第 V 部　Logistic Model

3　尤度関数の最大値

ポイント

- 単調関数としての対数尤度
- 対数尤度の上限
- 単調関数の和

3.1　はじめに

- 特別な状況を除き、ロジスティックモデルの尤度関数は最大値をただ一つ持ちます。
- 尤度関数で説明することは難しいので対数尤度関数で説明します。

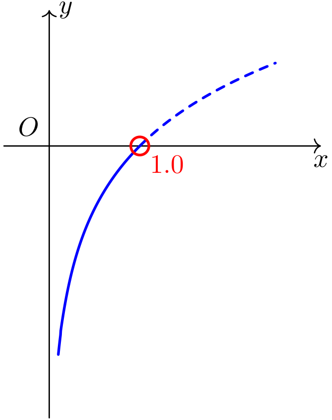
3.1.1　自然対数の値域

底をネイピア数 e とする対数関数

$$y = \ln x \tag{1}$$

は狭義単調増加関数であって、x の範囲が $(0, 1)$ のとき対応する y の範囲は $(-\infty, 0)$ である。

図 1　e を底とする対数関数



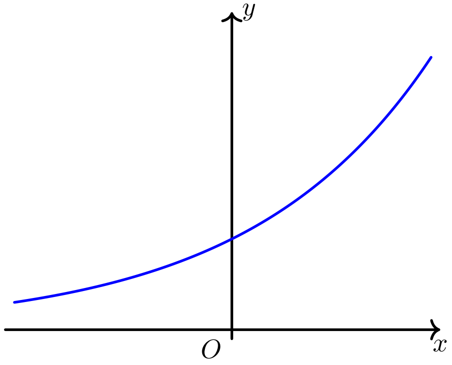
3.1.2　指数関数の値域

底を e とする指数関数

$$y = \exp(x) \tag{2}$$

の定義域は $(-\infty, +\infty)$ であって、値域は $(0, +\infty)$ である。

図 2　e を底とする指数関数



3.2　2 次導関数

関数 $f(x)$ の導関数 $f'(x)$ が更に微分可能であれば、 $f'(x)$ の導関数を求めることができる。 $f'(x)$ の導関数を

$$f''(x), \quad \frac{d^2 f}{dx^2}, \quad \Big(f(x)\Big)'' \tag{3}$$

などの記号であらわす。$f''(x)$ を $f(x)$ の『2 次導関数』あるいは『2 階導関数』という。

定義

関数 $f(x)$ が 2 回微分可能であるとする。 2 次導関数を次式で定義する。

$$f''(x) = \lim_{\Delta x \to 0} \frac{f'(x+\Delta x) - f'(x)}{\Delta x} \tag{4}$$

$f''(x)$ の $x = a$ における値を $f''(a)$ とかく。$f''(a)$ を $f(x)$ の $x = a$ における『2 次微分係数』あるいは『2 階微分係数』という。

問題 V–3 – 1

以下の関数を偏微分しなさい。

$$f(\alpha,\beta)=\frac{1}{1+\exp\big(-(\alpha+\beta x_i)\big)}$$

問題 V–3 – 2

以下の関数を偏微分しなさい。

$$f(\alpha,\beta)=\frac{1}{1+\exp(\alpha+\beta x_i)}$$

解例 V–3 – 1

$$f(\alpha,\beta)=\frac{1}{1+\exp(-\alpha-\beta x_i)}$$

において

$$u=-\alpha-\beta x_i$$

$$v=1+\exp(u)$$

$$f=v^{-1}$$

と置く。

それぞれ微分すると

$$\frac{\partial u}{\partial \alpha}=-1$$
$$\frac{\partial u}{\partial \beta}=-x_i$$
$$\frac{dv}{du}=\exp(u)$$
$$\frac{df}{dv}=-v^{-2}=-\frac{1}{v^2}$$

合成関数の微分より

$$
\begin{aligned}
\frac{\partial f}{\partial \alpha} &= \frac{df}{dv}\frac{dv}{du}\frac{\partial u}{\partial \alpha} \\
&= \left(-\frac{1}{v^2}\right)(\exp(u))\,(-1) \\
&= \frac{\exp\big(-(\alpha+\beta x_i)\big)}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)^2} \\
&= \frac{1}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}
\end{aligned}
$$

同様に

$$
\begin{aligned}
\frac{\partial f}{\partial \beta} &= \frac{df}{dv}\frac{dv}{du}\frac{\partial u}{\partial \beta} \\
&= \left(-\frac{1}{v^2}\right)(\exp(u))(-x_i) \\
&= \frac{x_i\ \exp\big(-(\alpha+\beta x_i)\big)}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)^2} \\
&= \frac{x_i}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}
\end{aligned}
$$

解例 V–3 – 2

$$f(\alpha, \beta) = \frac{1}{1 + \exp(\alpha + \beta x_i)}$$

において

$$u = \alpha + \beta x_i$$

$$v = 1 + \exp(u)$$

$$f = v^{-1}$$

と置く。

それぞれ微分すると

$$\frac{\partial u}{\partial \alpha} = 1$$
$$\frac{\partial u}{\partial \beta} = x_i$$
$$\frac{dv}{du} = \exp(u)$$
$$\frac{df}{dv} = -v^{-2} = -\frac{1}{v^2}$$

合成関数の微分より

$$\frac{\partial f}{\partial \alpha} = \frac{df}{dv}\frac{dv}{du}\frac{\partial u}{\partial \alpha} = \left(-\frac{1}{v^2}\right)(\exp(u))\,(1) = \frac{-\exp(\alpha+\beta x_i)}{\Big(1+\exp(\alpha+\beta x_i)\Big)^2}$$

$$= \frac{-1}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$

$$\frac{\partial f}{\partial \beta} = \frac{df}{dv}\frac{dv}{du}\frac{\partial u}{\partial \beta} = \left(-\frac{1}{v^2}\right)(\exp(u))\,(x_i) = \frac{(-x_i)\ \exp(\alpha+\beta x_i)}{\Big(1+\exp(\alpha+\beta x_i)\Big)^2}$$

$$= \frac{-x_i}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$

問題 V–3 – 3

以下の関数を偏微分しなさい。

$$g(\alpha,\beta)=\ln\left(\frac{1}{1+\exp\big(-(\alpha+\beta x_i)\big)}\right)$$

問題 V–3 – 4

以下の関数を偏微分しなさい。

$$g(\alpha,\beta)=\ln\left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right)$$

解例 V–3 – 3

$$f(\alpha,\beta)=\frac{1}{1+\exp(-(\alpha+\beta x_i))}$$

とおく。問題 V–3 – 1 より、

$$\frac{\partial f}{\partial \alpha}=\frac{1}{\Big(1+\exp(-(\alpha+\beta x_i))\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$
$$\frac{\partial f}{\partial \beta}=\frac{x_i}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$

対数微分法により、

$$
\begin{aligned}
\frac{\partial g}{\partial \alpha} = \frac{\partial}{\partial \alpha} \ln\left(f(\alpha,\beta)\right) &= \frac{\left(\frac{1}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))}\right)}{\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right)} \\
&= \frac{1}{1+\exp(\alpha+\beta x_i)}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial g}{\partial \beta} = \frac{\partial}{\partial \beta} \ln\left(f(\alpha,\beta)\right) &= \frac{\left(\frac{x_i}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))}\right)}{\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right)} \\
&= \frac{x_i}{1+\exp(\alpha+\beta x_i)}
\end{aligned}
$$

解例 V−3 − 4

$$f(\alpha,\beta)=\frac{1}{1+\exp(\alpha+\beta x_i)}$$

とおく。問題 V−3 − 2 より、

$$\frac{\partial f}{\partial \alpha}=\frac{-1}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$
$$\frac{\partial f}{\partial \beta}=\frac{-x_i}{\Big(1+\exp\big(-(\alpha+\beta x_i)\big)\Big)\Big(1+\exp(\alpha+\beta x_i)\Big)}$$

対数微分法により、

$$\frac{\partial g}{\partial \alpha} = \frac{\partial}{\partial \alpha} \ln\left(f(\alpha, \beta)\right) = \frac{\left(\frac{-1}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))}\right)}{\left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right)}$$

$$= \frac{-1}{1+\exp\left(-(\alpha+\beta x_i)\right)}$$

$$\frac{\partial g}{\partial \beta} = \frac{\partial}{\partial \beta} \ln\left(f(\alpha, \beta)\right) = \frac{\left(\frac{-x_i}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))}\right)}{\left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right)}$$

$$= \frac{-x_i}{1+\exp\left(-(\alpha+\beta x_i)\right)}$$

関数 $f(x_1, x_2)$ が 2 変数関数であって偏微分可能とする。偏導関数を

$$\frac{\partial f}{\partial x_1} \tag{5}$$

$$\frac{\partial f}{\partial x_2} \tag{6}$$

とあらわし 1 次偏導関数とよぶ。

1 次偏導関数がさらに微分可能とする。1 次偏導関数を微分した関数を 2 次偏導関数とよび

$$\frac{\partial^2 f}{\partial {x_1}^2} \tag{7}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} \tag{8}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} \tag{9}$$

$$\frac{\partial^2 f}{\partial {x_2}^2} \tag{10}$$

とあらわす。1 次偏導関数は変数の数だけあり、2 次偏導関数は変数の数の二乗だけある。

3.3　ロジスティックモデルを仮定した対数尤度関数

$\Pr(y=1)$ にロジスティックモデル

$$p_i = \frac{1}{1+\exp\left(-\left(\alpha+\beta x_i\right)\right)} \tag{11}$$

を仮定する。すると $\Pr(y=0)$ は

$$1-p_i = \frac{1}{1+\exp\left(\alpha+\beta x_i\right)} \tag{12}$$

である。

すると尤度 (l_i) は

$$l_i = \left(\frac{1}{1+\exp\left(-\left(\alpha+\beta x_i\right)\right)}\right)^{y_i}\left(\frac{1}{1+\exp\left(\alpha+\beta x_i\right)}\right)^{1-y_i} \tag{13}$$

であり、対数尤度 $(\ln l_i)$ は

$$\ln l_i = y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) + (1-y_i)\ln\left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right) \tag{14}$$

である。対数尤度 $(\ln l_i)$ を $i=1$ から n まで足し合わせたものが対数尤度関数 $(\ln L)$ である。

$$\ln L = \sum_{i=1}^{n}\left(y_i \ln\left(\frac{1}{1+\exp\left(-\left(\alpha+\beta x_i\right)\right)}\right) + (1-y_i)\ln\left(\frac{1}{1+\exp\left(\alpha+\beta x_i\right)}\right)\right) \tag{15}$$

3.4　対数尤度の偏導関数

$y_i = 1$ のときの、(14) を $f(\alpha, \beta)$ とすると、α, β による偏導関数は

$$f(\alpha, \beta) = \ln\left(\frac{1}{1 + \exp\big(-(\alpha + \beta x_i)\big)}\right) \tag{16}$$

$$\frac{\partial f}{\partial \alpha} = \frac{1}{1 + \exp(\alpha + \beta x_i)} \tag{17}$$

$$\frac{\partial f}{\partial \beta} = \frac{x_i}{1 + \exp(\alpha + \beta x_i)} \tag{18}$$

である。この (17) (18) は更に α, β で微分可能なのは明らかである。

(17) を α で偏微分すると

$$\frac{\partial^2 f}{\partial \alpha^2} = \frac{-1}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))} \tag{19}$$

であって、(18) を β で偏微分すると

$$\frac{\partial^2 f}{\partial \beta^2} = \frac{-x_i}{(1+\exp(-(\alpha+\beta x_i)))(1+\exp(\alpha+\beta x_i))} \tag{20}$$

である。

同様に、$y_i = 0$ のときの、(14) を $g(\alpha, \beta)$ とすると、

$$g(\alpha, \beta) = \ln\left(\frac{1}{1+\exp(\alpha+\beta x_i)}\right) \tag{21}$$

であって、偏導関数は

$$\frac{\partial g}{\partial \alpha} = \frac{-1}{1+\exp(-(\alpha+\beta x_i))} \tag{22}$$

$$\frac{\partial g}{\partial \beta} = \frac{-x_i}{1+\exp(-(\alpha+\beta x_i))} \tag{23}$$

である。

そして (22) の α による偏導関数と (23) の β による偏導関数は

$$\frac{\partial^2 g}{\partial \alpha^2} = \frac{-1}{(1 + \exp(-(\alpha + \beta x_i)))\,(1 + \exp(\alpha + \beta x_i))} \tag{24}$$

$$\frac{\partial^2 g}{\partial \beta^2} = \frac{-x_i}{(1 + \exp(-(\alpha + \beta x_i)))\,(1 + \exp(\alpha + \beta x_i))} \tag{25}$$

である。

ロジスティックモデルを仮定した対数尤度関数 $(\ln L)$ は α と β の関数である。そして対数尤度 $(\ln l_i)$ の偏導関数は

$$\frac{\partial}{\partial\alpha}\ln l_i = y_i\frac{1}{1+\exp(\alpha+\beta x_i)} - (1-y_i)\,\frac{1}{1+\exp(-(\alpha+\beta x_i))} \tag{26}$$

$$\frac{\partial}{\partial\beta}\ln l_i = y_i\frac{x_i}{1+\exp(\alpha+\beta x_i)} - (1-y_i)\,\frac{x_i}{1+\exp(-(\alpha+\beta x_i))} \tag{27}$$

である。

3.5　単調関数同士の和

$f(x), g(x)$ は共に狭義単調増加関数とする。$f(x), g(x)$ は狭義単調増加関数なので、その導関数は正である。従って

$$f'(x) > 0 \tag{28}$$

$$g'(x) > 0 \tag{29}$$

である。正値と正値の和は正値なので

$$f'(x) + g'(x) > 0 \tag{30}$$

である。

『導関数の和は和の導関数』なので

$$f'(x) + g'(x) = (f(x) + g(x))' \tag{31}$$

である。(31) 左辺が正ならば右辺も正である。従って

$$(f(x) + g(x))' > 0 \tag{32}$$

が成り立つ。つまり、狭義単調増加関数同士の和は狭義単調増加関数である。

問題 V–3 – 4

$f(x), g(x)$ は共に狭義単調減少関数とする。

$f(x) + g(x)$ が狭義単調減少関数であることを示しなさい。

解例 V–3 – 4

$f(x), g(x)$ は狭義単調減少関数なので、その導関数

$$f'(x) < 0$$

$$g'(x) < 0$$

は負である。
負値と負値の和

$$f'(x) + g'(x) < 0$$

は負値である。

導関数の和は和の導関数

$$f'(x) + g'(x) = (f(x) + g(x))'$$

である。ここで左辺は負なので、右辺も負である。導関数が負なので $f(x) + g(x)$ は狭義単調減少関数である。（終）

定理 V–3 – 1

狭義単調増加関数の和は狭義単調増加関数であり、狭義単調減少関数の和は狭義単調減少関数である。

3.6　対数尤度の和

$0 < \sum_{i=1}^{n} y_i < n$ とし、α に関する単峰性を示す。

3.6.1　$y_i = 1$ の対数尤度の和

$y_i = 1$ の対数尤度の和を $\ln L_1$ とあらわす。すると (15) の右辺シグマ括弧内第二項は 0 になるので

$$\ln L_1 = \sum_{i=1}^{n} y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \tag{33}$$

である。

(33) 右辺真数は

$$0 < \frac{1}{1 + \exp(-(\alpha + \beta x_i))} < 1 \tag{34}$$

なので

$$\ln\left(\frac{1}{1 + \exp(-(\alpha + \beta x_i))}\right) < 0 \tag{35}$$

である。

仮定により $\sum_{i=1}^{n} y_i > 0$ なので、

$$\sum_{i=1}^{n} y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) < 0 \tag{36}$$

である。負なので上界を持つ。上界を持つから上限をもつ。

また、

$$\ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \tag{37}$$

は α に関して狭義単調増加関数なので、その和である

$$\sum_{i=1}^{n} y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \tag{38}$$

も狭義単調増加関数である。

そして (33) の $\alpha \to +\infty$ の時の極限は

$$\begin{aligned}
\alpha \to +\infty \quad &\Longrightarrow \quad \frac{1}{1+\exp(-(\alpha+\beta x_i))} \to 1 &(39)\\
&\Longrightarrow \quad \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \to 0 &(40)\\
&\Longrightarrow \quad \sum_{i=1}^{n} y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \to 0 &(41)
\end{aligned}$$

である。

参考

$$\alpha \to -\infty \quad \Longrightarrow \quad \sum_{i=1}^{n} y_i \ln\left(\frac{1}{1+\exp(-(\alpha+\beta x_i))}\right) \to -\infty \tag{42}$$

(33) の α に関する偏導関数

$$\frac{\partial}{\partial \alpha} \ln L_1 = \sum_{i=1}^{n} y_i \frac{1}{1 + \exp(\alpha + \beta x_i)} \tag{43}$$

は正なので (33) は狭義単調増加関数である。

そして (43) の α に関する偏導関数 (2 次偏導関数)

$$\frac{\partial^2}{\partial \alpha^2} L_1 = -\sum_{i=1}^{n} y_i \frac{1}{(1 + \exp(-(\alpha + \beta x_i)))(1 + \exp(\alpha + \beta x_i))} \tag{44}$$

は負なので、α が増加するに従い、(33) の変化量は減少する。

3.6.2 $y_i = 0$ の対数尤度の和

$y = 0$ の対数尤度の和を $\ln L_0$ とあらわすと

$$\ln L_0 = \sum_{i=1}^{n} (1 - y_i) \ln \left(\frac{1}{1 + \exp(\alpha + \beta x_i)} \right) \tag{45}$$

である。(45) の極限は

$$\alpha \to +\infty \quad \Longrightarrow \quad \ln L_0 \to -\infty \tag{46}$$

$$\alpha \to -\infty \quad \Longrightarrow \quad \ln L_0 \to 0 \tag{47}$$

である。

(45) の α による偏導関数は

$$\frac{\partial}{\partial \alpha} \ln L_0 = -\sum_{i=1}^{n} (1 - y_i) \frac{1}{1 + \exp(-(\alpha + \beta x_i))} \tag{48}$$

である。偏導関数 (48) が負なので (45) は狭義単調減少関数である。

α による 2 次偏導関数

$$\frac{\partial^2}{\partial \alpha^2} \ln L_0 = -\sum_{i=1}^{n} (1 - y_i) \frac{1}{(1 + \exp(-(\alpha + \beta x_i)))(1 + \exp(\alpha + \beta x_i))} \tag{49}$$

が負なので α が増加するに伴い、$\ln L_0$ の変化量の絶対値は増加していく。

3.6.3 $\ln L_1$ と $\ln L_0$ の和

(33) と (45) の和が対数尤度関数 $\ln L$ である。

$$\ln L = \ln L_1 + \ln L_0 \tag{50}$$

$\ln L_0,\ \ln L_1 < 0$ なので、$\ln L < 0$ である。負値なので上界を持つ。上界があるから上限を持つ。

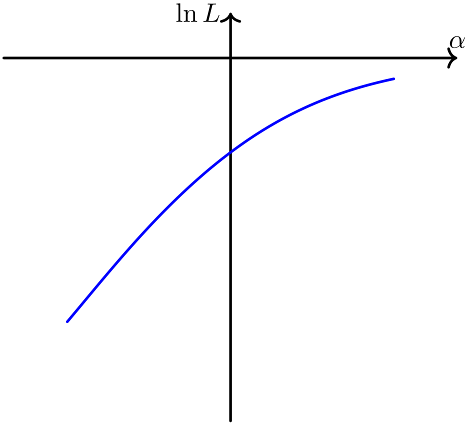
(50) を α で偏微分すると

$$\frac{\partial}{\partial \alpha} \ln L = \frac{\partial}{\partial \alpha} \ln L_1 + \frac{\partial}{\partial \alpha} \ln L_0 \tag{51}$$

ここで $\ln L_1$ は

- $\ln L_1 < 0$
- α に関して狭義単調増加関数
- $\sup(\ln L_1) = 0$
- $\ln L_1$ の増加量は α の増加に従い減少

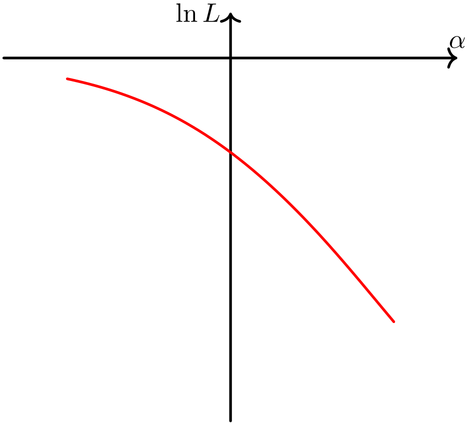
図 3　$\ln L_1$ の形状



また $\ln L_0$ は

- $\ln L_0 < 0$
- α に関して狭義単調減少関数
- $\sup(\ln L_0) = 0$
- $\ln L_0$ の減少量の絶対値は α の増加に従い増加

図 4　$\ln L_0$ の形状



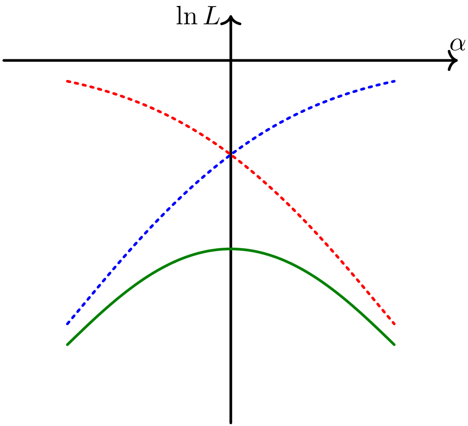
従って、$\ln L$ は α 軸において単峰であって、

$$\frac{\partial}{\partial \alpha} \ln L_1 + \frac{\partial}{\partial \alpha} \ln L_0 = 0 \tag{52}$$

のとき最大値をとる[1)]。

1）　図では $\ln L_1$ と $\ln L_0$ を対称に描いているが、通常対称ではない。あくまでもイメージ図として表現している。

図 5　$\ln L$ の形状



3.6.4 $\ln L$ の β における単峰性

対数尤度の β による偏導関数 (27) の符号は、y_i の値と x_1 の符号により影響を受ける。
$\{y=1, x_i>0\}$ および $\{y=0, x_i<0\}$ の場合、

$$\frac{\partial}{\partial\beta}\ln l_i > 0 \tag{53}$$

であって、$\{y=1, x_i<0\}$ および $\{y=0, x_i>0\}$ の場合、

$$\frac{\partial}{\partial\beta}\ln l_i < 0 \tag{54}$$

である。それぞれの場合において α と同様の議論により $\ln L$ が β 軸において単峰であって、最大値をとることを示すことができる。

3.6.5 $\ln L$ の単峰性

従って、$\ln L$ が α 軸、 β 軸それぞれにおいて単峰であって、最大値をとる。

ただし、以下の場合最尤推定値は発散する。

- $\sum_{i=1}^{n} y_i = n$
- $\sum_{i=1}^{n} y_i = 0$
- 完全分離
- 準完全分離

3.7　まとめ

- ロジスティックモデルを仮定した対数尤度関数は、単峰であって最大値を持つ。
- ロジスティックモデルを仮定した尤度関数は、単峰であって最大値を持つ。
- ただし、完全分離、準完全分離等の場合は除く。