。定理より、ヘッセ行列は正定値で、 f は狭義凸関数である。

正定値行列と行列式

[定理] (特殊例: 2×2 行列)

A を 2×2 行列とする.

- 1 $|A| > 0$ かつ $a > 0 \Leftrightarrow A$ は正定値
- 2 $|A| > 0$ かつ $a < 0 \Leftrightarrow A$ は負定値
- 3 $|A| < 0 \Leftrightarrow A$ は不定

Proof.

A の成分を文字でおいてから, 2 次形式を展開し, 平方完成する. 後は, (行列式) = (固有値) の積を用いて場合分け. \square

Example

$f(x, y) = 3x^2 - 2xy + 3y^2$ のヘッセ行列の $\nabla^2 f(x, y) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$ 正值性を行列式を用いて調べると, $|\nabla^2 f(x, y)| = 32 > 0$ かつ $f_{xx}(x, y) = 6 > 0$ なので, ヘッセ行列は正定値である.

[練習問題]

以下の関数の勾配ベクトルとヘッセ行列を求め、凸関数かどうか調べよ.

$$(1). f(x, y) = x^2 + xy + y^2 \quad (2). f(x, y) = 2x^2 + 4xy + y^2$$

$$(3). f(x, y) = x^2 + xy + y^3 \quad (4). f(x, y) = \sqrt{1 + x^2 + y^2}$$

$$(5). f(x, y, z) = x^2 + y^2 + z^2 - xy - yz - zx$$