

数学補習プログラム（社会人院生向け）

トピック 15：凹関数・凸関数

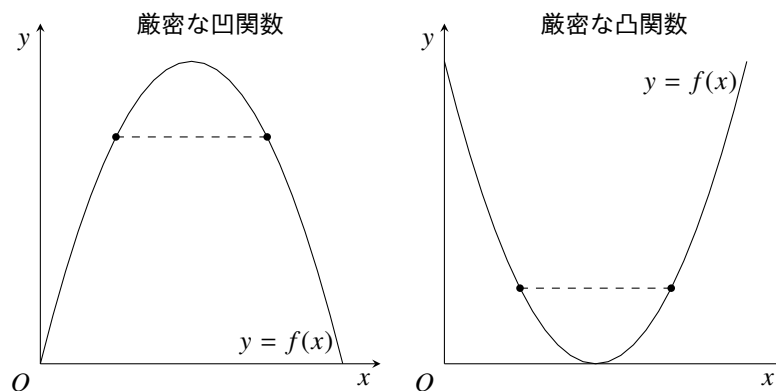
北村友宏*

2016年3月26日

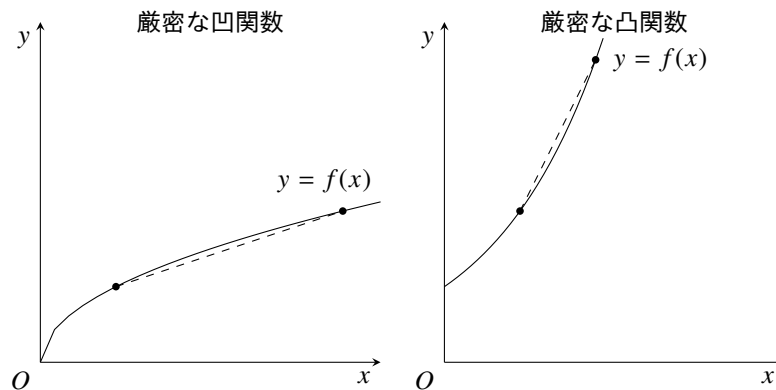
1 凹関数・凸関数（参考書上巻 pp.439-459）

1.1 1変数関数の場合

- 凹関数・厳密な凹関数・凸関数・厳密な凸関数の形状（定義ではない）は以下の通り。
 - ★ 凹関数：関数のグラフ上のどの2点をとっても、その2点を結ぶ直線上の（とってきた2点を除く）全ての点の高さが、その2点の間の関数のグラフ上の（とってきた2点を除く）全ての点の高さと同じかそれよりも低い。
 - ★ 厳密な凹関数：関数のグラフ上のどの2点をとっても、その2点を結ぶ直線上の（とってきた2点を除く）全ての点の高さが、その2点の間の関数のグラフ上の（とってきた2点を除く）全ての点の高さよりも低い。
 - ★ 凸関数：関数のグラフ上のどの2点をとっても、その2点を結ぶ直線上の（とってきた2点を除く）全ての点の高さが、その2点の間の関数のグラフ上の（とってきた2点を除く）全ての点の高さと同じかそれよりも高い。
 - ★ 厳密な凸関数：関数のグラフ上のどの2点をとっても、その2点を結ぶ直線上の（とってきた2点を除く）全ての点の高さが、その2点の間の関数のグラフ上の（とってきた2点を除く）全ての点の高さよりも高い。



* Email: kitamu.tom@gmail.com URL: <http://tomkitamura.html.xdomain.jp>



- 凹関数・厳密な凹関数・凸関数・厳密な凸関数の定義は以下の通り.

- ★ 関数 $y = f(x)$ の定義域における任意の u, v (ただし $u \neq v$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u) + (1 - \theta)f(v) \leq f[\theta u + (1 - \theta)v]$$

となる $f(x)$ を凹関数という.

- ★ 関数 $y = f(x)$ の定義域における任意の u, v (ただし $u \neq v$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u) + (1 - \theta)f(v) < f[\theta u + (1 - \theta)v]$$

となる $f(x)$ を厳密な凹関数という.

* 狭義凹関数または強い意味の凹関数ということもある.

- ★ 関数 $y = f(x)$ の定義域における任意の u, v (ただし $u \neq v$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u) + (1 - \theta)f(v) \geq f[\theta u + (1 - \theta)v]$$

となる $f(x)$ を凸関数という.

- ★ 関数 $y = f(x)$ の定義域における任意の u, v (ただし $u \neq v$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u) + (1 - \theta)f(v) > f[\theta u + (1 - \theta)v]$$

となる $f(x)$ を厳密な凸関数という.

* 狭義凸関数または強い意味の凸関数ということもある.

関数 $y = f(x)$ について, 以下のことがいえる.

- 定義域内の任意の x について, $f''(x) \leq 0$ ならば $f(x)$ は凹関数となる.
- 定義域内の任意の x について, $f''(x) < 0$ ならば $f(x)$ は厳密な凹関数となる.
⇒ 厳密な凹関数の場合, $f'(x^*) = 0$ となる停留点 $x = x^*$ において, $f(x^*)$ は最大値.
- 定義域内の任意の x について, $f''(x) \geq 0$ ならば $f(x)$ は凸関数となる.
- 定義域内の任意の x について, $f''(x) > 0$ ならば $f(x)$ は厳密な凸関数となる.
⇒ 厳密な凸関数の場合, $f'(x^*) = 0$ となる停留点 $x = x^*$ において, $f(x^*)$ は最小値.

1.2 2変数関数の場合

- 凹関数・厳密な凹関数・凸関数・厳密な凸関数の定義は以下の通り.

- ★ 関数 $z = f(x_1, x_2)$ の定義域における任意の $(u_1, u_2), (v_1, v_2)$ (ただし $(u_1, u_2) \neq (v_1, v_2)$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u_1, u_2) + (1 - \theta)f(v_1, v_2) \leq f[\theta u_1 + (1 - \theta)v_1, \theta u_2 + (1 - \theta)v_2]$$

となる $f(x_1, x_2)$ を凹関数という.

- ★ 関数 $z = f(x_1, x_2)$ の定義域における任意の $(u_1, u_2), (v_1, v_2)$ (ただし $(u_1, u_2) \neq (v_1, v_2)$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u_1, u_2) + (1 - \theta)f(v_1, v_2) < f[\theta u_1 + (1 - \theta)v_1, \theta u_2 + (1 - \theta)v_2]$$

となる $f(x_1, x_2)$ を厳密な凹関数という.

- ★ 関数 $z = f(x_1, x_2)$ の定義域における任意の $(u_1, u_2), (v_1, v_2)$ (ただし $(u_1, u_2) \neq (v_1, v_2)$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u_1, u_2) + (1 - \theta)f(v_1, v_2) \geq f[\theta u_1 + (1 - \theta)v_1, \theta u_2 + (1 - \theta)v_2]$$

となる $f(x_1, x_2)$ を凸関数という.

- ★ 関数 $z = f(x_1, x_2)$ の定義域における任意の $(u_1, u_2), (v_1, v_2)$ (ただし $(u_1, u_2) \neq (v_1, v_2)$) と, $0 < \theta < 1$ を満たす任意の θ について,

$$\theta f(u_1, u_2) + (1 - \theta)f(v_1, v_2) > f[\theta u_1 + (1 - \theta)v_1, \theta u_2 + (1 - \theta)v_2]$$

となる $f(x_1, x_2)$ を厳密な凸関数という.

関数 $z = f(x_1, x_2)$ について, その2階全微分 d^2z を考えると, 以下のことがいえる.

- どこでも $d^2z \leq 0$ (2階全微分がどこでも非正) ならば $f(x_1, x_2)$ は凹関数となる.
- どこでも $d^2z < 0$ (2階全微分がどこでも負) ならば $f(x_1, x_2)$ は厳密な凹関数となる.
- どこでも $d^2z \geq 0$ (2階全微分がどこでも非負) ならば $f(x_1, x_2)$ は凸関数となる.
- どこでも $d^2z > 0$ (2階全微分がどこでも正) ならば $f(x_1, x_2)$ は厳密な凸関数となる.

2 ヘッセ行列式 (参考書上巻 pp.417-433, 449-453)

2.1 判別式による凹・厳密な凹・凸・厳密な凸の判断

- 2変数関数 $z = f(x_1, x_2)$ の1階全微分は,

$$dz = \frac{\partial z}{\partial x_1} dx_1 + \frac{\partial z}{\partial x_2} dx_2 = f_1 dx_1 + f_2 dx_2$$

であり、2階全微分は、

$$\begin{aligned}
 d^2z &= d(dz) \\
 &= \left(\frac{\partial}{\partial x_1} [dz] \right) dx_1 + \left(\frac{\partial}{\partial x_2} [dz] \right) dx_2 \\
 &= \left(\frac{\partial}{\partial x_1} [f_1 dx_1 + f_2 dx_2] \right) dx_1 + \left(\frac{\partial}{\partial x_2} [f_1 dx_1 + f_2 dx_2] \right) dx_2 \\
 &= (f_{11} dx_1 + f_{12} dx_2) dx_1 + (f_{21} dx_1 + f_{22} dx_2) dx_2 \\
 &= f_{11} (dx_1)^2 + f_{12} (dx_2)(dx_1) + \underbrace{f_{21}}_{=f_{12}} (dx_1)(dx_2) + f_{22} (dx_2)^2 \\
 &= f_{11} dx_1^2 + 2f_{12} dx_1 dx_2 + f_{22} dx_2^2
 \end{aligned}$$

★ 微小な変化を表す dx_1 や dx_2 はそれぞれ1つの記号であり、定数として扱う。

★ f_1 は x_1 と x_2 の関数、 f_2 も x_1 と x_2 の関数。

★ ヤングの定理より、 $f_{21} = f_{12}$ 。

- d^2z を計算して符号・値を確認すれば、 $z = f(x_1, x_2)$ が凹関数か厳密な凹関数か凸関数か厳密な凸関数かを判断できる。 f_{11} 、 f_{12} 、 f_{22} は計算できるが、 d^2z には他に微小な変化を表す dx_1 や dx_2 が入っている。

⇒ d^2z が常に正であることや常に負であることを確認するには、判別式を用いる。

- 以下では厳密な凹関数か厳密な凸関数かの判断に限定して説明する。

- $f_{11} > 0$ とする。

★ $d^2z = f_{11} dx_1^2 + 2f_{12} dx_1 dx_2 + f_{22} dx_2^2 = f_{11} (dx_1)^2 + 2f_{12} dx_2 (dx_1) + f_{22} dx_2^2$ を dx_1 の2次関数であると考えて、「判別式が負」、すなわち

$$D = (2f_{12} dx_2)^2 - 4f_{11} f_{22} dx_2^2 < 0 \iff \underbrace{4dx_2^2}_{>0} (f_{12}^2 - f_{11} f_{22}) < 0 \iff f_{12}^2 - f_{11} f_{22} < 0$$

が成立すれば、任意の dx_1, dx_2 について d^2z が常に正となる。よって $f_{11} > 0$ と $f_{12}^2 - f_{11} f_{22} < 0$ が 定義域内の任意の x_1 と x_2 について成立すれば、定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について d^2z が常に正となり、 $f(x_1, x_2)$ は厳密な凸関数となる。

- $f_{11} < 0$ とする。

★ 判別式を適用するため、 dx_1^2 に掛かっている係数を正にする。そのために、次のように考える。

$$\begin{aligned}
 d^2z &= \overbrace{f_{11}}^{<0} dx_1^2 + 2f_{12} dx_1 dx_2 + f_{22} dx_2^2 \\
 \iff -d^2z &= \underbrace{-f_{11}}_{>0} (dx_1)^2 - 2f_{12} dx_2 (dx_1) - f_{22} dx_2^2
 \end{aligned}$$

を dx_1 の2次関数であると考えて、 $-d^2z$ に対して「判別式が負」、すなわち

$$\begin{aligned}
 D &= (-2f_{12} dx_2)^2 - 4(-f_{11})(-f_{22} dx_2^2) < 0 \iff 4f_{12}^2 dx_2^2 - 4f_{11} f_{22} dx_2^2 < 0 \\
 &\iff \underbrace{4dx_2^2}_{>0} (f_{12}^2 - f_{11} f_{22}) < 0 \\
 &\iff f_{12}^2 - f_{11} f_{22} < 0
 \end{aligned}$$

が成立すれば、任意の dx_1, dx_2 について $-d^2z$ が常に正となる。このとき、 d^2z は常に負となる。よって $f_{11} < 0$ と $f_{12}^2 - f_{11} f_{22} < 0$ が 定義域内の任意の x_1 と x_2 について成立すれば、

定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について d^2z が常に負となり, $f(x_1, x_2)$ は厳密な凹関数となる.

$\Rightarrow f_{11} > 0$ のときも $f_{11} < 0$ のときも, $f_{12}^2 - f_{11}f_{22} < 0$ が厳密な凹・厳密な凸関数の判断基準となる.

• 以上をまとめると,

- ★ 定義域内の任意の x_1 と x_2 について, $f_{11} < 0$ かつ $f_{12}^2 - f_{11}f_{22} < 0$ のとき, 定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について, $d^2z < 0$. よって $f(x_1, x_2)$ は厳密な凹関数となる.
- ★ 定義域内の任意の x_1 と x_2 について, $f_{11} > 0$ かつ $f_{12}^2 - f_{11}f_{22} < 0$ のとき, 定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について, $d^2z > 0$. よって $f(x_1, x_2)$ は厳密な凸関数となる.

2.2 判別式からヘッセ行列式へ

- $f_{12}^2 - f_{11}f_{22}$ はヘッセ行列式 (ヘッシアンともいう) にマイナスをつけたもの (-1 を掛けたもの). ヘッセ行列式は,

$$|\mathbf{H}| = \begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} = \underbrace{f_{11}f_{22}}_{\text{右下がりの要素の積}} - \underbrace{f_{12}f_{12}}_{\text{右上がりの要素の積}} = f_{11}f_{22} - f_{12}^2 = -(f_{12}^2 - f_{11}f_{22}).$$

\Rightarrow 変数関数 $z = f(x_1, x_2)$ について, ヘッセ行列式を用いた厳密な凹・厳密な凸関数の判断は以下の通り.

- 定義域内の任意の x_1 と x_2 について, $f_{11} < 0$ かつ $\begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} > 0$ のとき, 定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について, $d^2z < 0$. よって $f(x_1, x_2)$ は厳密な凹関数となる.
- 定義域内の任意の x_1 と x_2 について, $f_{11} > 0$ かつ $\begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} > 0$ のとき, 定義域内の任意の x_1, x_2 と任意の dx_1, dx_2 について, $d^2z > 0$. よって $f(x_1, x_2)$ は厳密な凸関数となる.

例題 2.2.1 $z = f(x_1, x_2) = 3x_1^2 + 2x_1x_2 + 4x_2^2$ が厳密な凹関数か, 厳密な凸関数かをヘッセ行列式を用いて判断しなさい.

解法

- まず、全ての変数 (x_1 と x_2) について 1 次偏導関数を求める。
- 「 x_1 に関する 1 次偏導関数」をさらに x_1 に関して偏微分すると「 x_1 に関する 2 次偏導関数」が、「 x_2 に関する 1 次偏導関数」をさらに x_2 に関して偏微分すると「 x_2 に関する 2 次偏導関数」が導出できる。
 \Rightarrow 2 次偏導関数は、 $f_{11} = \frac{\partial^2 z}{\partial x_1^2}$, $f_{22} = \frac{\partial^2 z}{\partial x_2^2}$ の 2 つ。
- 続いて、「 x_2 に関する 1 次偏導関数」をさらに x_1 に関して偏微分すると片方の交差偏導関数が導出できる。それはもう片方の交差偏導関数と同じになる (ヤングの定理)。
 \Rightarrow 交差偏導関数は、 $f_{12} = \frac{\partial^2 z}{\partial x_1 \partial x_2} = f_{21}$ 。
- 2 次偏導関数と交差偏導関数からヘッセ行列式を求め、その符号・値から判断する。

1 次偏導関数は、

$$f_1 = \frac{\partial z}{\partial x_1} = \underbrace{3}_{\text{定数}} \cdot 2x_1^{2-1} + \underbrace{2}_{\text{定数}} \cdot 1x_1^{1-1} \underbrace{x_2}_{\text{定数}} + \underbrace{0}_{\text{定数の微分}} = 6x_1 + 2 \underbrace{x_1^0}_{=1} x_2 = 6x_1 + 2x_2,$$

$$f_2 = \frac{\partial z}{\partial x_2} = \underbrace{0}_{\text{定数の微分}} + \underbrace{2x_1}_{\text{定数}} \cdot 1x_2^{1-1} + \underbrace{4}_{\text{定数}} \cdot 2x_2^{2-1} = 2x_1 \underbrace{x_2^0}_{=1} + 8x_2 = 2x_1 + 8x_2$$

である。よって、2 次偏導関数は、

$$f_{11} = \frac{\partial^2 z}{\partial x_1^2} = \underbrace{6}_{\text{定数}} \cdot 1x_1^{1-1} + \underbrace{0}_{\text{定数の微分}} = 6 \underbrace{x_1^0}_{=1} = 6,$$

$$f_{22} = \frac{\partial^2 z}{\partial x_2^2} = \underbrace{0}_{\text{定数の微分}} + \underbrace{8}_{\text{定数}} \cdot 1x_2^{1-1} = 8 \underbrace{x_2^0}_{=1} = 8$$

となる。また、交差偏導関数は、

$$f_{12} = \frac{\partial^2 z}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1} \left(\frac{\partial z}{\partial x_2} \right) = \frac{\partial}{\partial x_1} [2x_1 + 8x_2] = \underbrace{2}_{\text{定数}} \cdot 1x_1^{1-1} + \underbrace{0}_{\text{定数の微分}} = 2 \underbrace{x_1^0}_{=1} = 2$$

である。したがって、ヘッセ行列式は、

$$|\mathbf{H}| = \begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} = \begin{vmatrix} 6 & 2 \\ 2 & 8 \end{vmatrix} = \underbrace{6 \cdot 8}_{\text{右下がりの要素の積}} - \underbrace{2 \cdot 2}_{\text{右上がりの要素の積}} = 48 - 4 = 44$$

となる。よって、 $f_{11} = 6 > 0$ かつ $\begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} = 44 > 0$ であることから、 $z = f(x_1, x_2)$ は厳密な凸関数である。

例題 2.2.2 $z = f(x_1, x_2) = -x_1^2 + 2x_1x_2 - 3x_2^2$ が厳密な凹関数か、厳密な凸関数かをヘッセ行列式を用いて判断しなさい。

解法

- まず、全ての変数 (x_1 と x_2) について 1 次偏導関数を求める。
- 「 x_1 に関する 1 次偏導関数」をさらに x_1 に関して偏微分すると「 x_1 に関する 2 次偏導関数」が、「 x_2 に関する 1 次偏導関数」をさらに x_2 に関して偏微分すると「 x_2 に関する 2 次偏導関数」が導出できる。
 \Rightarrow 2 次偏導関数は、 $f_{11} = \frac{\partial^2 z}{\partial x_1^2}$, $f_{22} = \frac{\partial^2 z}{\partial x_2^2}$ の 2 つ。
- 続いて、「 x_2 に関する 1 次偏導関数」をさらに x_1 に関して偏微分すると片方の交差偏導関数が導出できる。それはもう片方の交差偏導関数と同じになる (ヤングの定理)。
 \Rightarrow 交差偏導関数は、 $f_{12} = \frac{\partial^2 z}{\partial x_1 \partial x_2} = f_{21}$ 。
- 2 次偏導関数と交差偏導関数からヘッセ行列式を求め、その符号・値から判断する。

1 次偏導関数は、

$$f_1 = \frac{\partial z}{\partial x_1} = \underbrace{-1}_{\text{定数}} \cdot 2x_1^{2-1} + \underbrace{2}_{\text{定数}} \cdot 1x_1^{1-1} \cdot \underbrace{x_2}_{\text{定数}} - \underbrace{0}_{\text{定数の微分}} = -2x_1 + 2 \underbrace{x_1^0}_{=1} x_2 = -2x_1 + 2x_2,$$

$$f_2 = \frac{\partial z}{\partial x_2} = \underbrace{0}_{\text{定数の微分}} + \underbrace{2x_1}_{\text{定数}} \cdot 1x_2^{1-1} - \underbrace{3}_{\text{定数}} \cdot 2x_2^{2-1} = 2x_1 \underbrace{x_2^0}_{=1} - 6x_2 = 2x_1 - 6x_2$$

である。よって、2 次偏導関数は、

$$f_{11} = \frac{\partial^2 z}{\partial x_1^2} = \underbrace{-2}_{\text{定数}} \cdot 1x_1^{1-1} + \underbrace{0}_{\text{定数の微分}} = -2 \underbrace{x_1^0}_{=1} = -2,$$

$$f_{22} = \frac{\partial^2 z}{\partial x_2^2} = \underbrace{0}_{\text{定数の微分}} - \underbrace{6}_{\text{定数}} \cdot 1x_2^{1-1} = -6 \underbrace{x_2^0}_{=1} = -6$$

となる。また、交差偏導関数は、

$$f_{12} = \frac{\partial^2 z}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1} \left(\frac{\partial z}{\partial x_2} \right) = \frac{\partial}{\partial x_1} [2x_1 - 6x_2] = \underbrace{2}_{\text{定数}} \cdot 1x_1^{1-1} - \underbrace{0}_{\text{定数の微分}} = 2 \underbrace{x_1^0}_{=1} = 2$$

である。したがって、ヘッセ行列式は、

$$|\mathbf{H}| = \begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} = \begin{vmatrix} -2 & 2 \\ 2 & -6 \end{vmatrix} = \underbrace{(-2) \cdot (-6)}_{\text{右下がりの要素の積}} - \underbrace{2 \cdot 2}_{\text{右上がりの要素の積}} = 12 - 4 = 8$$

となる。よって、 $f_{11} = -2 < 0$ かつ $\begin{vmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{vmatrix} = 8 > 0$ であることから、 $z = f(x_1, x_2)$ は厳密な凹関数である。