

非線形回帰分析入門
—生物実験データ解析の基礎—

第 16 回 高橋セミナー

2004 年 2 月 7 日

高橋 行雄

10/31/2005 7:33 PM

表紙裏

2004年1月19日, 新規作成,

ファイル名 : G:¥R3_高橋_2004_から¥R_ANZ_高橋 seminar_16_非線形回帰入門¥Semi_16_非線形回帰.doc

非線形回帰分析入門

目 次

1. はじめに	1
2. データ例示	4
2.1. 用いるデータの選択.....	4
2.2. 細胞毒性試験のデータ	5
2.3. 反応率にした細胞毒性試験データ	6
3. 解析方法	8
3.1. シグモイド曲線.....	8
3.2. 反復計算の考え方.....	9
3.3. 線形化法の計算手順.....	10
3.4. 計算例.....	13
3.5. 解説, 実用上の諸問題.....	19
4. 統計ソフトによる検証	22
4.1. SASのNLINプロシジャによる計算.....	22
SASのNLINプログラム	22
SASのNLINによる偏微分式の出力	22
SASのNLINによる繰り返し計算.....	23
4.2. JMPのスク립ト言語による繰り返し計算	24
JMPのスク립トによるプログラム.....	24
JMPのスク립トによる計算結果.....	24
ロジスティック曲線の95%信頼区間.....	25
4.3. JMPの非線形プログラムによる計算	25
4.4. Excelによる繰り返し計算.....	27
第1反復.....	27
第2反復.....	28
信頼区間の計算.....	30

5. Draper (3rd edition) の 24 章本を基にした入門.....	31
5.1. 非線形モデルのための最小 2 乗法.....	31
Introduction.....	31
非線形モデル.....	31
非線形の場合の最小 2 乗法.....	32
5.2. 非線形系のパラメータ推定.....	34
線形化の幾何学的な解釈 (3 版での追加)	38
最急勾配法.....	41
マルカート法.....	41
Gridst and Plots.....	42
The Importance of Good Starting Value	42
Getting Initial Estimation θ_0	42
5.3. AN EXAMPLE.....	42
A Solution Through the Normal Equations.....	44
A Solution Through the Linearization Technique.....	44
Further Analysis	48
Confidence Regions	49
Some Typical Nonlinear Program Output Features	49
5.4. ガウス・ニュートン法による逐次計算, ドレーパー・スミスの事例.....	52
JMPのスク립ト言語での結果.....	52
JMPの非線形モデルでの結果	53
SAS/NLINでの結果.....	55
6. 光毒性試験データについてのJMPによる追加解析	56
6.1. ロジスティック曲線の推定値の 95%信頼区間	56
SASによる残差 2 乗和のグリッド計算.....	59
6.2. 第 1 反復での線形化によるパラメータの推定.....	59
SASによる線形化による最小 2 乗解.....	60
6.3. パラメータの信頼区間.....	61

図 表 目 次

図 2.1	光細胞毒性試験の吸光度の用量反応関係	6
図 2.2	反応率に対するシグモイド曲線のあてはめ	7
図 3.1	初期値 $\beta_{10} = -1.1$ と $\beta_{20} = 3.0$ のシグモイド曲線 (a) と残差プロット(b).....	9
図 3.2	残差の2乗和の自然対数の等高線図	10
図 3.3	シグモイド曲線の95%信頼区間.....	20
図 4.1	JMPの非線形プログラムによるあてはめ	26
図 0.1	残差平方和Sの等高線図	53
図 0.2	パラメータの信頼区間の等高線	54
表 2.1	化合物Aの光細胞毒性試験の吸光度と反応率.....	6
表 2.2	光の照射ある場合の化合物Aの細胞毒性の反応率.....	7
表 3.1	第1回目の反復計算表	14
表 3.2	第2回目の反復計算表	16
表 3.3	第3回目の反復計算表	16
表 3.4	パラメータの信頼区間の計算シート	19
表 5.1	Standard Notations for Linear and Nonlinear Least Squares.....	32

1. はじめに

多くの化学物質の *in vitro* および *in vivo* の実験データの解析に非線形回帰分析が中心的な役割を果たしていることが、これまでのセミナーを通じて再認識した。薬物動態学におけるコンパートメント解析では非線形回帰分析が古くから用いられており、さらに現在のポピュレーション PK では非線形回帰分析に変量効果 (random effects) を含めた解析のために非線形混合モデルが適用されるようになってきた。

しかしながら、他の生物実験の分野では、実験結果が非線形となる現象を各種の変数変換を用いて線形化し、線形回帰分析により各種のパラメータの推定をおこなう努力がされ、非線形のままでのデータ解析を目にすることは稀である。

Peace, K.E. 編 (1988), 中上節夫・森川俊彦監訳 (1992), 「医薬統計学—医薬品開発のための統計学」に医薬品の創薬から市販後の安全性・製造と品質管理まで幅広い場面での統計的な話題が豊富にある。この本に含まれている内容で、本質的に非線形の課題を拾い出してみよう。2章の「薬物の創製」であるが、III節の「標的を定めたスクリーニング」は、ほとんどが非線形の実験結果の線形化する方法論で占められている。この線形化の限界については、2003年の統計連合大会で、馬場淳、高橋行雄で、「シグモイド型反応に関する古典的な線形化解析手法」で示した。

3章の「薬理学的活性の評価」であるが、III節の「生物検定法、薬物および用量」、V節の「逆推定、フィラーの定理、ジャックナイフ法」、VI節の「用量—反応関係：量的な反応についての中央量の推定」、VII節の「相対力価の推定」、などは、第14回および第15回のセミナーで、非線形回帰分析を用いた比の推定によって全て取り扱えることを示した。同じく3章のXIVの「薬物—受容体の相互作用」は、2003年の計量生物学会で、私（高橋行雄）が「非線形回帰モデルを用いた EC_{50} の比較」で示した課題である。

5章の前臨床での安全性試験のVI節の「*in vitro* 試験」で取り上げられている、復帰突然変異試験 (Ames 試験) は、第15回のセミナーのテーマであった。これは、2つの化学物質に対する変異コロニー数の比活性値の比較であるが、古典的な生物検定法の効力比とその信頼区間を非線形回帰モデルで直接求める方法を示した。

平均値の比の信頼区間を出すためには、比の分散が必要である。平均値の差の分散は、多くの統計の教科書で取り上げられているが、比の分散については稀である。医薬系の統計学の教科書である、Armitage・Berry 著、椿美智子・椿広計訳 (2001), 「医学研

究のための統計的方法」の 83 ページに 2 つの平均値の比、積、および、一般の関数についての分散の求め方が示されている。しかしながら、第 17 章の「生物検定法」では、平行線検定法、勾配比検定法などの伝統的な生物検定法の計算公式が示されているのみである

椿美智子・椿広計訳 (2001) は第 3 版の訳であり、原著は、2002 年に第 4 版 *Statistical Methods in Medical Research* が、Armitage, P., Berry, G. and Matthews, J.N.S. によって出版された。第 3 版は、医学研究者向けの統計入門書であったが、第 4 版は、医学研究に関係する試験統計家 (Biostatistician) 向けに版あらたに起こす必要性があつたが前書きに記されている。

この第 4 版の 12.4 節に「非線形回帰, Non-linear regression」が新たに登場し、*in vitro* 試験での非線形関係の実験データを線形化して線形回帰分析した場合の問題点が述べられている。さらに、20 章は「Laboratory assay」と題し Biological assay は 20.1 節に移され、20.5 節に「Some special assay」として再度、線形化の問題点が示されている。しかしながら、この本は、非線形回帰分析の *in vitro* および *in vivo* 試験での適用上の注意になっており、非線形回帰分析の入門ではない。

本質的に非線形の課題に対して、線形化の方法ではなく非線形のままの解析が有益であることと、逆に、線形の問題でも、推定したいパラメータが、線形回帰分析のパラメータの比であるような場合に、その比について非線形化することにより信頼区間を直接推定できることを、これまでのセミナーで示してきた。多くの化学物質の *in vitro* および *in vivo* の実験データの標準的な解析に非線形回帰分析を適用することは、実験結果を統計的にも簡潔に表現するために不可欠と認識した。

ところが、非線形回帰分析についての統計の入門書は非常に少ない。古くは、Draper, N.R. and Smith, H. (1966) 著、中村慶一訳 (1967) 「応用回帰分析」の 10 章に「非線形推定序説」がある。最近では、芳賀俊郎・野澤昌弘・岸本淳順司 (1966) 「SAS による回帰分析」の 8 章の「非線形回帰分析」ぐらいで、他に見当たらない。前者の日本語版は絶版になっているが、英語版は、1998 年に第 3 版「Applied Regression Analysis」が出版されていて 24 章に「An Introduction to Nonlinear Estimation」がある。第 2 版に対して加筆がかなりされている。芳賀らの本は、SAS を用いた非線形回帰分析の使い方の手ほどきであり、Draper, N. R. and Smith, H. (1998) が非線形回帰分析を詳しく知るための入門書として適している。

これまでのセミナーで、Draper, N. R. and Smith, H. (1998) が取り上げている例題について、JMP スクリプト言語の行列表記により計算プロセスを示したが、「入門」ではない。そこで、今回は、細胞毒性のデータをもちいて、P 本の原稿となるような「非線形回帰分析入門」を示すことにした。

2. データ例示

2.1. 用いるデータの選択

これまでのセミナーで計量値のシグモイド曲線の例示に、平滑筋の収縮反応データ、および、環境ホルモンの子宮重量に与える実験データを用いてきた。これらの実験データは、医薬安全研の Q&A で取り上げてきたものであるが、P 本の改訂版に用いるデータとしては幾つかの難点があった。そこで、新たな例示用のデータを示し、これを用いて P 本の原稿で必要となる非線形回帰の計算手順を必要かつ十分に、なおかつコンパクトにする努力をしてみた。

用いるデータは、細胞毒性試験データが適していると判断した。これは、医薬品のみならず、化学物質の安全性、医療器具の安全性など実験対象の幅が広いことである。また、細胞毒性試験は、細胞の生死の量を吸光度で測定したり、直接コロニー数をカウントしたりして、*in-vivo* の実験系に比べ小規模な実験系でもある。

その中で、光の照射によって毒性が増強されるかを検討する光毒性試験データが、様々な議論がしやすい事例である。これは、基本的に光の照射の有無による 2 つのシグモイド曲線の比較のみならず、陰性対象、陽性対象（ブランク）などの取り扱いの課題もあり、基準化して単純化することも、生データのままである取り扱いにも言及でき、各種の解析モデルの説明がしやすい。

これまでの平滑筋のデータは、投与量に対してデータが互いに独立でないとの弱点もあり、例題としては使い難い。この例は、受容体モデルと含めて解説のときに使うとよいだろう。

環境ホルモンの子宮重量に与える実験データは、社会的に興味を引くと思われるが、実験の関係者は細胞毒性に比べると少ないと思われる。D10 の推定などの必要性を解説するときに使うとよいだろう。

2.2. 細胞毒性試験のデータ

表 2.1 に光細胞毒性試験の実験データを示す。このデータは、大森（2003：統計連合大会）の抄録にある図から、用量反応関係の基本は変えずに、吸光度を他の実験データを参考の計算し直したものである。また、事例として取り扱いやすいようにデータを変換したものである。

ある物質 A を調整して光の照射をしないプレート上のウェル（穴）に 8 用量， $2^3 \sim 2^{10}$ mg/mL（単位は要確認），光を照射するプレートでも 8 用量とするが， $2^0 \sim 2^7$ mg/mL と濃度を 3 段階低く設定し，それぞれ 2 回の繰り返しとする。陰性対照は，，光の照射（具体的に光の波長，強さ，時間などを書く）によって影響がない細胞を含む溶液とする。陰性対照は，細胞を含まない溶液のみとする。

ある一定時間経過後に，プレート上のそれぞれのウェルで物質Aの濃度によって死亡細胞が変化する。生存した細胞に色のつく処理を施すと，死亡した細胞が多くなるウェル内の溶液の濁りが澄んで光が透過しやすくなる。微弱な測定用の光をそれぞれのウェルにあてて，ウェルから出てくる光の量と入れた光の量との差を測定する。溶液が濁っていると入れた光は拡散し出てくる光が見かけ上少なる。言い換えると光が吸収されてしまったようになり，この相対的な大きさが 0 から 2（要確認）ぐらいになるように設定し，それぞれのウェルの吸光度を測定する。これが，表 2.1 の吸光度の欄のデータである。

吸光度が実際の測定データであるが，一般的に陰性対照を 100%，陽性対照を 0%となるような反応率にして毒性の強さを基準化されている。反応率が高ければ毒性がなく，低ければ毒性が強いと判定する。反応率は相対的な毒性の強さなので，100%を超えることもあれば，0%以下になることもある。これは，光照射がある場合の反応率で 2^0 と 2^1 mg/mL 溶液に 109.9%と 102.2%が 100%を超えるデータ例である。

反応率は，2つのプレートの陰性対照，および，陽性対照の吸光度データをプールし，陰性対照の場合は 8 個のデータの平均 1.216 を 100%，陽性対照の場合は，8 個のデータの平均 0.052 を 0%となるように算出した。

光を照射するプレートとしないプレートそれぞれに陰性，および陽性対照があるので，別々に反応率を計算するほうが合理的と思われるかもしれないが，統計的には問題である。光の照射によって毒性が増強または減弱し，陰性対照の吸光度が異なる場合のときに，別々

に反応率を計算すると、被験物質の毒性と光照射による毒性との相互作用を無視して解析することになる。光照射したプレート上の陽性対象の吸光度が統計的にも異なるならば、光照射しないプレートの陽性対象の吸光度のみを用いて全ての吸光度データの反応率を算出すべきである。

表 2.1 化合物 A の光細胞毒性試験の吸光度と反応率

濃度	光：なし		濃度	光：あり		
	吸光度：y	反応率：%		吸光度：y	反応率：%	
陰性 対照	1.177 1.196	1.146 1.247	陰性 対照	1.194 1.270	1.330 1.170	—
2 ³	1.097	1.018	2 ⁰	1.331	1.084	109.9 88.7
2 ⁴	0.795	0.886	2 ¹	1.250	1.062	102.9 86.8
2 ⁵	0.731	0.701	2 ²	0.810	1.022	65.1 83.3
2 ⁶	0.409	0.426	2 ³	0.566	0.533	44.2 41.3
2 ⁷	0.217	0.220	2 ⁴	0.268	0.293	18.6 20.7
2 ⁸	0.178	0.156	2 ⁵	0.119	0.102	5.8 4.3
2 ⁹	0.122	0.133	2 ⁶	0.092	0.066	3.4 1.2
2 ¹⁰	0.135	0.111	2 ⁷	0.126	0.142	6.4 7.7
陽性 対照	0.041 0.016	0.054 0.075	陽性 対照	0.040 0.109	0.064 0.092	—

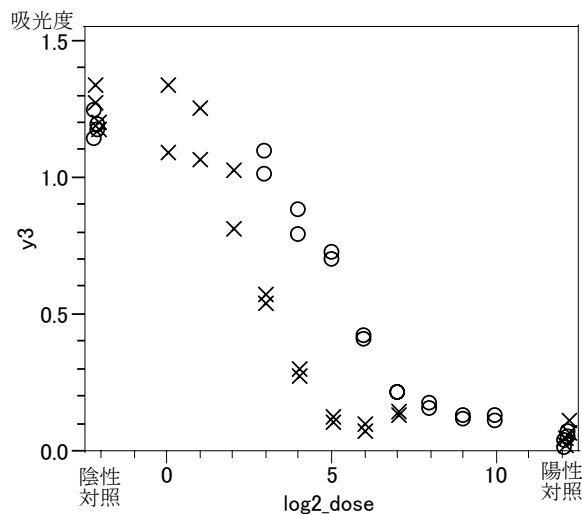


図 2.1 光細胞毒性試験の吸光度の用量反応関係

○：光の照射なし，×：光の照射あり

2.3. 反応率にした細胞毒性試験データ

表 2.1 の光毒性試験データで、光を照射した場合のデータについて非線形回帰分析

により毒性の発現が 50%となるような化合物Aの濃度D50, およびD90 を計算する手順を示す. 計算の手順を簡潔にするために陰性対照を 100%, 陽性対照を 0%とし, 吸光度を反応率としたデータを用いる.

表 2.2 光の照射ある場合の化合物 A の細胞毒性の反応率

濃度	反応率 : %	
2 ⁰	109.9	88.7
2 ¹	102.9	86.8
2 ²	65.1	83.3
2 ³	44.2	41.3
2 ⁴	18.6	20.7
2 ⁵	5.8	4.3
2 ⁶	3.4	1.2
2 ⁷	6.4	7.7

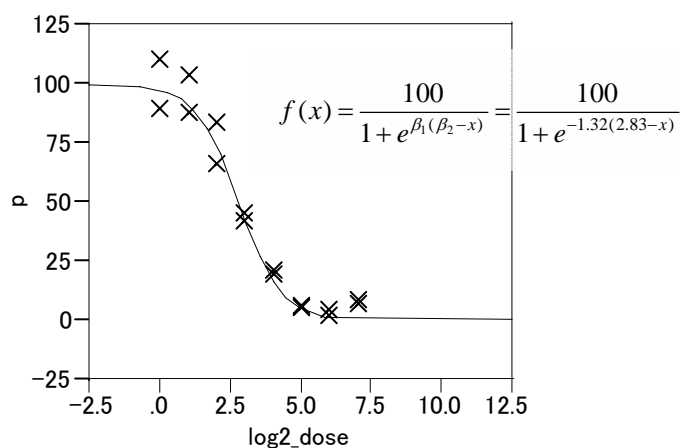


図 2.2 反応率に対するシグモイド曲線のあてはめ

3. 解析方法

非線形回帰分析を行いためには、逐次計算が必要である。実際には、統計ソフトを使うことが前提であるが、計算の考え方とその方法は、統計ソフトを適切に使うために必要である。このために、非線形回帰分析の標準的な計算方法であるガウス・ニュートン法の繰り返し手順を、計算の途中結果を示しつつ解説する。ガウス・ニュートン法以外の方法については解説の項で若干の説明をする。

3.1. シグモイド曲線

反応率にロジスティック曲線をあてはめる場合には、上限が 100 および下限が 0 に収束するようなシグモイド曲線は、 $x = \log(\text{dose})$ としたときに、

$$Y = \frac{100}{1 + e^{\beta_1(\beta_2 - x)}} + \varepsilon \quad (1)$$

となる。ここで、 β_1 は曲線の傾き、 β_2 は $D50$ を表すパラメータである。なお、用量 dose の対数は常用対数、自然対数、底が 2 の対数でもよい。用量を $\sqrt{10}$ 倍で設定したときは、底を $\sqrt{10}$ とする x が整数となり用量で端数が出ないので扱いやすくなる。式 (1) は、関数表記を用いて、

$$Y = f(x) + \varepsilon$$

のように書ける。また、 $E(\varepsilon) = 0$ と仮定するならば

$$E(Y) = f(x) \quad (2)$$

となり、誤差が互いに無相関で、分散が $V(\varepsilon) = \sigma^2$ の正規分布 $\varepsilon \sim N(0, \sigma^2)$ に従うとする。

n 個の観測値が $i = 1, 2, \dots, n$, に対して、式 (1) は、

$$Y_i = \frac{100}{1 + e^{\beta_1(\beta_2 - x_i)}} + \varepsilon_i$$

となり、

$$Y_i = f(x_i; \beta_1, \beta_2) + \varepsilon_i \quad (3)$$

と書くこともできる。さらに簡略化して、

$$Y_i = f(x_i) + \varepsilon_i \quad (4)$$

とする。非線形モデルおよび与えられたデータに対して残差の 2 乗和を次のように定義する。

$$S_T(\beta_1, \beta_2) = \sum_{i=1}^n \{Y_i - f(x_i; \beta_1, \beta_2)\}^2 \quad (5)$$

$f(x)$ が線形であれば、回帰分析の手順に従って、切片と傾きの最小 2 乗推定量を求めることができる。しかし、 $f(x)$ が非線形の場合には、最小 2 乗推定量 $\hat{\beta}_1$ と $\hat{\beta}_2$ を見出すためには、式 (5) を β_1 と β_2 について偏微分した 2 つの正規方程式について反復法を用いて解くことになる。いくつかの反復法の中で SAS などの統計ソフトで標準的に用いられているのがテーラー展開法を用いた線型化法の方法で、ガウス・ニュートン法ともいわれている。これは、一回一回の反復計算の過程で線形回帰分析を使って非線形の問題を解くことから線形化法といわれている。

3.2. 反復計算の考え方

$\beta_1^{(1)}, \beta_2^{(1)}$ を、最終的に推定したいパラメータ β_1, β_2 の初期値としよう。これらの初期値は、データをプロットし、シグモイド曲線を描くことにより “大体正しい” と示唆された値とする。 $\beta_2^{(1)}$ は、反応率が 50% 前後の対数用量であるので、おおよそ 3 と推定できる。 $\beta_1^{(1)}$ は、シグモイド曲線の傾きであるが、おおよその目安は、反応率が 50% の対数用量から対数用量で 1 単位増加した反応率を y_0 としたときに $\log_e\{y_0/(100-y_0)\}$ で求められる。 y_0 はおおよそ 25% であるので、 $\hat{\beta}_1^{(1)} = \log_e\{25/(100-25)\} = -1.1$ となる。これは、式 (1) の x に $\beta_2 + 1$ を代入して β_1 について解いた結果である。

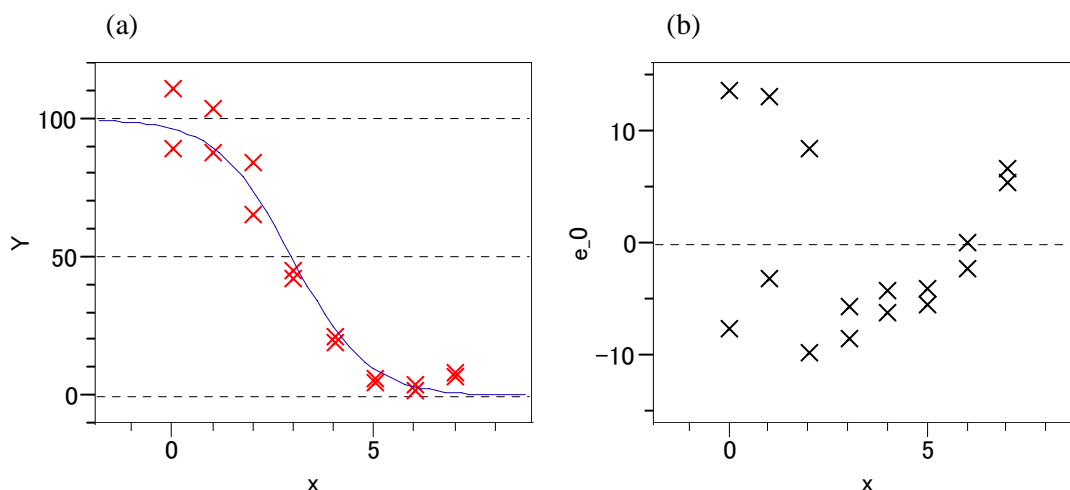


図 3.1 初期値 $\beta_{10} = -1.1$ と $\beta_{20} = 3.0$ のシグモイド曲線 (a) と残差プロット (b)

$x = 3 \sim 5$ 残差が全てマイナスとなりあてはめが悪い

式 (5) に初期値を代入して残差の 2 乗和を求めると $S_T(\hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)}) = 876.63$ となる。初期値 $\beta_1^{(1)}, \beta_2^{(1)}$ にある値を加え残差の 2 乗和が小さくなるような $\delta_1^{(1)}, \delta_2^{(1)}$ を線型化法の方法で計算し、第 2 の反復のためのパラメータとして $\beta_1^{(2)} = \beta_1^{(1)} + \delta_1^{(1)}$ 、および $\beta_2^{(2)} = \beta_2^{(1)} + \delta_2^{(1)}$ を計算する。 $\delta_1^{(1)}, \delta_2^{(1)}$ は、線形化法によって、 -0.172 と -0.171 が計算されるので、 $\hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)}$ は、 -1.272 と 2.829 になる。残差の 2 乗和は、 $S_T(\hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)}) = 664.56$ となり $S_T(\hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})$ に比べ 201.07 減少する。

第 2 回目の反復では、 $\hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)}$ を使って線型化法の方法で $\hat{\delta}_1^{(2)} = -0.049$ 、 $\hat{\delta}_2^{(2)} = -0.005$ が計算され、 $\hat{\beta}_1^{(3)}, \hat{\beta}_2^{(3)}$ は、それぞれ -1.321 と、 2.834 となる。残差の 2 乗和 $S_T(\hat{\beta}_1^{(3)}, \hat{\beta}_2^{(3)}) = 659.60$ と $S_T(\hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)})$ に比べて 4.96 減少している。このような反復を繰り返して、推定したい残差の 2 乗和の減少、あるいは、推定したいパラメータの変化が、あらかじめ定めた小さな値よりも小さくなったら反復をやめる。第 3 回目で反復が終わったとしたときパラメータの推定値は、 $\hat{\beta}_1 = \hat{\beta}_1^{(3)} = -1.322$ 、 $\hat{\beta}_2 = \hat{\beta}_2^{(3)} = 2.832$ となる。この反復の過程を残差の 2 乗和を 図 3.2 に β_1 を X 軸 β_2 を Y 軸とした等高線図に示す。

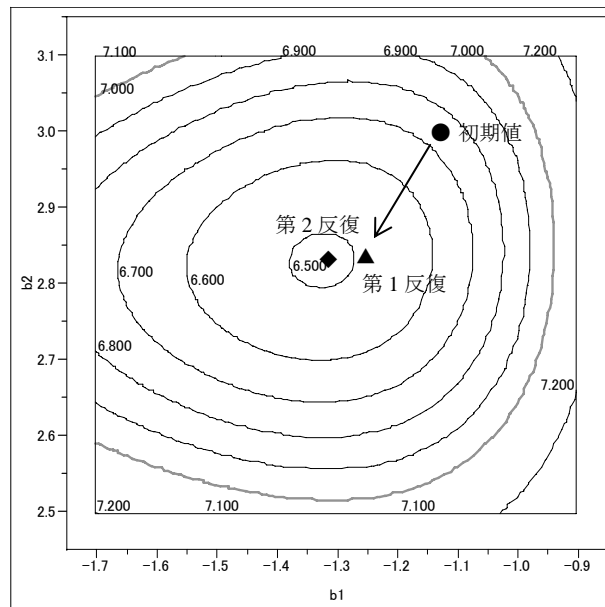


図 3.2 残差の 2 乗和の自然対数の等高線図

第 3 反復は、第 2 反復と図上では重なっているので省略した

3.3. 線形化法の計算手順

シグモイド曲線

$$f(x) = \frac{100}{1 + e^{\beta_1(\beta_2 - x)}} \quad (6)$$

をパラメータ β_1 と β_2 で偏微分すると次の結果をえる.

$$\frac{\partial y}{\partial \beta_1} = \frac{-(\beta_2 - x) e^\eta f(x)}{1 + e^\eta} \quad (7)$$

$$\frac{\partial y}{\partial \beta_2} = \frac{-\beta_1 e^\eta f(x)}{1 + e^\eta} \quad (8)$$

ただし, $\eta = \beta_1(\beta_2 - x)$.

【手順 1】

パラメータの初期値 $\hat{\beta}_1^{(1)} = -1.1$ と $\hat{\beta}_2^{(1)} = 3.0$ を式 (6) に代入し, n 個分のデータについて予測値 $f_i^{(1)}(x_i; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})$ を計算し, 反応率 y_j との残差 $\hat{e}_i^{(1)}$ を計算する.

$$\hat{e}_i^{(1)} = y_i - f_i^{(1)}(x_i; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)}) = y_i - \frac{100}{1 + \exp\{\hat{\beta}_1^{(1)}(\hat{\beta}_2^{(1)} - x_i)\}} \quad (9)$$

【手順 2】

パラメータの初期値 $\hat{\beta}_1^{(1)} = -1.1$ と $\hat{\beta}_2^{(1)} = 3.0$ を式 (7) と (8) に代入し, n 個分のデータについて微分係数 $\hat{z}_{1i}^{(1)}$, $\hat{z}_{2i}^{(1)}$ を計算する.

$$\hat{z}_{1i}^{(1)} = \frac{-(\hat{\beta}_2^{(1)} - x_i) e^{\hat{\eta}_i^{(1)}} f(x_i; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_i^{(1)}}} \quad (10)$$

$$\hat{z}_{2i}^{(1)} = \frac{-\hat{\beta}_1^{(1)} e^{\hat{\eta}_i^{(1)}} f(x_i; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_i^{(1)}}} \quad (11)$$

ただし, $\hat{\eta}_i^{(1)} = \hat{\beta}_1^{(1)}(\hat{\beta}_2^{(1)} - x_i)$, $i = 1, 2, \dots, n$

【手順 3】

手順 1 で求めた残差 $\hat{e}_i^{(1)}$ を従属変数 $e_i^{(1)}$ とし, 手順 2 で求めた微分係数 $\hat{z}_{1i}^{(1)}$, $\hat{z}_{2i}^{(1)}$ を独立変数 $z_{1i}^{(1)}$, $z_{2i}^{(1)}$ として, 次の切片なしの線形回帰式をたてる.

$$e_i^{(1)} = \delta_1^{(1)} z_{1i}^{(1)} + \delta_2^{(1)} z_{2i}^{(1)} + \varepsilon_i \quad i = 1, 2, \dots, n$$

【手順 4】

手順 3 の線形回帰式のパラメータ $\delta_1^{(1)}$ と $\delta_2^{(1)}$ を, 切片なしの回帰分析を用いて, $\hat{\delta}_1^{(1)}$ と $\hat{\delta}_2^{(1)}$ を推定する.

【手順 5】

シグモイド曲線のパラメータの初期値 $\hat{\beta}_1^{(1)}$ と $\hat{\beta}_2^{(1)}$ に、回帰分析によって推定された $\hat{\delta}_1^{(1)}$ と $\hat{\delta}_2^{(1)}$ を加えて、次の反復のためのパラメータ $\hat{\beta}_1^{(2)}$ と $\hat{\beta}_2^{(2)}$ を計算する。

$$\begin{aligned}\hat{\beta}_1^{(2)} &= \hat{\beta}_1^{(1)} + \hat{\delta}_1^{(1)} \\ \hat{\beta}_2^{(2)} &= \hat{\beta}_2^{(1)} + \hat{\delta}_2^{(1)}\end{aligned}$$

次のパラメータの相対誤差が、

$$|\hat{\delta}_1^{(1)} / \hat{\beta}_1^{(1)}| < 0.001 \quad \text{かつ} \quad |\hat{\delta}_2^{(1)} / \hat{\beta}_2^{(1)}| < 0.001$$

となったときに繰り返しを終え手順 6 に行く。(収束の判定基準値を 0.001 としたが、実際には、求めたいパラメータの精度を高めるために 0.000001 のように小さい値を用いる。)

手順 1 に戻り、で $\hat{\beta}_1^{(1)} = \hat{\beta}_1^{(2)}$ と $\hat{\beta}_2^{(1)} = \hat{\beta}_2^{(2)}$ とおき直して、手順 2 から手順 5 を繰り返す。

【手順 6】

r 回目の繰り返しで収束したとする。計算されたパラメータの推定値を $\hat{\beta}_1 = \hat{\beta}_1^{(r+1)}$, $\hat{\beta}_2 = \hat{\beta}_2^{(r+1)}$ としてシグモイド曲線の最終的なパラメータの推定値とする。

【手順 7】

式 (5) 残差の 2 乗和を次の計算式により計算し、

$$S_T(\hat{\beta}_1, \hat{\beta}_2) = \sum_{i=1}^n \{y_i - f(x_i; \hat{\beta}_1, \hat{\beta}_2)\}^2 = \sum_{i=1}^n \{y_i - f(x_i; \hat{\beta}_1^{(r+1)}, \hat{\beta}_2^{(r+1)})\}^2$$

誤差分散 σ^2 を、次式のように

$$\hat{\sigma}^2 = \frac{S_T(\hat{\beta}_1, \hat{\beta}_2)}{n-2}$$

残差の 2 乗和を自由度 $n-2$ で割り推定する。

【手順 8】

r 回目の繰り返しで計算されたパラメータの推定値 $\hat{\beta}_1^{(r+1)}$, $\hat{\beta}_2^{(r+1)}$ からあらためて微分係数 $\hat{z}_{1i} = \hat{z}_{1i}^{(r+1)}$, $\hat{z}_{2i} = \hat{z}_{2i}^{(r+1)}$ を計算し $n \times 2$ の行列 \mathbf{Z}

$$\mathbf{Z} = \begin{bmatrix} \hat{z}_{11} & \hat{z}_{21} \\ \hat{z}_{12} & \hat{z}_{22} \\ \vdots & \vdots \\ \hat{z}_{1n} & \hat{z}_{2n} \end{bmatrix}$$

とする。2×2の積和行列 $\mathbf{Z}'\mathbf{Z}$ の逆行列 $(\mathbf{Z}'\mathbf{Z})^{-1}$ を計算し、誤差分散の推定値 $\hat{\sigma}^2$ を $(\mathbf{Z}'\mathbf{Z})^{-1}$ に掛けて分散共分散行列

$$\mathbf{V} = (\mathbf{Z}'\mathbf{Z})^{-1} \hat{\sigma}^2$$

を計算する。Vの対角要素 \hat{v}_{11} , \hat{v}_{22} の平方根が、 $\hat{\beta}_1$ と $\hat{\beta}_2$ の標準誤差となる。

【手順9】

手順8で求めた標準誤差が、自由度 $n-2$ の t 分布に従うことを用いて、 $100(1-\alpha)\%$ 信頼区間を求める。 $\hat{\beta}_2$ が $E50$ の推定値なので、

$$\hat{\beta}_2 \pm t(n-2; 1-\alpha) \sqrt{\hat{v}_{22}}$$

が $D50$ の信頼区間となる。

3.4. 計算例

《第1回目の反復》

【手順1】

1番目の反応率 y_1 との残差の推定値 $\hat{e}_1^{(1)}$ を、式(9)に $\hat{\beta}_1^{(1)} = -1.1$ と $\hat{\beta}_2^{(1)} = 3$ を代入し、次の結果を得る。

$$f_1^{(1)}(x_1; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)}) = \frac{100}{1 + \exp\{\hat{\beta}_1^{(1)}(\hat{\beta}_2^{(1)} - x_1)\}} = \frac{100}{1 + \exp\{-1.1 \cdot (3 - 0)\}} = 96.443$$

$$\hat{e}_1^{(1)} = y_1 - \frac{100}{1 + \exp\{-1.1 \cdot (3 - x_1)\}} = 109.9 - 96.443 = 13.457$$

2番目以降も同様に計算をする。16番の反応率 y_{16} との残差の推定値 $\hat{e}_{16}^{(1)}$ は、

$$\hat{e}_{16}^{(1)} = y_{16} - \frac{100}{1 + \exp\{-1.1 \cdot (3 - x_{16})\}} = 7.7 - 1.231 = 6.478$$

となる。すべての $\hat{e}_i^{(1)}$, $i=1,2,\dots,16$ について、表 3.1 に示す。

【手順2】

初期値 $\hat{\beta}_1^{(1)} = -1.1$ と $\hat{\beta}_2^{(1)} = 3.0$ を式(10)と(11)に代入し、1番目のデータについて微分係数 $\hat{z}_{11}^{(1)}$, $\hat{z}_{21}^{(1)}$ を計算する。

$$\hat{\eta}_1^{(1)} = \hat{\beta}_1^{(1)}(\hat{\beta}_2^{(1)} - x_1) = -1.1 \cdot (3 - 0) = -3.3$$

$$\hat{z}_{11}^{(1)} = \frac{-(\hat{\beta}_2^{(1)} - x_1) e^{\hat{\eta}_1^{(1)}} f(x_1; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_1^{(1)}}} = \frac{-(3 - 0) e^{-3.3} \cdot 96.443}{1 + e^{-3.3}} = -10.292$$

$$\hat{z}_{21}^{(1)} = \frac{-\hat{\beta}_1^{(1)} e^{\hat{\eta}_1^{(1)}} f(x_1; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_1^{(1)}}} = \frac{1.1 e^{-3.3} \cdot 96.443}{1 + e^{-3.3}} = 3.774$$

2 番目以降も同様に計算をする。16 番の 微分係数 $\hat{z}_{1,16}^{(1)}$, $\hat{z}_{2,16}^{(1)}$ を計算する、

$$\hat{\eta}_{16}^{(1)} = \hat{\beta}_1^{(1)} (\hat{\beta}_2^{(1)} - x_{16}) = -1.1 \cdot (3 - 7) = 4.4$$

$$\hat{z}_{1,16}^{(1)} = \frac{-(\hat{\beta}_2^{(1)} - x_{16}) e^{\hat{\eta}_{16}^{(1)}} f(x_{16}; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_{16}^{(1)}}} = \frac{-(3 - 7) e^{4.4} \cdot 1.213}{1 + e^{4.4}} = 4.793$$

$$\hat{z}_{2,16}^{(1)} = \frac{-\hat{\beta}_1^{(1)} e^{\hat{\eta}_{16}^{(1)}} f(x_{16}; \hat{\beta}_1^{(1)}, \hat{\beta}_2^{(1)})}{1 + e^{\hat{\eta}_{16}^{(1)}}} = \frac{1.1 e^{4.4} \cdot 1.213}{1 + e^{4.4}} = 1.318$$

となる。すべての $\hat{z}_{1i}^{(1)}$, $\hat{z}_{2i}^{(1)}$, $i = 1, 2, \dots, 16$ について、表 3.1 に示す。

表 3.1 第 1 回目の反復計算表

x	y			(y)	(x1)	(x2)
		eta(1)	f(1)	y·f(1) e(1)	f'(beta1) z1(1)	f'(beta2) z2(1)
0	109.9	-3.300	96.443	13.457	-10.292	3.774
0	88.7	-3.300	96.443	-7.743	-10.292	3.774
1	102.9	-2.200	90.025	12.875	-17.960	9.878
1	86.8	-2.200	90.025	-3.225	-17.960	9.878
2	65.1	-1.100	75.026	-9.926	-18.737	20.611
2	83.3	-1.100	75.026	8.274	-18.737	20.611
3	44.2	0.000	50.000	-5.800	0.000	27.500
3	41.3	0.000	50.000	-8.700	0.000	27.500
4	18.6	1.100	24.974	-6.374	18.737	20.611
4	20.7	1.100	24.974	-4.274	18.737	20.611
5	5.8	2.200	9.975	-4.175	17.960	9.878
5	4.3	2.200	9.975	-5.675	17.960	9.878
6	3.4	3.300	3.557	-0.157	10.292	3.774
6	1.2	3.300	3.557	-2.357	10.292	3.774
7	6.4	4.400	1.213	5.187	4.793	1.318
7	7.7	4.400	1.213	6.487	4.793	1.318

【手順 3】

残差の推定値 $\hat{\varepsilon}_i^{(1)}$ を従属変数 $e_i^{(1)}$ とし、微分係数を説明変数とした線形回帰式は次のようになる。

$$13.457 = \delta_1^{(1)} \cdot (-10.292) + \delta_2^{(1)} \cdot (3.774) + \varepsilon_1$$

$$-7.743 = \delta_1^{(1)} \cdot (-10.292) + \delta_2^{(1)} \cdot (3.774) + \varepsilon_2$$

$$12.875 = \delta_1^{(1)} \cdot (-17.960) + \delta_2^{(1)} \cdot (9.878) + \varepsilon_3$$

...

$$6.487 = \delta_1^{(1)} \cdot (4.793) + \delta_2^{(1)} \cdot (4.793) + \varepsilon_{16}$$

【手順 4】

手順 3 の線形回帰式のパラメータ $\delta_1^{(1)}$ と $\delta_2^{(1)}$ を, 切片なしの回帰分析を用いて推定し, 次の標準的な分散分析表と回帰係数の推定値を得る.

要因	自由度	平方和	平均平方	F 値
Model	2	201.259	100.630	2.086
Error	14	675.371	48.241	0.161
C. Total	16	876.630		

要因	推定値 e	SE	t 値	P 値
z1	-0.172	0.123	-1.396	0.184
z2	-0.171	0.115	-1.486	0.159

これから, $z_{i1}^{(1)}$, $z_{i2}^{(1)}$ の回帰係数の推定値として, それぞれ $\hat{\delta}_1^{(1)} = -0.172$, $\hat{\delta}_2^{(1)} = -0.171$ を得る.

【手順 5】

次の反復のためのパラメータ $\hat{\beta}_1^{(2)}$ と $\hat{\beta}_2^{(2)}$ を次のように計算し,

$$\hat{\beta}_1^{(2)} = \hat{\beta}_1^{(1)} + \hat{\delta}_1^{(1)} = -1.1 + (-0.172) = -1.272$$

$$\hat{\beta}_2^{(2)} = \hat{\beta}_2^{(1)} + \hat{\delta}_2^{(1)} = 3 + (-0.171) = 2.829$$

推定された $\hat{\delta}_1^{(1)}$ と $\hat{\delta}_2^{(1)}$ が, $\beta_1^{(1)}$ と $\beta_2^{(1)}$ に比べて, 次のように

$$|\hat{\delta}_1^{(1)} / \beta_1^{(1)}| = |(-0.172) / (-1.1)| = 0.156 > 0.001$$

$$|\hat{\delta}_2^{(1)} / \beta_2^{(1)}| = |(-0.171) / 3| = 0.057 > 0.001$$

相対的に大きいので, 手順 1 に戻る.

《第 2 回目の反復》

【手順 1】 $\hat{\beta}_1^{(1)}$ を $\hat{\beta}_1^{(2)} = -1.272$ に, $\hat{\beta}_2^{(1)}$ を $\hat{\beta}_2^{(2)} = 2.829$ におき直し, $f_i^{(2)}(x_i; \hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)})$, および $\hat{e}_i^{(2)} = y_i - f_i^{(2)}(x_i; \hat{\beta}_1^{(2)}, \hat{\beta}_2^{(2)})$, $i = 1, 2, \dots, n$ を計算する. 結果は 表 3.2 に示す.

【手順 2】 微分係数 $\hat{z}_{i1}^{(2)}$, $\hat{z}_{i2}^{(2)}$, $i = 1, 2, \dots, n$ を計算する. 結果は 表 3.2 に示す.

表 3.2 第 2 回目の反復計算表

x	y	$\eta(2)$	$f(2)$	$e(2)$	$z1(2)$	$z2(2)$
0	109.9	-3.598	97.336	12.564	-7.335	3.298
0	88.7	-3.598	97.336	-8.636	-7.335	3.298
1	102.9	-2.326	91.105	11.795	-14.822	10.308
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
7	7.7	5.306	0.494	7.206	2.050	0.625

【手順 3】手順 1 と 2 で推定された $\hat{e}_i^{(2)}$ を改めて従属変数 $e_i^{(2)}$ に、微分係数の推定値 $\hat{z}_{1i}^{(2)}$ と $\hat{z}_{2i}^{(2)}$ を独立変数 $z_{1i}^{(2)}$ と $z_{2i}^{(2)}$ におき直して、切片なしの線形回帰式をたてる。

$$e_i^{(2)} = \delta_1^{(2)} z_{1i}^{(2)} + \delta_2^{(2)} z_{2i}^{(2)} + \varepsilon_i \quad i=1,2,\dots,n$$

【手順 4】切片なしの回帰分析により回帰係数の推定値 $\hat{\delta}_1^{(2)} = -0.049$, $\hat{\delta}_2^{(2)} = 0.0049$ を得る。

【手順 5】次の反復のためのパラメータ $\hat{\beta}_1^{(3)}$ と $\hat{\beta}_2^{(3)}$ を計算する。

$$\hat{\beta}_1^{(3)} = \hat{\beta}_1^{(2)} + \hat{\delta}_1^{(2)} = -1.272 + (-0.049) = -1.321$$

$$\hat{\beta}_2^{(3)} = \hat{\beta}_2^{(2)} + \hat{\delta}_2^{(2)} = 2.829 + (-0.0049) = 2.834$$

反復計算が必要かの判定を行なう。

$$|\hat{\delta}_1^{(2)} / \hat{\beta}_1^{(2)}| = |(-0.049) / (-1.271)| = 0.038 > 0.001$$

$$|\hat{\delta}_2^{(2)} / \hat{\beta}_2^{(2)}| = |(0.0049) / 2.829| = 0.0017 > 0.001$$

相対的に大きいので、手順 1 に戻る。

《第 3 回目の反復》

【手順 1】 $\hat{\beta}_1^{(1)}$ を $\hat{\beta}_1^{(3)} = -1.321$ と $\hat{\beta}_2^{(1)}$ を $\hat{\beta}_2^{(3)} = 2.834$ とおき直し、 $f_i^{(3)}(x_i; \hat{\beta}_1^{(3)}, \hat{\beta}_2^{(3)})$, および $\hat{e}_i^{(3)} = y_i - f_i^{(3)}(x_i; \hat{\beta}_1^{(3)}, \hat{\beta}_2^{(3)})$, $i=1,2,\dots,n$ を計算する。結果を表 3.3 に示す。

【手順 2】微分係数 $\hat{z}_{1i}^{(3)}$, $\hat{z}_{2i}^{(3)}$, $i=1,2,\dots,n$ を計算する。結果を表 3.3 に示す。

表 3.3 第 3 回目の反復計算表

x	y	$\eta(3)$	$f(3)$	$e(3)$	$z1(3)$	$z2(3)$
0	109.9	-3.744	97.688	12.212	-6.400	2.983
0	88.7	-3.744	97.688	-8.988	-6.400	2.983
1	102.9	-2.423	91.854	11.046	-13.722	9.884
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
7	7.7	5.503	0.406	7.294	1.683	0.534

【手順3】 手順1と2で推定された $\hat{e}_i^{(3)}$ を改めて従属変数 $e_i^{(3)}$ に、微分係数の推定値 $\hat{z}_{1i}^{(3)}$ と $\hat{z}_{2i}^{(3)}$ を独立変数 $z_{1i}^{(3)}$ と $z_{2i}^{(3)}$ におき直して、切片なしの線形回帰式をたてる。

$$e_i^{(3)} = \delta_1^{(3)} z_{1i}^{(3)} + \delta_2^{(3)} z_{2i}^{(3)} + \varepsilon_i \quad i=1,2,\dots,n$$

【手順4】 切片なしの回帰分析により、 $\hat{\delta}_1^{(3)} = -0.00056$ 、 $\hat{\delta}_2^{(3)} = 0.0022$ を得る。

【手順5】 次の反復のためのパラメータ $\hat{\beta}_1^{(4)}$ と $\hat{\beta}_2^{(4)}$ を計算する。

$$\hat{\beta}_1^{(4)} = \hat{\beta}_1^{(3)} + \hat{\delta}_1^{(3)} = -1.321 + (-0.00056) = -1.322$$

$$\hat{\beta}_2^{(4)} = \hat{\beta}_2^{(3)} + \hat{\delta}_2^{(3)} = 2.834 + (-0.0022) = 2.832$$

反復計算が必要かの判定を行なう。

$$|\hat{\delta}_1^{(3)} / \hat{\beta}_1^{(3)}| = |(-0.00056) / (-1.321)| = 0.00042 < 0.001$$

$$|\hat{\delta}_2^{(3)} / \hat{\beta}_2^{(3)}| = |(0.0022) / 2.834| = 0.00078 < 0.001$$

と収束の条件を満たしたので、手順6に行く。

【手順7】

シグモイド曲線の最終的なパラメータの推定値は、

$$\hat{\beta}_1 = \hat{\beta}_1^{(4)} = -1.322$$

$$\hat{\beta}_2 = \hat{\beta}_2^{(4)} = 2.832$$

となる。

【手順8】

式(5) 残差の2乗和を次の計算式により計算する。計算表を

$$\begin{aligned} S_T(\hat{\beta}_1, \hat{\beta}_2) &= \sum_{i=1}^n \{y_i - f(x_i; \hat{\beta}_1, \hat{\beta}_2)\}^2 \\ &= (109.9 - 97.686)^2 + (88.7 - 97.686)^2 + \dots + (7.7 - 0.404)^2 = 659.580 \end{aligned}$$

誤差分散 σ^2 を次式で推定する。

$$\hat{\sigma}^2 = \frac{S_T(\hat{\beta}_1, \hat{\beta}_2)}{n-2} = \frac{659.580}{16-2} = 47.113$$

【手順8】

微分係数 $\hat{z}_{1i}^{(4)}$ 、 $\hat{z}_{2i}^{(4)}$ を再度計算し、 $\hat{z}_{1i} = \hat{z}_{1i}^{(4)}$ 、 $\hat{z}_{2i} = \hat{z}_{2i}^{(4)}$ とおき直して $n \times 2$ の行列 \mathbf{Z} とする。

$$\mathbf{Z} = \begin{bmatrix} \hat{z}_{11} & \hat{z}_{21} \\ \hat{z}_{12} & \hat{z}_{22} \\ \vdots & \vdots \\ \hat{z}_{1n} & \hat{z}_{2n} \end{bmatrix} = \begin{bmatrix} -6.402 & 2.987 \\ -6.402 & 2.987 \\ \vdots & \vdots \\ 1.676 & 0.531 \end{bmatrix}$$

2×2 の積和行列 $\mathbf{Z}'\mathbf{Z}$ および逆行列を計算する．逆行列は，固有値の逆数を余因子行列に掛ける方法を用いる．

$$\mathbf{Z}'\mathbf{Z} = \begin{bmatrix} -6.402 & -6.402 & \cdots & 1.676 \\ 2.987 & 2.987 & \cdots & 0.531 \end{bmatrix} \begin{bmatrix} -6.402 & 2.987 \\ -6.402 & 2.987 \\ \vdots & \vdots \\ 1.676 & 0.531 \end{bmatrix} = \begin{bmatrix} 187.245 & 6.713 \\ 6.713 & 4404.606 \end{bmatrix}$$

$$\begin{aligned} (\mathbf{Z}'\mathbf{Z})^{-1} &= \frac{1}{187.245 \times 4404.606 - 6.713 \times 6.713} \begin{bmatrix} 4404.606 & -6.713 \\ -6.713 & 187.245 \end{bmatrix} \\ &= \begin{bmatrix} 0.00054135 & -0.00000083 \\ -0.00000083 & 0.00022704 \end{bmatrix} \end{aligned}$$

誤差の 2 乗和 S_T から分散 σ^2 を

$$\hat{\sigma}^2 = \frac{S_T}{n-2} = \frac{659.580}{16-2} = 47.113$$

で推定し， $(\mathbf{Z}'\mathbf{Z})^{-1}$ に掛けて分散共分散行列

$$\mathbf{V} = (\mathbf{Z}'\mathbf{Z})^{-1} \hat{\sigma}^2 = 47.113 \begin{bmatrix} 0.00054135 & -0.00000083 \\ -0.00000083 & 0.00022704 \end{bmatrix} = \begin{bmatrix} 0.025505 & -0.000039 \\ -0.000039 & 0.010696 \end{bmatrix}$$

を計算する． \mathbf{V} の対角要素 \hat{v}_{11} ， \hat{v}_{22} の平方根が，

$$\text{s.e.}(\hat{\beta}_1) = \sqrt{\hat{v}_{11}} = 0.160$$

$$\text{s.e.}(\hat{\beta}_2) = \sqrt{\hat{v}_{22}} = 0.103$$

$\hat{\beta}_1$ と $\hat{\beta}_2$ 標準誤差となる．

【手順 9】

手順 8 で求めた標準誤差が，自由度 $n-2$ の t 分布に従うことを用いて， $100(1-\alpha)\%$ 信頼区間を求める． $\hat{\beta}_2$ が $\log_2(D50)$ の推定値となるので，

$$\hat{\beta}_2 \pm \sqrt{\hat{v}_{22}} t(2, n-2; 1-\alpha) = 2.832 \pm 0.103 \times 2.145 = (2.611, 3.054)$$

が $\log_2(D50)$ の信頼区間となる．この信頼区間は，底が 2 の対数となっているので，元の溶液濃度では， $D50 = 2^{2.832} = 7.121$ mg/mL、95% の信頼区間の下限は $L95 = 6.106$ mg/mL，上限は $U95 = 8.305$ mg/mL となる．

表 3.4 パラメータの信頼区間の計算シート

x	y	eta(4)		y-f(4)		f(beta1)	f(beta2)
		eta(4)	f(4)	e(4)	e(4)^2	z1(4)	z2(4)
0	109.9	-3.743	97.686	12.214	149.183	-6.402	2.987
0	88.7	-3.743	97.686	-8.986	80.747	-6.402	2.987
1	102.9	-2.421	91.843	11.057	122.254	-13.725	9.900
1	86.8	-2.421	91.843	-5.043	25.433	-13.725	9.900
2	65.1	-1.100	75.020	-9.920	98.410	-15.594	24.765
2	83.3	-1.100	75.020	8.280	68.555	-15.594	24.765
3	44.2	0.222	44.476	-0.276	0.076	4.146	32.635
3	41.3	0.222	44.476	-3.176	10.090	4.146	32.635
4	18.6	1.543	17.604	0.996	0.991	16.940	19.169
4	20.7	1.543	17.604	3.096	9.583	16.940	19.169
5	5.8	2.865	5.392	0.408	0.167	11.058	6.741
5	4.3	2.865	5.392	-1.092	1.191	11.058	6.741
6	3.4	4.186	1.497	1.903	3.620	4.672	1.949
6	1.2	4.186	1.497	-0.297	0.088	4.672	1.949
7	6.4	5.508	0.404	5.996	35.955	1.676	0.531
7	7.7	5.508	0.404	7.296	53.235	1.676	0.531
				S_T=	659.580		
				siguma^2=	47.113		

3.5. 解説, 実用上の諸問題

D10, および D90 の求め方 ==> 第 12 回高橋セミナー p9.

シグモイド曲線の関数

幾つかの線形化の方法と対比しつつ、線形化は元の現象がロジスティック曲線であることを示し、シグモイド曲線の事例としてロジスティック曲線を取り上げたことを解説する。

誤差の正規性の仮定が困難な例とその対処法

誤差の正規性の仮定の許容範囲、観測データが細胞数の場合に、その分布がポアソン分布に従っている場合には、 $1/y_i$ 重み付き非線形最小 2 乗法が必要なことを解説する。観測データの SD が、平均値に比例して大きくなる場合には、 $1/y_i^2$ の重み付き非線形最小 2 乗法が必要なことを解説する。

実際の計算

専用ソフト，汎用ソフトの使い分けなど、簡単に解説する。解散事例は Excel でも追試できるように工夫してあることえを解説する。

微分式

- ・あらかじめ統計ソフトに組み込まれていることを確認することを注意する.
- ・ダミー変数を用いる非線形回帰分析の場合には、計算アルゴリズムが組み込まれてる汎用ソフトを用いることの解説をする。

ロジスティク曲線の 95%信頼区間

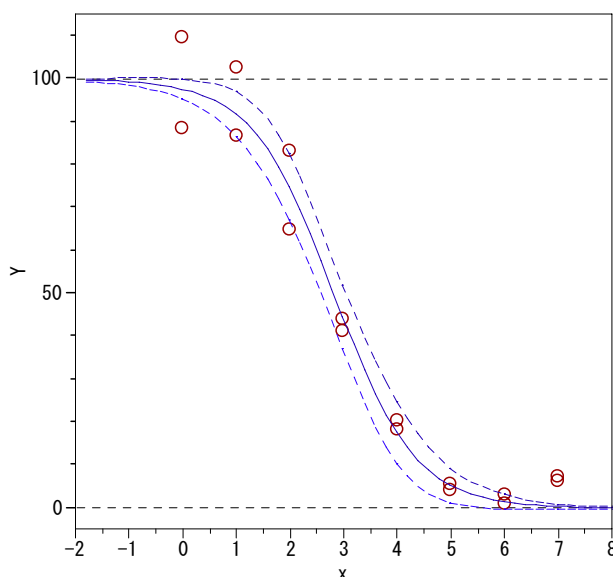


図 3.3 シグモイド曲線の 95%信頼区間

計算の方法を解説し、2 ポイント程度計算例を示す。

生データでの解析と反応率での解析を解説する。

陰性対照あるいは陰性対照のデータの繰り返しが多くあり、分散が大きくなければ反応率での計算でよい。

最大反応をデータから推定する方法考え方を解説する。

最小反応をデータから推定する方法考え方解説する。

最大反応が陰性対照として別にあり，同時に推定する場合の解析を解説する。

最小反応が対陽性照として別にあり，同時に推定する場合の解析を解説する。

対数用量に変換しないで直接 $D50$ を求める方法との対比解説し、用量を対数変換することの良さを解説する。

4. 統計ソフトによる検証

4.1. SAS の NLIN プロシジャによる計算

SASのNLINプログラム

```
Title 'photo_TX 2004-01-19 Y.Takahashi ' ;

data d01 ;
  input id x y ;
datalines ;
  1 0 109.9
  2 0 88.7
  3 1 102.9
  4 1 86.8
  5 2 65.1
  6 2 83.3
  7 3 44.2
  8 3 41.3
  9 4 18.6
  0 4 20.7
 11 5 5.8
 12 5 4.3
 13 6 3.4
 14 6 1.2
 15 7 6.4
 16 7 7.7
;
proc print data=d01 ;
run ;

proc nlin data=d01 method=gauss listall maxiter=3 ;
  parms beta1=-1.1 beta2=3 ;
  model y = 100 / ( 1 + exp(beta1*(beta2 - x)) ) ;
run ;
```

SASのNLINによる偏微分式の出力

```
MODEL. y = 100 / ( 1 + EXP(beta1 * (beta2 - x))) ;

@MODEL. y/@beta1 = ( 0 - (beta2 - x) * EXP(beta1 * (beta2 - x)) * MODEL. y)
                  / ( 1 + EXP(beta1 * (beta2 - x))) ;

@MODEL. y/@beta2 = ( 0 - beta1 * EXP(beta1 * (beta2 - x)) * MODEL. y) / ( 1
                  + EXP(beta1 * (beta2 - x))) ;
```

SASのNLINによる繰り返し計算

(反復は0から始まり、3回の反復で強制終了)

Iterative Phase			
Iter	beta1	beta2	Sum of Squares
0	-1.1000	3.0000	876.6
1	-1.2724	2.8295	664.5
2	-1.3210	2.8343	659.6
3	-1.3215	2.8321	659.6

WARNING: Maximum number of iterations exceeded.
WARNING: PROC NLIN failed to converge.

Estimation Summary (Not Converged)

Method	Gauss-Newton
Iterations	3
R	0.000377
PPC(beta1)	0.000167
RPC(beta2)	0.00078
Object	0.000034
Objective	659.5795
Observations Read	16
Observations Used	16
Observations Missing	0

NOTE: An intercept was not specified for this model.

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Regression	2	53184.8	26592.4	564.44	<.0001
Residual	14	659.6	47.1128		
Uncorrected Total	16	53844.4			
Corrected Total	15	24062.3			

Parameter	Estimate	Approx		
		Std Error	Approximate 95% Confidence Limits	
beta1	-1.3215	0.1597	-1.6641	-0.9790
beta2	2.8321	0.1034	2.6103	3.0540

Approximate Correlation Matrix

	beta1	beta2
beta1	1.0000000	-0.0023535
beta2	-0.0023535	1.0000000

4.2. JMP のスクリプト言語による繰り返し計算

JMPのスクリプトによるプログラム

```
// The Nonlinear regression model Photo_TX    2004-1-29  Y.Takahashi

x=[ 0, 0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6,
6, 7, 7 ] ;
y=[109.9, 88.7, 102.9, 86.8, 65.1, 83.3, 44.2, 41.3, 18.6, 20.7, 5.8, 4.3, 3.4, 1.2,
6.4, 7.7 ] ;

b =[ -1.1 , 3 ] ;
delta=[0, 0] ;

for( i=1, i<=4, i++,
    b = b + delta ;                               show(i) ;
    eta    = b[1] :* (b[2] - x) ;
    fx     = 100 / (1 + exp (eta));
    z1    =  -( b[2] - x ) :* exp( eta ) :* fx ) :* (1/ (1 + exp( eta ) )) ;
    z2    =  - b[1] :* exp(eta) :* fx  :* (1/ (1 + exp( eta ) )) ;
    z     = z1 || z2 ;
    delta =  inverse(z`*z)*z`(y-fx) ;
    delta_b = delta:*(1/b) ;
    show(round(b, 6), round(delta, 6), round(delta_b, 6)) ;
);

ST = (y-fx)`*(y-fx) ;                               show(round(ST, 6)) ;
sigma2= ST / (nrow(z)-ncol(z)) ;   show(round(sigma2, 6)) ;
zpz=z`*z ;                                           show(round(zpz, 6)) ;
izpz=inv(zpz) ;                                       show(round(izpz, 8)) ;
cov=inv(z`*z) :* sigma2 ;                             show(round(cov, 6)) ;
```

JMPのスクリプトによる計算結果

```
i:1 Round(b, 6):      [-1.1,      3      ]
    Round(delta, 6): [-0.172356, -0.170542] Round(delta_b, 6): [0.156687, -0.056847]

i:2 Round(b, 6):      [-1.272356, 2.829458]
    Round(delta, 6): [-0.048621, 0.004884] Round(delta_b, 6): [0.038213, 0.001726]

i:3 Round(b, 6):      [-1.320977, 2.834342]
    Round(delta, 6): [-0.000558, -0.002212] Round(delta_b, 6): [0.000422, -0.00078]

i:4 Round(b, 6):      [-1.321535, 2.83213]
    Round(delta, 6): [-0.000221, -0.000027] Round(delta_b, 6): [0.000167, -0.000009]

Round(ST, 6):      [659.579548]
```

```
Round(sigma2, 6):[ 47.112825 ]
Round(zpz, 6):   [1847.244564  6.713097,
                  6.713097  4404.606164]
Round(izpz, 8):  0.00054135 -0.00000083,
                  -0.00000083  0.00022704]
Round(cov, 6):   [0.025505 -0.000039,
                  -0.000039  0.010696]
```

ロジスティック曲線の95%信頼区間

```
h=VecDiag(z*izpz*z`); show(round(H, 6)) ;
L95=fx - sqrt(sigma2 :* h) :*2.1448 ; show(round(L95, 6)) ;
U95=fx + sqrt(sigma2 :* h) :*2.1448 ; show(round(U95, 6)) ;
yhat= fx || L95 || U95 ; show(round(yhat, 6)) ;
```

predict	L95M	U95M	H
97.686	95.3937	99.9783	0.0242
97.686	95.3937	99.9783	0.0242
91.8432	86.6495	97.0368	0.1245
91.8432	86.6495	97.0368	0.1245
75.0202	67.349	82.6914	0.2715
75.0202	67.349	82.6914	0.2715
44.4765	37.1026	51.8503	0.2509
44.4765	37.1026	51.8503	0.2509
17.6044	10.4188	24.7901	0.2382
17.6044	10.4188	24.7901	0.2382
5.3915	1.3227	9.4604	0.0764
5.3915	1.3227	9.4604	0.0764
1.4973	-0.1595	3.154	0.0127
1.4973	-0.1595	3.154	0.0127
0.4038	-0.182	0.9896	0.0016
0.4038	-0.182	0.9896	0.0016

4.3. JMP の非線形プログラムによる計算

JMP の標準的な収束条件による打ち切り

	平方和	beta1	beta2
Taking Steps 2	664.461	-1.27236	2.829458
Taking Steps 3	659.602	-1.32098	2.834342
Taking Steps 4	659.58	-1.32154	2.83213
Taking Steps 5	659.579	-1.32176	2.832103
Converged in the Gradient 5	659.579	-1.32176	2.832094

/* ここがスクリプトでの打ち切り */

Solution

	SSE	DFE	MSE	RMSE	
	659.57945234	14	47.112818	6.8638778	
Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL	
beta1	-1.321756743	0.15974088	-1.706492	-1.0505052	
beta2	2.8320939341	0.10341431	2.62571906	3.04059641	

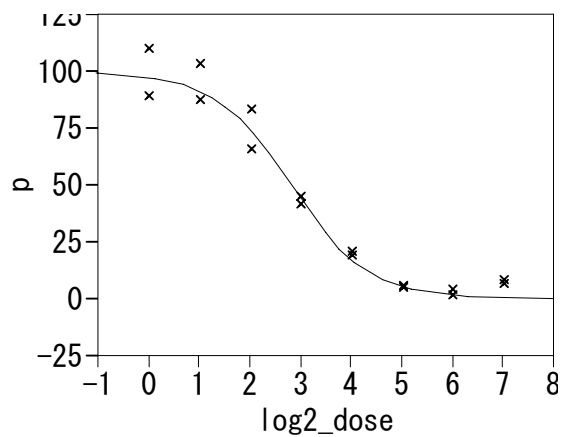


図 4.1 JMP の非線形プログラムによるあてはめ

4.4. Excel による繰り返し計算

第1反復

シグモイド曲線の線形化の方式による非線形回帰

<<<第1ステップ>>>

$$(y) = \text{delta1} (x1) + \text{delta2} (x2) : e(1) = \text{delta} * z1 + \text{delta2} * z2, \text{ 切片なし}$$

	beta_i (1)	delta (1)	beta_i (2)
beta1=	-1.1	-0.1724	-1.2724
beta2=	3	-0.1705	2.8295

x	y	eta(1)	f(1)	(y) y-f(1) e(1)	(x1) f(beta1) z1(1)	(x2) f(beta2) z2(1)
0	109.9	-3.300	96.443	13.457	-10.292	3.774
0	88.7	-3.300	96.443	-7.743	-10.292	3.774
1	102.9	-2.200	90.025	12.875	-17.960	9.878
1	86.8	-2.200	90.025	-3.225	-17.960	9.878
2	65.1	-1.100	75.026	-9.926	-18.737	20.611
2	83.3	-1.100	75.026	8.274	-18.737	20.611
3	44.2	0.000	50.000	-5.800	0.000	27.500
3	41.3	0.000	50.000	-8.700	0.000	27.500
4	18.6	1.100	24.974	-6.374	18.737	20.611
4	20.7	1.100	24.974	-4.274	18.737	20.611
5	5.8	2.200	9.975	-4.175	17.960	9.878
5	4.3	2.200	9.975	-5.675	17.960	9.878
6	3.4	3.300	3.557	-0.157	10.292	3.774
6	1.2	3.300	3.557	-2.357	10.292	3.774
7	6.4	4.400	1.213	5.187	4.793	1.318
7	7.7	4.400	1.213	6.487	4.793	1.318

LINEST 関数による最小2乗法解

	beta2(1)	beta1(1)	切片
パラメタ	-0.1705	-0.1724	0
SE	0.1148	0.1235	#N/A
	0.2214	6.9456	#N/A
	1.9907	14	#N/A
	192.0696	675.3714	#N/A

× r^2	○ root_ms
× reg_F	○ error_df
× reg_ss	○ error_ss

注) Excelで切片なしの場合分散分析表に不具合がある

第2反復

シグモイド曲線の線形化の方式による非線形回帰

<<<第2ステップ>>>

$(y) = \text{delta1}(x1) + \text{delta2}(x2) : e(1) = \text{delta} * z1 + \text{delta2} * z2$, 切片なし

	beta_i (2)	delta (1)	beta_i (3)
beta1=	-1.27236	-0.0486	-1.3210
beta2=	2.829458	0.0049	2.8343

	(y)	(x1)	(x2)
	y-f(1)	f(beta1)	f(beta2)
	e(1)	z1(1)	z2(1)
x	y	eta(1)	f(1)
0	109.9	-3.600	97.341
0	88.7	-3.600	97.341
1	102.9	-2.328	91.115
1	86.8	-2.328	91.115
2	65.1	-1.055	74.180
2	83.3	-1.055	74.180
3	44.2	0.217	44.596
3	41.3	0.217	44.596
4	18.6	1.489	18.402
4	20.7	1.489	18.402
5	5.8	2.762	5.943
5	4.3	2.762	5.943
6	3.4	4.034	1.739
6	1.2	4.034	1.739
7	6.4	5.306	0.494
7	7.7	5.306	0.494

LINEST 関数による最小2乗法解

	beta2(1)	beta1(1)	切片
パラメタ	0.0049	-0.0486	0
SE	0.1054	0.1510	#N/A
	-0.0430	6.8634	#N/A
	-0.2884	14	#N/A
	-27.1683	659.4806	#N/A

$\times r^2$	○root_ms
$\times \text{reg}_F$	○error_df
$\times \text{reg}_{ss}$	○error_ss

注)Excelで切片なしの場合分散分析表に不具合がある

第3反復

シグモイド曲線の線形化の方式による非線形回帰

<<<第3ステップ>>>

$(y) = \text{delta1} (x1) + \text{delta2} (x2) : e(1) = \text{delta}*z1 + \text{delta2}*z2$, 切片なし

	beta_i (3)	delta (1)	beta_i (4)
beta1=	-1.32098	-0.0006	-1.3215
beta2=	2.834342	-0.0022	2.8321

	x	y	eta(1)	f(1)	(y) y-f(1) e(1)	(x1) f(beta1) z1(1)	(x2) f(beta2) z2(1)
	0	109.9	-3.744	97.689	12.211	-6.399	2.982
	0	88.7	-3.744	97.689	-8.989	-6.399	2.982
	1	102.9	-2.423	91.857	11.043	-13.720	9.880
	1	86.8	-2.423	91.857	-5.057	-13.720	9.880
	2	65.1	-1.102	75.066	-9.966	-15.616	24.725
	2	83.3	-1.102	75.066	8.234	-15.616	24.725
	3	44.2	0.219	44.551	-0.351	4.092	32.632
	3	41.3	0.219	44.551	-3.251	4.092	32.632
	4	18.6	1.540	17.656	0.944	16.947	19.206
	4	20.7	1.540	17.656	3.044	16.947	19.206
	5	5.8	2.861	5.413	0.387	11.087	6.763
	5	4.3	2.861	5.413	-1.113	11.087	6.763
	6	3.4	4.182	1.504	1.896	4.690	1.957
	6	1.2	4.182	1.504	-0.304	4.690	1.957
	7	6.4	5.503	0.406	5.994	1.684	0.534
	7	7.7	5.503	0.406	7.294	1.684	0.534

LINEST 関数による最小2乗法解

	beta2(1)	beta1(1)	切片
パラメタ	-0.0022	-0.0006	0
SE	0.1034	0.1596	#N/A
	-0.0481	6.8639	#N/A
	-0.3212	14	#N/A
	-30.2684	659.5800	#N/A

<input checked="" type="checkbox"/> r^2	<input type="checkbox"/> root_ms
<input checked="" type="checkbox"/> reg_F	<input type="checkbox"/> error_df
<input checked="" type="checkbox"/> reg_ss	<input type="checkbox"/> error_ss

注)Excelで切片なしの場合分散分析表に不具合がある

信頼区間の計算

シグモイド曲線の線形化の方式による非線形回帰

<<<信頼区間>>>

$(y) = \text{delta1} (x1) + \text{delta2} (x2) : e(1) = \text{delta} * z1 + \text{delta2} * z2, \text{ 切片なし}$

	estimate	s.e. t(14,0.05)		L95	U95		
beta1=	-1.32154	0.1597	2.1448	-1.66406	-0.97901		
beta2=	2.83213	0.1034	2.1448	2.61031	3.053951		
				(y)		(x1)	(x2)
				y-f(1)		f(beta1)	f(beta2)
				e(1)		z1(1)	z2(1)
x	y	eta(1)	f(1)				
0	109.9	-3.743	97.686	12.214	149.183	-6.402	2.987
0	88.7	-3.743	97.686	-8.986	80.747	-6.402	2.987
1	102.9	-2.421	91.843	11.057	122.254	-13.725	9.900
1	86.8	-2.421	91.843	-5.043	25.433	-13.725	9.900
2	65.1	-1.100	75.020	-9.920	98.410	-15.594	24.765
2	83.3	-1.100	75.020	8.280	68.555	-15.594	24.765
3	44.2	0.222	44.476	-0.276	0.076	4.146	32.635
3	41.3	0.222	44.476	-3.176	10.090	4.146	32.635
4	18.6	1.543	17.604	0.996	0.991	16.940	19.169
4	20.7	1.543	17.604	3.096	9.583	16.940	19.169
5	5.8	2.865	5.392	0.408	0.167	11.058	6.741
5	4.3	2.865	5.392	-1.092	1.191	11.058	6.741
6	3.4	4.186	1.497	1.903	3.620	4.672	1.949
6	1.2	4.186	1.497	-0.297	0.088	4.672	1.949
7	6.4	5.508	0.404	5.996	35.955	1.676	0.531
7	7.7	5.508	0.404	7.296	53.235	1.676	0.531
					659.580		
					47.113		

行列計算による分散共分散の推定

Z'Z=	1847.244564	6.713096941
	6.713096941	4404.606164
(Z'Z) ⁻¹ =	0.00054135	-8.25076E-07
	-8.25076E-07	0.000227036
V=	0.02550452	-3.88717E-05
	-3.88717E-05	0.010696322

元の用量に戻した結果

D50	L95	U95
7.121	6.106	8.305

5. Draper (3rd edition) の 24 章本を基にした入門

5.1. 非線形モデルのための最小 2 乗法

(原著の式をこの節では、パワーモデルとロジスティック曲線に説明の都合上変えてある)

Introduction

多くの統計の教科書で取り上げられている回帰モデルは、線形回帰モデルであり、非線形回帰モデルは極めてまれである。非線形回帰モデルの正規方程式は線形ではなく、一般的には解くことができない。偏差平方和を最小化するために、繰り返し計算が必要となる。非線形となる現象は、多種多様であり、どのような非線形モデルが適切なのかという選択の問題も絡み、扱いにくい問題ではあるが、生物系で応用範囲が広いと思われるロジスティック曲線を例として取り上げ、非線形回帰モデルの考え方、繰り返し計算による最小 2 乗法による解の求め方を示す。

非線形モデル

線形モデルは、

$$y_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \varepsilon_i, \quad i=1,2,\dots,n \quad (24.1.1^*)$$

のように、従属変数 y_i が、複数のパラメータ $\beta_1, \beta_2, \dots, \beta_p$ が独立変数 z_1, z_2, \dots, z_p の和の形となっている場合である。この線形モデルが適当でないような場合も数多くある。

式 (24.1.1*) の形を取らないモデルはすべて非線形モデルと呼ばれる。投与量を z に対して観察された薬物濃度曲線下の面積 (AUC) を y とした場合のパワーモデル、

$$y_i = \beta_0 z^{\beta_1}, \quad i=1,2,\dots,n \quad (24.1.2^*)$$

は、 x に β_1 のべき乗が付いているので非線形モデルである。本章で取り扱うロジスティック曲線、

$$y_i = \frac{\beta_3}{1 + \exp(-\beta_1(\ln(x_i) - \ln(\beta_2)))} + \varepsilon_i \quad i=1,2,\dots,n \quad (24.1.3^*)$$

も非線形モデルである。

式 (1) および式 (2) は、共に非線形モデルであるが、本質的には異なる特質を持っている。式 (1) は、 e を底とする自然対数を取ることによって、つぎの形に変換で

きる.

$$\ln y_i = \ln \beta_0 + \beta_1 x_i, \quad i=1,2,\dots,n \quad (24.1.4^*)$$

$\ln \beta_1$ を β_1' と置き換えることにより, 線形モデルとなる. 線形に変換可能な場合に本質的線型である.

しかし, 式 (3) は, パラメータ $\beta_1, \beta_2, \beta_3$ について線形にすることは不可能であり, このようなモデルを, 本質的に非線形という.

非線形の場合の最小2乗法

(ここからは, 原著の訳, 線形化の記述に限定し, 他の記述は省略した)

非線形最小2乗法の数式の表記は, 線形の場合とは異なり, まぎらわしいのであるが文献では確立している. その違いを 表 5.1 に示す.

仮定したモデルを次のような式とする.

$$Y = f(\xi_1, \xi_2, \dots, \xi_k; \theta_1, \theta_2, \dots, \theta_k) + \varepsilon \quad (24.1.5)$$

次のベクトル表記を用いれば,

$$\begin{aligned} \xi &= (\xi_1, \xi_2, \dots, \xi_k)' \\ \theta &= (\theta_1, \theta_2, \dots, \theta_k)' \end{aligned}$$

式 (24.1.5) は,

$$Y = f(\xi, \theta) + \varepsilon$$

または, $E(\varepsilon) = 0$ と仮定するならば

$$E(Y) = f(\xi, \theta) \quad (24.1.6)$$

のように簡略化できる. また, 誤差が互いに無相関で, 分散が $V(\varepsilon) = \sigma^2$ である場合は, 通常, $\varepsilon \sim N(0, \sigma^2)$ と書くことができ, 従って誤差は互いに独立である.

表 5.1 Standard Notations for Linear and Nonlinear Least Squares

	線形	非線形
反応変数	Y	Y
添字	$i=1,2,\dots,n$	$u=1,2,\dots,n$
説明変数	X_1, X_2, \dots, X_k	$\xi_1, \xi_2, \dots, \xi_k$ (Sometimes $_t$ = time, or T = temperature, etc., sometimes even X_1, X_2, \dots, X_k)
パラメータ	$\beta_1, \beta_2, \dots, \beta_k$	$\theta_1, \theta_2, \dots, \theta_k$ (Sometimes, $\alpha, \beta, \dots, \phi, \dots$, etc.)

n 個の観測値が $u = 1, 2, \dots, n$, に対して

$$Y_u, \xi_{1u}, \xi_{2u}, \dots, \xi_{ku},$$

という形をとるとするならば, モデルを次ように書くことができる.

$$Y_u = f(\xi_{1u}, \xi_{2u}, \dots, \xi_{ku}; \theta_1, \theta_2, \dots, \theta_k) + \varepsilon_u \quad (24.1.7)$$

ここで, ε_u , $u = 1, 2, \dots, n$, は u 番目の誤差であり, さらに, $\xi = (\xi_1, \xi_2, \dots, \xi_k)'$ とすれば,

$$Y_u = f(\xi, \theta) + \varepsilon_u \quad (24.1.8)$$

と書くこともできる. 正規性および誤差の独立性を, $\varepsilon \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$ と書くことができる. ここに, $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k)'$ で $\mathbf{0}$ はゼロベクトル, \mathbf{I} は, 単位行列であり, とともに適切な大きさを持っているとする. 非線形モデルおよび与えられたデータに対して誤差の 2 乗和を次のように定義する.

$$S(\theta) = \sum_{u=1}^n \{Y_u - f(\xi, \theta)\}^2. \quad (24.1.9)$$

最小 2 乗推定量 $\hat{\theta}$ を見出すためには, 式 (24.1.9) を θ について微分する必要がある. これらは, $\hat{\theta}$ について解くための p 個の正規方程式となる. 正規方程式は $i = 1, 2, \dots, p$ に対してつぎの形をとる

$$\sum_{u=1}^n \{Y_u - f(\xi, \hat{\theta})\} \left[\frac{\partial f(\xi, \theta)}{\partial \theta_i} \right]_{\theta = \hat{\theta}} = 0 \quad (24.1.10)$$

ここで大括弧で示された量は $f(\xi_u, \theta)$ の θ_i についての導関数となり, θ 's は対応する添字の $\hat{\theta}$'s で置き換えられる. 関数 $f(\xi_u, \theta)$ が線型るとき, この量は, ξ_u のみの関数となり, $\hat{\theta}$'s を全く含まない. たとえば

$$f(\xi, \theta) = \theta_1 \xi_{1u} + \theta_2 \xi_{2u} + \dots + \theta_k \xi_{ku},$$

ならば,

$$\frac{\partial f}{\partial \theta_i} = \xi_{iu}, \quad i = 1, 2, \dots, p$$

は, θ について独立である. これは, $\theta_1, \theta_2, \dots, \theta_p$ の線形方程式の形での正規方程式となる. モデルが, $\hat{\theta}$'s において非線形であるとき, 正規方程式も非線形となる. 簡単な例, 非線形モデルのなかに 1 つだけのパラメータ θ を含む場合を示してみよう.

例題 モデル $Y = f(\theta, t) + \varepsilon$ の θ の最小 2 乗推定量 $\hat{\theta}$ を得るために正規方程式を作りたいとしよう. ここで, n 組の観測値 $(Y_1, t_1), (Y_2, t_2), \dots, (Y_n, t_n)$ について $f(\theta, t) = e^{-\theta t}$ を考えよう. θ についての偏微分は,

$$\frac{\partial f}{\partial \theta} = -te^{-\theta t}$$

であり，式(24.1.10) を用いると，一つの正規方程式が得られる．

$$\sum_{u=1}^n \{Y_u - e^{-\hat{\theta}t_u}\} \left[-t_u e^{-\hat{\theta}t_u} \right] = 0.$$

または，

$$? \sum_{u=1}^n Y_u t_u e^{-\hat{\theta}t} - \sum_{u=1}^n t_u e^{-2\hat{\theta}t_u} = 0. \quad (24.1.11)$$

$$\sum_{u=1}^n t_u e^{-2\hat{\theta}t_u} - \sum_{u=1}^n Y_u t_u e^{-\hat{\theta}t_u} = 0. \quad (24.1.12)$$

である．

パラメータが 1 個で比較的簡単なモデルの場合でも，このたった 1 つの正規方程式を解いて $\hat{\theta}$ を見出すことは容易でない．もつと多くのパラメータ・モデルに含まれていたり，さらに複雑な場合には，正規方程式の解を得ることは，極端に困難になり，反復法は，ほぼすべての場合に使わなければならなくなる．

困難さの複合的要因は，関数 $S(\hat{\theta})$ の多くの数値に対応して，複数の解が存在することである．非線形系でパラメータの推定するために用いられてきた方法を論ずることにしよう．

5.2. 非線形系のパラメータ推定

ある種の非線形の問題において，式 (24.1.10) を書き下して反復法を用いて解くのが最も便利である．この方法がうまくゆくかどうかは，方程式の形と用いられる反復法とによって依存する．いくつかの現在使われている方法は，あるコンピュータ計算手順によってパラメータ推定値を得ることが可能である．ここでは，そのうちの 3 つ，(1) 線型化法，(2) 最急勾配法，(3) Marquardt 法について述べよう．

線型化法 (linearization) またはテーラー展開法は，段階的に線形最小 2 乗法の結果を使う．仮定したモデルが，式 (24.1.8) であるとしよう． $\theta_{10}, \theta_{20}, \dots, \theta_{p0}$ が，パラメータ $\theta_1, \theta_2, \dots, \theta_p$ の初期値としよう．これらの初期値は，知的な推測によるか，または，利用可能な情報に基づいた予備的な推定値でもよい．

(たとえば，異なる実験室での似たような方程式を当てはめたときの情報によって示唆された値でもよいし，また実験者の経験や知識に基づいて“大体正しい”と示唆された

値であつてもよい).

これらの初期値は, つぎに説明する連続的な反復によって改良されることを期待している.

もし, $\boldsymbol{\theta}_0 = (\theta_{10}, \theta_{20}, \dots, \theta_{p0})'$ であるようなポイント $\boldsymbol{\theta}_0$ の近傍で $f(\boldsymbol{\xi}_u, \boldsymbol{\theta})$ のテーラー展開を行ない, 一次微分で展開を止めるならば, $\boldsymbol{\theta}$ が $\boldsymbol{\theta}_0$ に近似的に近づくとと言える.

$$f(\boldsymbol{\xi}_u, \boldsymbol{\theta}) = f(\boldsymbol{\xi}_u, \boldsymbol{\theta}_0) + \sum_{i=1}^p \left[\frac{\partial f(\boldsymbol{\xi}_u, \boldsymbol{\theta})}{\partial \theta_i} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\theta_i - \theta_{i0}) \quad (24.2.1)$$

もし,

$$\begin{aligned} f_u^0 &= f(\boldsymbol{\xi}_u, \boldsymbol{\theta}_0), \\ \beta_i^0 &= \theta_i - \theta_{i0}, \\ Z_{iu}^0 &= \left[\frac{\partial f(\boldsymbol{\xi}_u, \boldsymbol{\theta})}{\partial \theta_i} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \end{aligned} \quad (24.2.2)$$

とおけば,, 式(24.1.8) が近似的に次の形に

$$Y_u - f_u^0 = \sum_i^p \beta_i^0 Z_{iu}^0 + \varepsilon_u; \quad (24.2.3)$$

なる. 言い換えれば, 選択した近似まで式 (24.1.1) で示した線形の形となる.

線形最小 2 乗法の理論を適用することにより, 今パラメータ $\beta_i^0, i=1, 2, \dots, p$ を推定することができる.

もし, 次のように書けば,

$$\mathbf{Z}_0 = \begin{bmatrix} \mathbf{Z}_{11}^0 & \mathbf{Z}_{21}^0 & \cdots & \mathbf{Z}_{p1}^0 \\ \mathbf{Z}_{12}^0 & \mathbf{Z}_{22}^0 & \cdots & \mathbf{Z}_{p2}^0 \\ \vdots & \vdots & & \vdots \\ \mathbf{Z}_{1u}^0 & \mathbf{Z}_{2u}^0 & \cdots & \mathbf{Z}_{pu}^0 \\ \vdots & \vdots & & \vdots \\ \mathbf{Z}_{1n}^0 & \mathbf{Z}_{2n}^0 & \cdots & \mathbf{Z}_{pn}^0 \end{bmatrix} = \{\mathbf{Z}_{iu}^0\}, \quad n \times p, \quad (24.2.4)$$

$$\mathbf{b}_0 = \begin{bmatrix} b_1^0 \\ b_2^0 \\ \vdots \\ b_p^0 \end{bmatrix}, \quad \mathbf{y}_0 = \begin{bmatrix} Y_1 - f_1^0 \\ Y_2 - f_2^0 \\ \vdots \\ Y_u - f_u^0 \\ \vdots \\ Y_n - f_n^0 \end{bmatrix} = \mathbf{Y} - \mathbf{f}^0, \quad (24.2.5)$$

$\boldsymbol{\beta}_0 = (\beta_1^0, \beta_2^0, \dots, \beta_p^0)'$ の推定量は、自明な表記で

$$\mathbf{b}_0 = (\mathbf{Z}_0' \mathbf{Z}_0)^{-1} \mathbf{Z}_0' (\mathbf{Y} - \mathbf{f}^0). \quad (24.2.6)$$

によって与えられる。

ベクトル \mathbf{b}_0 は、したがって、2乗和

$$SS(\boldsymbol{\theta}) = \sum_{u=1}^n \left\{ Y_u - f(\boldsymbol{\xi}_u, \boldsymbol{\theta}_0) - \sum_{i=1}^p \beta_i^0 Z_{iu}^0 \right\}^2 \quad (24.2.7)$$

を、 $\beta_i^0 = \theta_i - \theta_{i0}$ であるような $\beta_i^0, i=1, 2, \dots, p$ を最小化するであろう。

$b_i^0 = \theta_{i1} - \theta_{i0}$ と書くとすると、 $\theta_{i1}, i=1, 2, \dots, p$ は、改定された $\boldsymbol{\theta}$ の最良推定量と考えてよい。

適切な非線形モデルを使用した式 (24.1.9) における平方和 $S(\boldsymbol{\theta})$ と、モデルの線形近似展開を用いた式 (24.2.7) における平方和 $SS(\boldsymbol{\theta})$ との違いに注意を払う必要がある。

ここで、上の θ_{i0} と同じ役割を果たす改定された推定量 θ_{i1} に代えることができ、上の式 (24.2.1) から式 (24.2.7) で記述された手順をすべての添字 0 を 1 に置き換え、正確に実行する。これによって、改定された推定量 θ_{i2} が導かれ、以下同様に反復される。ベクトルの形で前の表記法を拡張するならば、つぎのように書くことができる。

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + \mathbf{b}_j, \quad (24.2.8)$$

$$\boldsymbol{\theta}_{j+1} = \boldsymbol{\theta}_j + (\mathbf{Z}_j' \mathbf{Z}_j)^{-1} \mathbf{Z}_j' (\mathbf{Y} - \mathbf{f}^j)$$

ここで、

$$\mathbf{Z}_j = \{Z_{iu}^j\},$$

$$\mathbf{f}^j = (f_1^j, f_2^j, \dots, f_n^j)', \quad (24.2.9)$$

$$\boldsymbol{\theta}_j = (\theta_{1j}, \theta_{2j}, \dots, \theta_{pj})'$$

である。

この反復の手順は、解が収束するまで続けられる。つまり、連続する $j, (j+1)$ の反復において、

$$|\{\theta_{i(j+1)} - \theta_{ij}\} / \theta_{ij}| < \delta, \quad i=1, 2, \dots, p$$

となることで、ここで、 δ は、あらかじめ定められた量 (たとえば 0.000001) である。

反復手順の各段階で、 $S(\theta_j)$ は、その値が実際に減少したかどうかを確かめるために使われる。

線形化の手順は、ある種の問題に対して好ましくない次の性質を持つ可能性がある。

1. その収束はきわめて遅い；すなわち、平方和 $S(\theta_j)$ が j の増加に伴って一貫して減少していても、解が安定するまでにはきわめて多くの反復を必要とする。この種の振る舞いは、時たまおきる。
2. たえず反対方向に広い範囲で振動することがあり、平方和の減少と同様の増大もしばしばおきる。それにもかかわらず、解は結局安定する。
3. 全く収束しないばかりか、発散さえするので、平方和は反復ごとに限りなく増大することがある。

これらの欠陥と戦うために、G. E. P. Box の指導のもとに、Booth と Peterson (1958) によって非線形推定プログラム、IBM, SHARE Program Pa. No. 687 (WLNLI), が書かれた。これは、式 (24.2.8) の中の修正ベクトル \mathbf{b}_j を、もし、

$$S(\theta_{j+1}) > S(\theta_j)$$

ならば、半分にし、

$$S(\theta_{j+1}) < S(\theta_j).$$

ならば、倍にするような修正を行なった。この半分あるいは倍にする手順は、 θ_j と θ_{j+1} の間の3点が、 $S(\theta)$ の局所的な最小値を見出すまで続けられる。2次の補間法が、最小値を見出すために使われ、再び、反復のサイクルが始まる。

理論上、この方法は常に収束する（ハートレー、1961年を参照）が、實際上、困難なことが生じるかもしれない。線形化方法は一般的に有用であり、多くの非線形問題をうまく解決するだろう。それがうまく行かないときに、モデル（24.4節を参照）の再パラメータ化、あるいは、マルカートの折衷法の使用を考慮すべきである。

A Remark on Derivatives. Many computer programs that use a method needing the values of the derivatives of a function at certain points do not use the functional values of the derivatives at all. Instead they compute ratios such as

$$\{f(\xi_u, \theta_{10}, \theta_{10} + h_i, \dots, \theta_{p0}) - f(\xi_u, \theta_{10}, \theta_{10}, \dots, \theta_{p0})\} / h_i,$$

$$i = 1, 2, \dots, p,$$

where h_i , is a selected small increment. The ratio given above is an approximation to the expression

$$\left[\frac{\partial f(\xi_u, \theta)}{\partial \theta_i} \right]_{\theta=\theta_0}$$

since, if h_i tends to zero, the limit of the ratio is this expression by definition.

線形化の幾何学的な解釈 (3版での追加)

平方和の関数 $S(\theta)$ は、 θ だけのパラメータを要素とする関数であり、データは $S(\theta)$ 内の数値係数を用意し、これらは任意の特定の非線形の推定問題のために固定される。母数空間で、すなわち、 p -次元の幾何学的な $\theta_1, \theta_2, \dots, \theta_p$ 空間であり、関数 $S(\theta)$ は、等高線 (the contours of a surface) によって表わすことができる。 θ 's の中でモデルが線形だった場合、最小二乗推定量 $\hat{\theta}$ によって定義された位置 $S(\hat{\theta})$ の等高線は楕円体で、一つの局所的な最小値 (広域的にも一つの) になるであろう。モデルが非線形の場合、等高線は楕円体ではないが、不規則で、しばしば“バナナ形”となし、恐らくいくつかの極小値があり、おそらく1つ以上の広域的な極小があり、それは、 $p=2$ に対する図 24.1 の例での一つ θ -点よりもっと多くの点に到達するかもしれない。

図 24.1a は、 $S(\theta)$ の線形モデルのための等高線“楕円のボウル”を示し、図 24.1b が、非線形モデルのための $S(\theta)$ 等高線を示している。

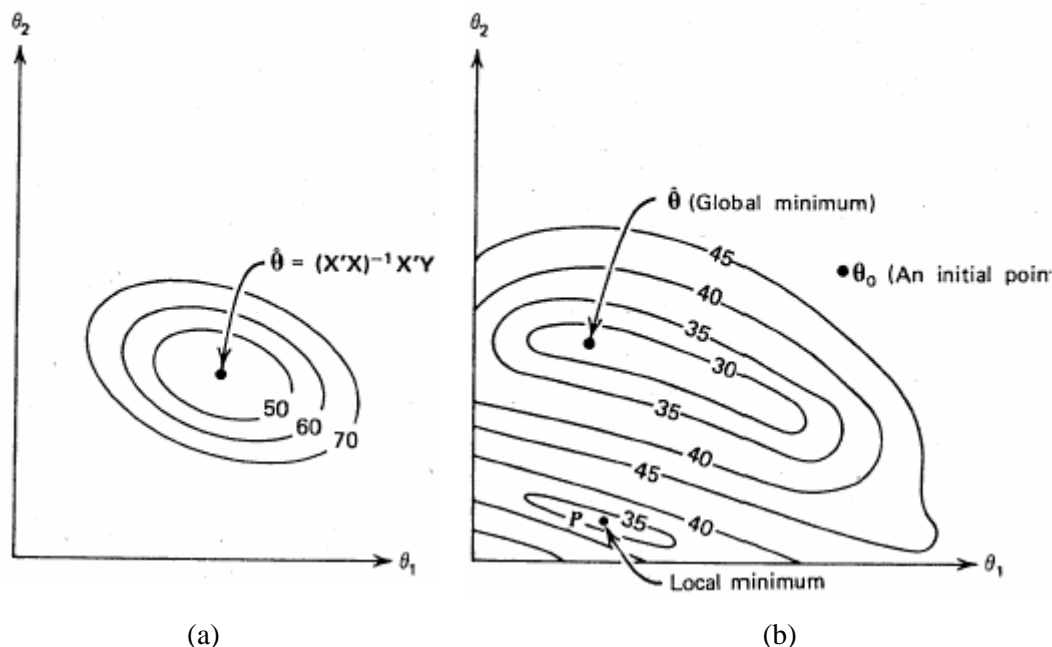


図 24.1. (a) 線形模型 $Y = \theta_1 X_1 + \theta_2 X_2 + \varepsilon$ に付いての楕円状の $S(\theta)$ 等高線; このボウルは唯一の最小点 $\hat{\theta}$ を持っている. (b) 2つの極小値を備えた非線形モデルのための不規則な $S(\theta)$ の等高線. 希望の解は $\hat{\theta}$ であるが、ポイント P のへの反復もおきる. チェックとして、反復法での解は、いくつかのうまく散らばった初期値 θ_0 からスタートして得るべきである.

正確な形および $S(\theta)$ の等高線の方向は、モデルとデータに依存する。最小二乗推定量 $\hat{\theta}$ を囲む等高線は、非常に伸びてて、 $S(\theta)$ の“ボウル高さ”値が、 $S(\hat{\theta})$ に近いようなという意味での $\hat{\theta}$ として、“ほぼよい”ような多くの可能な θ -値が、存在している時には、その問題はたちが悪いと言われ、 $\hat{\theta}$ は、計算上得ることが困難かもしれない。悪く条件は、過剰パラメータ化されたモデルであるとも指摘できる。すなわち、必要であるより多くのパラメータを持っている場合、あるいは、パラメータを推定することを可能としない不適当なデータと見なせる場合である。これらは、表裏一体であり、ある1つか、他方が原因かの選択は、実際のところ各自の予備的知識または視点に依存する。

例えば、式(24.1.3) の $f(t; \theta_1, \theta_2)$ を考えてみよう。これは、 $t=0$ で始まり、 $t=\infty$ で終わる共にゼロである曲線で、どこかに中間にピークをもつ。その曲線の $t=0$ での傾斜は、 θ_1 で、ピークは

$$t_{\text{peak}} = \ln(\theta_1 / \theta_2) / (\theta_1 - \theta_2).$$

である。

データが単に曲線の初めの部分をカバー（図 24.2a を参照）するのであれば、 θ_1 をうまく推定できるが、 θ_2 はできない。後者については、図 24.2b のように曲線がどこでピークに達するかについての情報を得なければならない。1つのパラメータ・モデル $Y = \theta t + \varepsilon$ は 図 24.2a のデータに適切であるが、これらのデータは、2つのパラメータ・モデル、式 (24.1.3) の推定に不適切である。

線形化方法は、始点 $S(\theta)$ からスタートする非線形モデル用の θ_0 の最小値を見つける問題を一連の線形模型問題に変える。 θ_0 に関する $f(\xi, \theta)$ の最初の線形化式 (24.2.1) は、不規則な $S(\theta)$ ボウルを「正しい形を見る」楕円のボウル $SS(\theta)$ に取り替える、すなわち、ちょうど θ_0 で対応するモデル関数の同じ一次導関数を持っている。

我々が 24.3 節の例で見ると、それは、実際の状況（すなわち仮定されたモデル、利用可能なデータ、 θ_0 の相対的位置）、および、 θ 空間でに依存して、実際の $S(\theta)$ の等高線に悪く、あるいは、うまく接近するかもしれない。どんな場合も、解決する“ θ_0 で線形化された” θ_0 （比較的容易な線形の最小二乗法計算）で線形化されたボウルの最下段ポイントに 図 24.3 の中で示されるような θ_1 に達するために移ることによる問題。その後、 θ_1 で全体の線形化プロセスを繰り返す。

我々の望みは、分岐するのではなく、図 24.4 の中で指摘されるように、逐次反復が $\hat{\theta}$ に集中するであろうということである。典型的には、その後、線形化されたものが通常よく実際の等高線に接近するので、反復の始点が $\hat{\theta}$ に接近している場合、線形化はうまくやる。

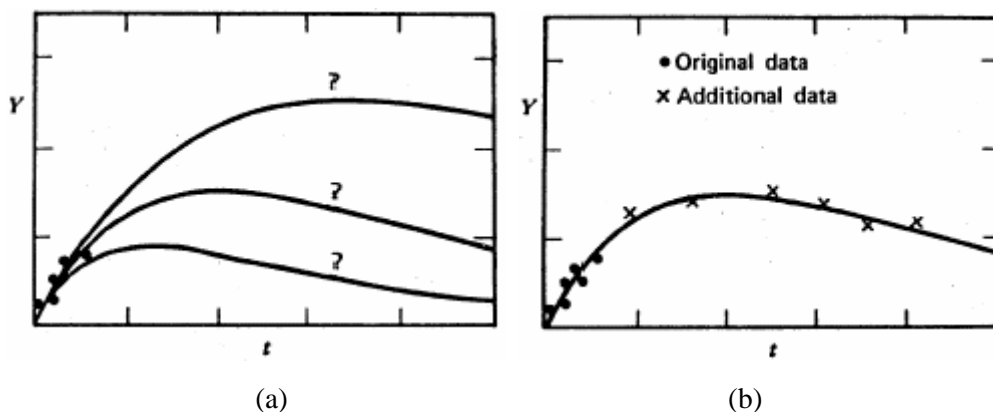


図 24.2. (a) これらのデータは、最初の傾斜 θ_1 を推定することは可能であるが、 θ_2 を含んでいるような曲線がピークになっているところでは可能でない。もし式 (24.1.3) が使用されれば、 $S(\theta)$ の表面はたちが悪いだらう。もし式 (24.1.3) が使用されれば、 $S(\theta)$ の表面はたちが悪くなるであらう。(b) 我々は、示された追加のデータによって式 (24.1.3) 中の θ_1 および θ_2 の両方をうまく評価することができるだろう。また、 $S(\theta)$ 表面は比較的によい状態になるだろう。

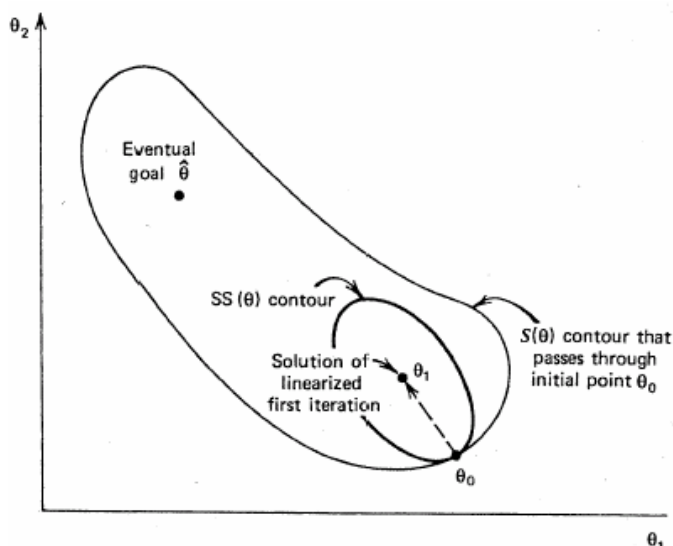


図 24.3. 線形化プロセスでの最初の反復。 $SS(\theta)$ の等高線の楕円は、テキストに記述されたように θ_0 で $S(\theta)$ の等高線接近する。線形化された問題は解 θ_1 を用意する。また、手続きは次にそこから繰り返される。

線形・非線形の最小 2 乗の幾何学的な詳しい記述は、24.5 と 24.6 節で示される。

最急勾配法

略

マルカート法

略

Confidence Contours

Some idea of the nonlinearity in the model under study can be obtained, after the estimation of θ , by evaluating the ellipsoidal confidence region obtained on the assumption that the linearized form of the model is valid around $\hat{\theta}$, the final estimate of θ .

This is given by the formula

$$(\theta - \hat{\theta})' \hat{\mathbf{Z}}' \mathbf{Z} (\theta - \hat{\theta}) \leq ps^2 F(p, n - p, 1 - \alpha),$$

where $\hat{\mathbf{Z}}$ denotes a matrix of the form shown in Eq. (24.2.4) but with $\hat{\theta}$ substituted into the elements in place of θ_0 everywhere, and where

$$s^2 = \mathbf{S}(\theta)/(n - p).$$

Note that when the difference between successive values θ_{j+1} and θ_j is sufficiently small so

that the linearization procedure terminates with $\theta_{j+1} = \hat{\theta}$ (θ_j , for practical purposes), then

$S(\hat{\theta})$ is a minimum value of $S(\theta)$ in Eq. (24.1.9) to the accuracy imposed by the termination procedure selected. This can be seen by examining Eq. (24.2.7) with $\hat{\theta}$, β_i^{j+1} , and \mathbf{Z}_{iu}^{j+1} replacing θ_0 , β_i^0 , and \mathbf{Z}_{iu}^0 , respectively, and remembering that, to the order of accuracy imposed by the termination procedure, $\mathbf{b}_{j+1} = \mathbf{0}$. The ellipsoid above will *not* be a true confidence region when the model is nonlinear. We can, however, determine the end points on the major axes of this ellipsoid by canonical reduction (see, e.g., Box and Draper, 1987). The *actual values* of $S(\theta)$ can be evaluated at these points and compared with each other. Under linear theory the values would all be the same.

An exact confidence contour is denned by taking $S(\theta) = \text{constant}$, but since we do not know the correct distribution properties in the general nonlinear case, we are unable to obtain a specified probability level. However, we can, for example, choose the contour such that

$$S(\boldsymbol{\theta}) = S(\hat{\boldsymbol{\theta}}) \left\{ 1 + \frac{p}{n-p} F(p, n-p, 1-\alpha) \right\},$$

which, if the model is linear, provides an exact, ellipsoidal $100(1-\alpha)\%$ boundary, and label it as an approximate $100(1-\alpha)\%$ confidence contour in the nonlinear case. Note that the contour so determined *will be a proper correct confidence contour in this case* (and will not be elliptical in general), *and it is only the probability level that is approximate*. When only two parameters are involved the confidence contour can be drawn. For more parameters, sectional drawings can be constructed if desired.

In general, when a linearized form of a nonlinear model is used, all the usual formulas and analyses of linear regression theory can be applied. Any results obtained, however, are valid only to the extent that the linearized form provides a good approximation to true model.

Gridst and Plots

略

The Importance of Good Starting Value

略

Getting Initial Estimation θ_0

略

5.3. AN EXAMPLE

(Emax モデルについて、これを参考に詳細に記したので、原著の例題については訳していないが、掲載する)

The example that follows is taken from an investigation performed at Procter & Gamble and reported by Smith and Dubey (1964). We shall use this example to illustrate how a solution can be obtained for a nonlinear estimation problem by solving the normal equations, directly, or alternatively by the linearization method. We shall not provide an example of the use of steepest descent (but see, e.g., Box and Coutie, 1956), nor an example of Marquardt's compromise procedure. The investigation involved a product A, which must have a fraction of 0.50 of Available Chlorine at the time of manufacture. The fraction of Available Chlorine in the product

decreases with time; this is known. In the 8 weeks before the product reaches the consumer a decline to a level 0.49 occurs but since many uncontrolled factors then arise (such as warehousing environments, handling facilities), theoretical calculations are not reliable for making extended predictions of the Available Chlorine fraction present at later times. To assist management in decisions--such as (1) When should warehouse material be scrapped? and (2) When should store stocks be replaced?--cartons of the product were analyzed over a period to provide the data of Table 24.2. (Note that the product is made only every other week and code-dated only by the week of the year. The predicted values shown in the table are obtained from the fitted equation to be found in what follows.) It was postulated that a nonlinear model of the form

$$Y = \alpha + (0.59 - \alpha)e^{-\beta(X-8)} + \varepsilon \tag{24.3.1}$$

would suitably account for the variation observed in the data, for $X \geq 8$. This model provides a true level, without error, of $\eta = 0.49$ when $X = 8$, and it exhibits the proper sort of decay. An additional point of information, agreed upon by knowledgeable chemists, was that an equilibrium, asymptotic level of Available Chlorine somewhere above 0.30 should be expected.

T A B L E 24.2. Percent of Available Chlorine in a Unit of Product

Length of Time Since Produced (weeks) X	Available Chlorine Y	Average Available Chlorine \bar{Y}	Predicted Y , Using the Model \hat{Y}
8	0.49, 0.49	0.490	0.490
10	0.48, 0.47, 0.48, 0.47	0.475	0.472
12	0.46, 0.46, 0.45, 0.43	0.450	0.457
14	0.45, 0.43, 0.43	0.437	0.445
16	0.44, 0.43, 0.43	0.433	0.435
18	0.46, 0.45	0.455	0.427
20	0.42, 0.42, 0.43	0.423	0.420
22	0.41, 0.41, 0.40	0.407	0.415
24	0.42, 0.40, 0.40	0.407	0.410
26	0.41, 0.40, 0.41	0.407	0.407
28	0.41, 0.40	0.405	0.404
30	0.40, 0.40, 0.38	0.393	0.401
32	0.41, 0.40	0.405	0.399
34	0.40	0.400	0.397
36	0.41, 0.38	0.395	0.396
38	0.40, 0.40	0.400	0.395
40	0.39	0.390	0.394
42	0.39	0.390	0.393

The problem is to estimate the parameters α and β of the nonlinear model given in Eq. (24.3.1) using the data given in the table. The residual sum of squares for this model can be written as

$$S(\alpha, \beta) = \sum_{u=1}^n [Y_u - \alpha - (0.59 - \alpha)e^{-\beta(X_u-8)}]^2 \quad (24.3.2)$$

where $(X_u, Y_u), u = 1, 2, 3, \dots, 44$, are the corresponding pairs of observations from the table (e.g., $X_1 = 8, Y_1 = 0.49, \dots, X_{44} = 42, Y_{44} = 0.39$).

A Solution Through the Normal Equations

略

A Solution Through the Linearization Technique

To linearize the model into the form of Eq. (24.2.1) we need to evaluate the first derivatives of

$$\begin{aligned} f(\xi_u, \theta) &= f(X_u; \alpha, \beta) \\ &= \alpha + (0.49 - \alpha)e^{-\beta(X_u-8)}, \end{aligned} \quad (24.3.11)$$

namely,

$$\begin{aligned} \frac{\partial f}{\partial \alpha} &= 1 - e^{-\beta(X_u-8)} \\ \frac{\partial f}{\partial \beta} &= -(0.49 - \alpha)(X_u - 8)e^{-\beta(X_u-8)}. \end{aligned} \quad (24.3.12)$$

Thus if $\alpha = \alpha_j, \beta = \beta_j$ are the values inserted at the j th stage, as described in Section 24.2, we have, in the notation implied in that section, a model of form (at the j th stage)

$$\begin{aligned} Y_u - f_u^j &= [1 - e^{-\beta_j(X_u-8)}](\alpha - \alpha_j) \\ &\quad + [-(0.49 - \alpha_j)(X_u - 8)e^{-\beta_j(X_u-8)}](\beta - \beta_j) + \varepsilon \end{aligned}$$

or in matrix form

$$Y - \mathbf{f}^j = \mathbf{Z}_j \begin{bmatrix} \alpha - \alpha_j \\ \beta - \beta_j \end{bmatrix} + \boldsymbol{\varepsilon},$$

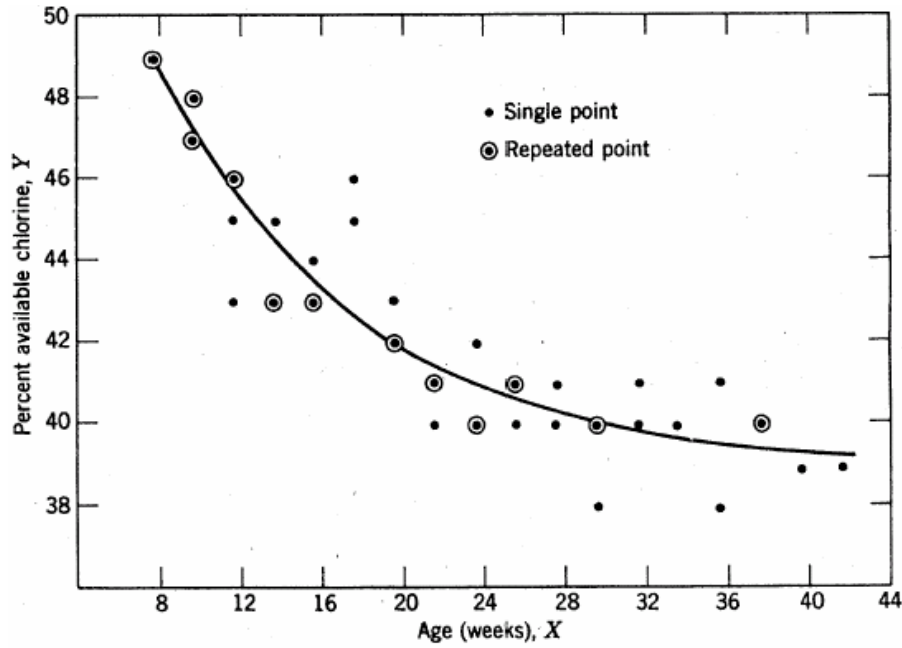


Figure 24.6. The fitted curve and the observations.

where

$$f_u^j = \alpha_j + (0.49 - \alpha_j)e^{-\beta(X_u - 8)} \quad (24.3.13)$$

$$\mathbf{Z}_j = \begin{bmatrix} 1 - e^{-\beta(X_1 - 8)} & -(0.49 - \alpha)(X_1 - 8)e^{-\beta(X_1 - 8)} \\ \vdots & \vdots \\ 1 - e^{-\beta(X_u - 8)} & -(0.49 - \alpha)(X_u - 8)e^{-\beta(X_u - 8)} \\ \vdots & \vdots \\ 1 - e^{-\beta(X_n - 8)} & -(0.49 - \alpha)(X_n - 8)e^{-\beta(X_n - 8)} \end{bmatrix}$$

and the vector of quantities to be estimated is

$$\begin{bmatrix} \alpha - \alpha_j \\ \beta - \beta_j \end{bmatrix} \quad (24.3.15)$$

with estimate given by

$$\begin{bmatrix} \alpha - \alpha_j \\ \beta - \beta_j \end{bmatrix} = (\mathbf{Z}'_j \mathbf{Z}_j)^{-1} \mathbf{Z}'_j \begin{bmatrix} Y_1 - f_1^j \\ Y_2 - f_2^j \\ \vdots \\ Y_n - f_n^j \end{bmatrix} \quad (24.3.16)$$

If we begin the iterations with initial guesses $\alpha_0 = 0.30$ and $\beta_0 = 0.02$ as before, and apply Eq. (24.3.16) iteratively, we obtain estimates as follows:

Iteration	α_j	β_j	$S(\alpha_j, \beta_j)$
0	0.3000	0.2000	0.0263
1	0.8416	0.1007	4.4881
2	0.3901	0.1004	0.0050
3	0.3901	0.1016	0.0050
4	0.3901	0.1016	0.0050

Note: These figures were rounded from the end results of computer calculations, which carried more significant figures. Numerical differences might occur if a parallel calculation were made on a pocket calculator.

This process converges to the same least squares estimates as before, namely, $\hat{\alpha} = 0.39$ and $\hat{\beta} = 0.10$ to give the fitted model, Eq. (24.3.10). Note that this happens in spite of the rather alarming fact that, after the first stage, $S(\alpha_1, \beta_1) = 4.4881$, which is about 170 times the initial $S(\alpha_0, \beta_0) = 0.0263$. The reduction in the next iteration is dramatic and practically final, the subsequent reduction in $S(\alpha, \beta)$ being in the sixth place of decimals, which is not shown. In some nonlinear problems no correction, dramatic or otherwise, occurs and the process diverges, providing larger and larger values for $S(\theta)$. (For a possible reason for this sort of behavior, see Section 24.6.)

Figure 24.7_a shows the nonlinear sum of squares contours $S(\alpha, \beta)$ in the region $0 \leq \beta \leq 0.20$, $0.28 \leq \alpha \leq 0.90$ together with the progress from the initial point (β_0, α_0) to (β_2, α_2) . The reason for this path is shown in Figures 24.7b and 24.7c. In Figure 24.7b the $SS(\theta)$ contours are relatively ill-conditioned and lead to a minimum $SS(\alpha, \beta)$ value at a point (β_1, α_1) well away from the actual minimum of $S(\alpha, \beta)$. In the next iteration the linearized sum of squares contours are relatively well-conditioned (Figure 24.7c) and also their center happens to be close to the minimum point of the actual sum of squares contours, as we have seen. Note that, in both Figures 24.7b and 24.7c, the direction of steepest descent, which would be perpendicular to the $S(\alpha, \beta)$ contours at the starting point, would take us off in a different direction from that determined by linearization.

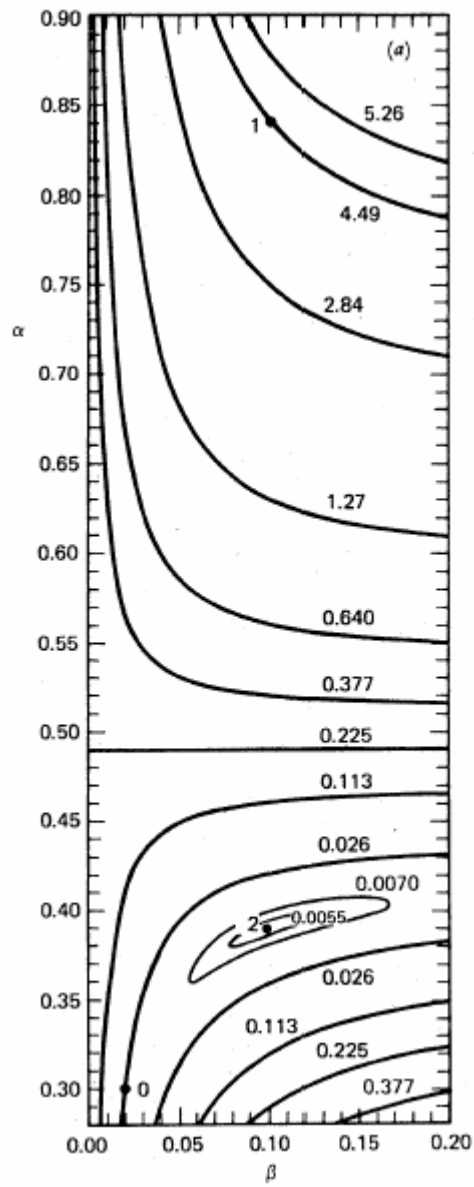


Figure 24.7.(a) Nonlinear sum of squares contours $S(\alpha, \beta)$ $S(a, j3)$ for the example. The points marked 0, 1, 2 are the (β_j, α_j) for the initial point and the first and second iterations.

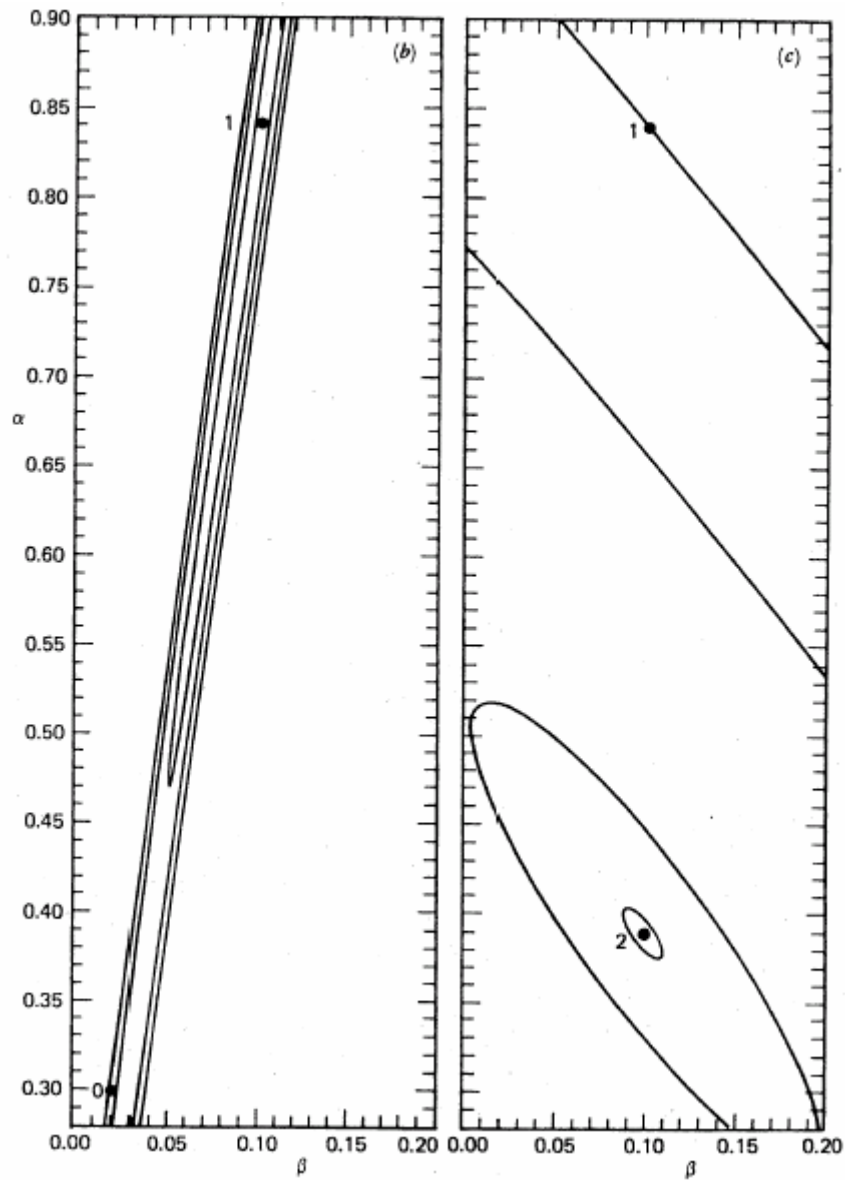


Figure 24.7, (Continued) (b) Linearized sum of squares contours $SS(\alpha, \beta)$ at the point $(\beta, \alpha) = (\beta_0, \alpha_0) = (0.02, 0.30)$. The center of the elliptical system is at $(\beta_1, \alpha_1) = (0.1007, 0.8416)$ and the next iteration begins from there. (c) Linearized sum of squares contours $SS(\alpha, \beta)$ at the point $(\beta, \alpha) = (\beta_1, \alpha_1) = (0.1007, 0.8416)$. The center of the elliptical system is at $(\beta_2, \alpha_2) = (0.1004, 0.3901)$ and the next iteration begins from there.

Further Analysis

略

Confidence Regions

We can calculate approximate $100(1 - q)\%$ confidence contours (described in Section 24.2) by finding points (α, β) that satisfy

$$\begin{aligned} S(\alpha, \beta) &= S(\hat{\alpha}, \hat{\beta}) \left[1 + \frac{p}{n-p} F(p, n-p, 1-q) \right] \\ &= 0.0050 [1 + F(2, 42, 1-q) / 21] \\ &= S_q \end{aligned}$$

以下 略

Some Typical Nonlinear Program Output Features

The printed output from nonlinear least squares programs varies from program to program. However, some features are common to many outputs and we now describe and illustrate these briefly, using the output from a Marquardt-type analysis of the example given in this section.

1. The *singular values* are the square roots of the eigenvalues of $\mathbf{Z}'_j \mathbf{Z}_j$, where \mathbf{Z}_j is derived from Eq. (24.2.4) but employs the current θ_j value and not θ_0 . Widely disparate singular values indicate a tendency toward ill-conditioning. The corresponding singular vectors provide the orientation of the axes of the “linearization approximation bowl” with respect to the θ -axes.

Example 1. At $\hat{\theta}$ we have the following:

Singular values:	4.964	0.773
α :	-0.946	0.326
β :	0.326	0.946

The ratio of the largest and smallest singular values (there are only two singular values here, of course, but in general there would be p) is about 6.4, implying a fairly well conditioned problem. (In an ill-conditioned problem, this ratio can be in the thousands. A ratio of 1 would imply circular linearizing approximating contours.) To achieve the same change in $SS(\theta)$, one would have to move about 6.4 units in the β direction compared with one unit in the α direction. Reading down the columns, we see that the first singular direction

$$\phi_1 = -0.946\alpha + 0.326\beta$$

is mostly in the $_{-}$ -a direction (the sign is not important here and is determined by the direction of the labeling only), while

$$\phi_2 = 0.326\alpha + 0.946\beta$$

is mostly in the $-\beta$ direction, as Figure 24.8 confirms should be the case. Note that the sum of squares of the coefficients (e.g., -0.946 and 0.326) is 1, apart from rounding error, and this sort of result is true in general for all the columns.

2. The *normalizing elements* are the square roots of the diagonal entries of the $\mathbf{Q} = ((q_{il})) = (\mathbf{Z}'_j \mathbf{Z}_j)^{-1}$ matrix. The linearized approximate *correlation matrix* $\mathbf{C} = ((c_{il}))$ of the $\hat{\theta}$'s is obtained by dividing each row and column of $(\mathbf{Z}'_j \mathbf{Z}_j)^{-1}$ by the two corresponding normalizing elements, which correspond to that row and column, namely, $c_{il} = q_{il} / (q_{ii} q_{ll})^{1/2}$.

Example 2.

Parameter:	α	β
Normalizing elements:	0.462	1.226
Parameter: α	1.000	
β	0.888	1.000

The correlation between α and β is positive, indicating that variations in (α, β) from $(\hat{\alpha}, \hat{\beta})$ in such a way that *both* increase, or *both* decrease, will give roughly similar $SS(\theta)$ values. We can see from the $S(\alpha, \beta)$ plot (Figure 24.8) how this can happen. [Remember that the $SS(\theta)$ ellipses around $\hat{\theta}$ will mimic the $S(\theta)$ contours but not represent them perfectly.] Although 0.888 is a high correlation, it is not high enough for us to feel that the model is overparameterized. (For more on this point, see the “overparameterization” subsection, below.)

3. Confidence limits for the true values of the θ 's can be evaluated on the basis of the linearized approximation, evaluated at $\hat{\theta}$. The limits are, typically, $\hat{\theta}_i \pm 2\text{se}(\hat{\theta}_i)$, where $\text{se}(\hat{\theta}_i) = \{\text{appropriate diagonal element of } (\hat{\mathbf{Z}}'_j \hat{\mathbf{Z}}_j)^{-1} s^2\}^{1/2}$, where $\hat{\mathbf{Z}}'$ is the \mathbf{Z} of Eq. (24.2.4) with $\hat{\theta}$ replacing θ_0 and $s^2 = S(\hat{\theta}) / (n - p)$. These are, roughly, 95% limits in general.

Example 3.

	Lower Limit	Final Value	Upper Limit
α	0.380	0.390	0.400
β	0.075	0.102	0.128

The final estimate of α is $\hat{\alpha} = 0.390$ and an approximate 95% confidence band for α is (0.380, 0.400). The final estimate of β is $\hat{\beta} = 0.102$ with an approximate 95% confidence band for β of (0.075, 0.128). Note that the message implied by the width of the bands is confirmed by Figure 24.8. (The reader may wish to reread Section 5.4. Not only is the caution there valid but also the confidence statements made here apply only to the extent that the linearized model is a reliable approximation to the nonlinear model) Both bands exclude zero, indicating nonzero values for α and β .

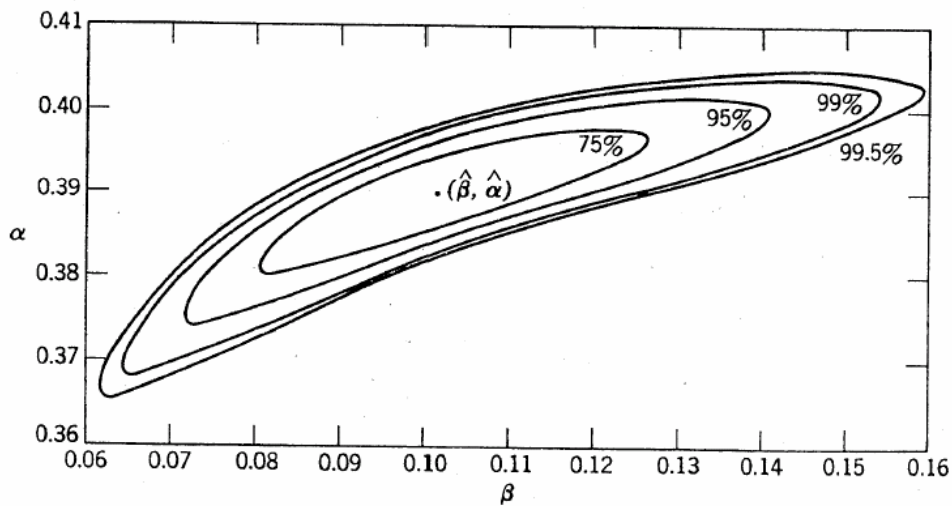


Figure 24.8. Confidence regions for (β, α) . The regions are exact, the confidence levels approximate.

5.4. ガウス・ニュートン法による逐次計算, ドレーパー・スミスの事例

(JMP スクリプトで分散共分散行列の追加)

JMPのスク립ト言語での結果

```
// The Nonlinear regression model Dreper & Smith    2003-1-21    Y.Takahashi

x=[8, 8, 10, 10, 10, 10, 12, 12, 12, 12, 14, 14, 14, 16, 16, 16, 18, 18, 20, 20, 20, 22, 22, 22, 24, 24, 24, 26, 26, 26, 28, 28, 30, 30, 30, 32, 32, 34, 36, 36, 38, 38, 40, 42] ;

y=[0.49, 0.49, 0.48, 0.47, 0.48, 0.47, 0.46, 0.46, 0.45, 0.43, 0.45, 0.43, 0.43, 0.44, 0.43, 0.43, 0.46, 0.45, 0.42, 0.42, 0.43, 0.41, 0.41, 0.4, 0.42, 0.4, 0.4, 0.41, 0.4, 0.41, 0.41, 0.4, 0.4, 0.4, 0.38, 0.41, 0.4, 0.4, 0.41, 0.38, 0.4, 0.4, 0.39, 0.39] ;

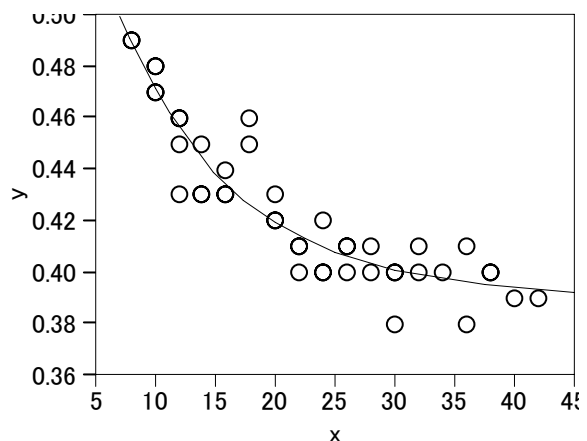
A = x || y ;
b = [0.30 , 0.02 ] ;    show(round(b, 4)) ;

for( i=1, i<=4, i++,
    alpha=b[1] ;
    beta=b[2] ;

    fx      = alpha+(0.49-alpha) :* exp(-beta :* (x-8)) ;
    d_alpha = 1-exp(-beta :* (x-8)) ;
    d_beta  = -(0.49-alpha) :* (x-8) :* exp(-beta :* (x-8)) ;
    z = d_alpha || d_beta ;

    delta=inverse(z`*z)*z`*(y-fx) ;
    b=b+delta ; show(round(b, 4)) ;
);
```

```
Round(b, 4): [0.3, 0.02]
Round(b, 4): [0.8416, 0.1007]
Round(b, 4): [0.3901, 0.1004]
Round(b, 4): [0.3901, 0.1016]
Round(b, 4): [0.3901, 0.1016]
           alpha  beta
```



```
s = (y-fx)`*(y-fx) ;    show(round(s, 6)) ;
n=nrow(z) ;
p=ncol(z) ;
sigma2= s / (nrow(z)-ncol(z)) ;    show(round(sigma2, 6)) ;
```

```

zpz=z`*z ;
cov=inv(z`*z) :* sigma2 ; show(round(cov,6)) ;
Round(s, 6):[0.005002]
Round(sigma2, 6):[0.000119]
Round(cov, 8):
[ 0.00002545 0.00005984,
  0.00005984 0.0001785 ] /* 分散共分散 */

```

```

SE(alpha)=0.005045      0.007736
                0.007736 SE(beta)=0.01336

```

JMPの非線形モデルでの結果

解				
	SSE	DFE	MSE	RMSE
	0.0050016796	42	0.0001191	0.0109127
パラメータ	推定値	近似標準誤差	下側信頼限界	上側信頼限界
alpha	0.3901398407	0.00504641	0.37794914	0.39916741
beta	0.1016321795	0.01335831	0.07679523	0.13207624

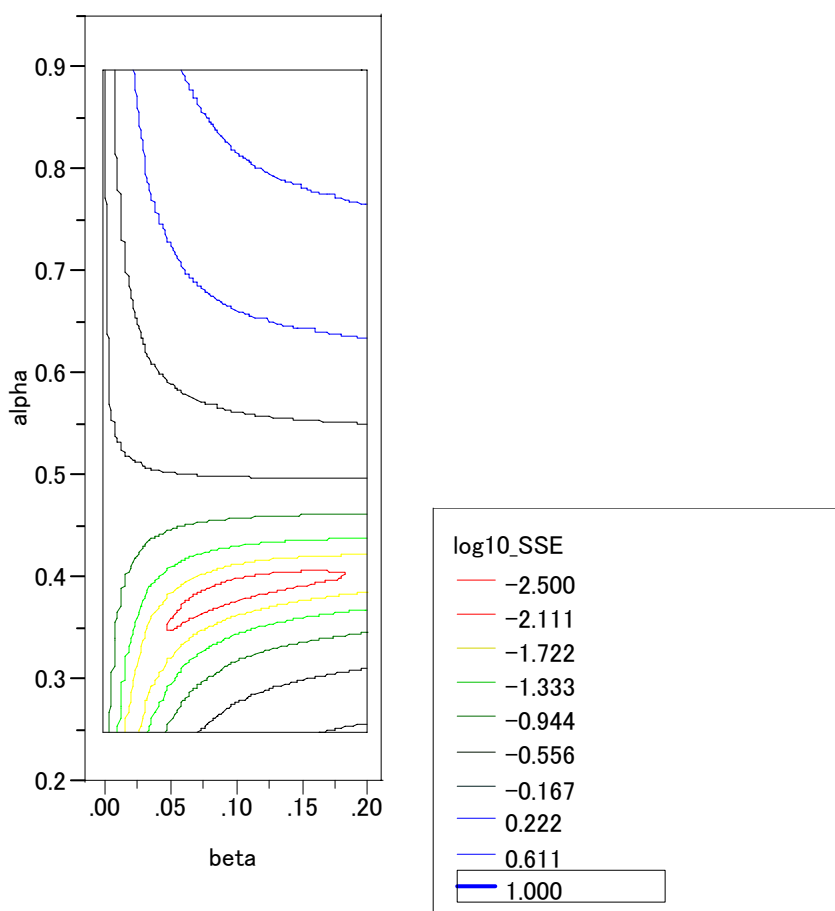
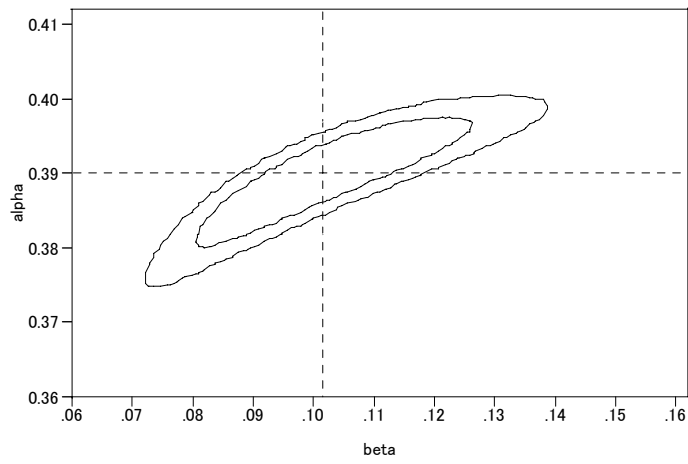
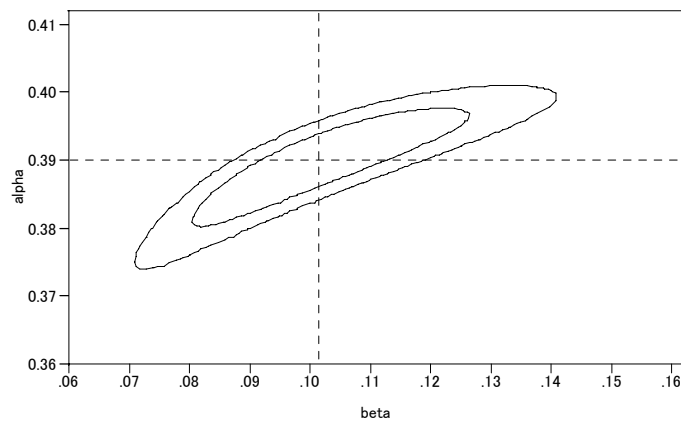


図 0.1 残差平方和 S の等高線図



75%, 99.5%



75%, 95%

図 0.2 パラメータの信頼区間の等高線

SAS/NLINでの結果

```
proc nlin data=d01 method=gauss listall ;
  parms alpha=0.30 beta=0.02 ;
  model y = alpha + (0.49 - alpha) * exp(-beta*(x-8)) ;
run ;
```

The NLIN Procedure Dependent Variable y Method: Gauss-Newton
Iterative Phase

Iter	alpha	beta	Sum of Squares
0	0.3000	0.0200	0.0263
1	0.3169	0.0225	0.0259
2	0.3312	0.0253	0.0253
3	0.3552	0.0313	0.0252
4	0.3806	0.0423	0.0249
5	0.4097	0.0729	0.0245
6	0.3968	0.1023	0.00594
7	0.3902	0.1016	0.00500
8	0.3901	0.1016	0.00500
9	0.3901	0.1016	0.00500

Estimation Summary

Method	Gauss-Newton
Iterations	9
Subiterations	9
Average Subiterations	1
R	1.511E-7
PPC(beta)	1.287E-7
RPC(beta)	5.613E-6
Object	1.42E-10
Objective	0.005002
Observations Read	44
Observations Used	44
Observations Missing	0

Source	DF	Sum of Squares	Mean Square	F Value	Approx Pr > F
Regression	2	7.9820	3.9910	33513.1	<.0001
Residual	42	0.00500	0.000119		
Uncorrected Total	44	7.9870			
Corrected Total	43	0.0395			

Parameter	Estimate	Approx Std Error	Approximate 95% Confidence Limits	
alpha	0.3901	0.00504	0.3800	0.4003
beta	0.1016	0.0134	0.0747	0.1286

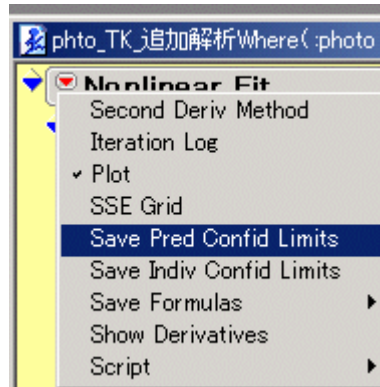
Approximate Correlation Matrix

	alpha	beta
alpha	1.000000	0.8878601
beta	0.8878601	1.000000

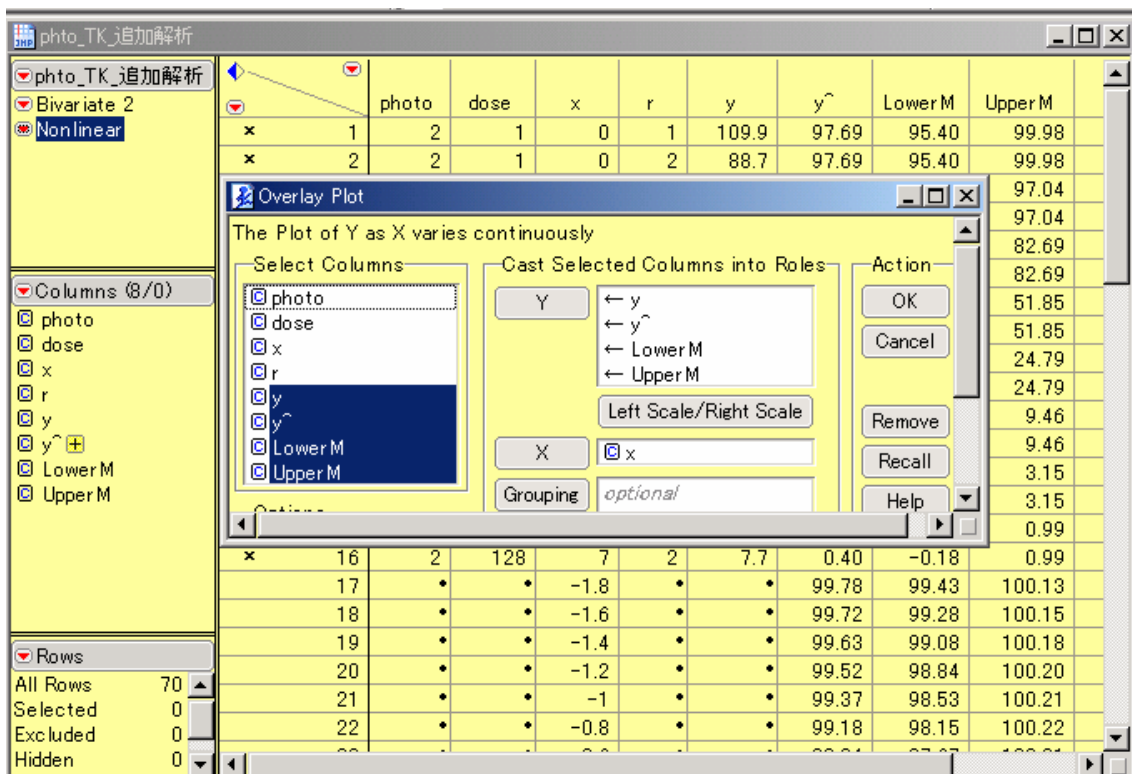
6. 光毒性試験データについての JMP による追加解析

6.1. ロジスティック曲線の推定値の 95%信頼区間

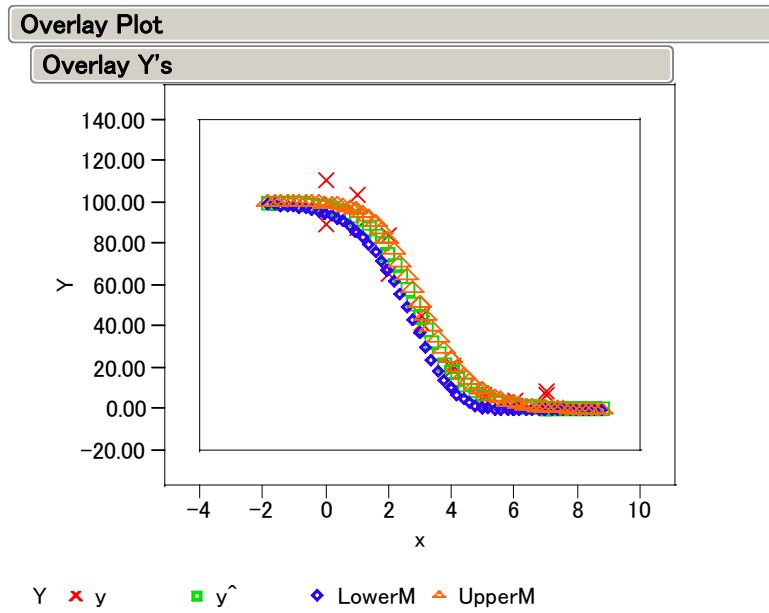
JMP データファイルへ 95%信頼区間の出力



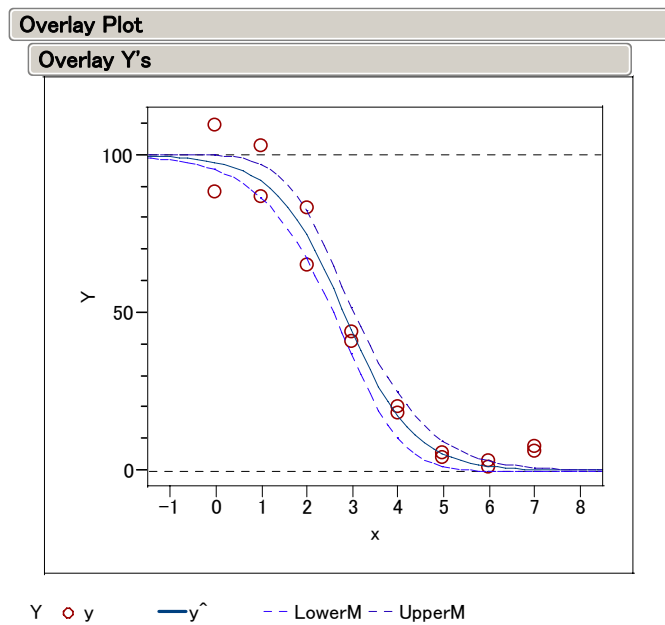
オーバーレイ・プロットの活用



編集前のオーバーレイ・プロット



編集後のオーバーレイ・プロット



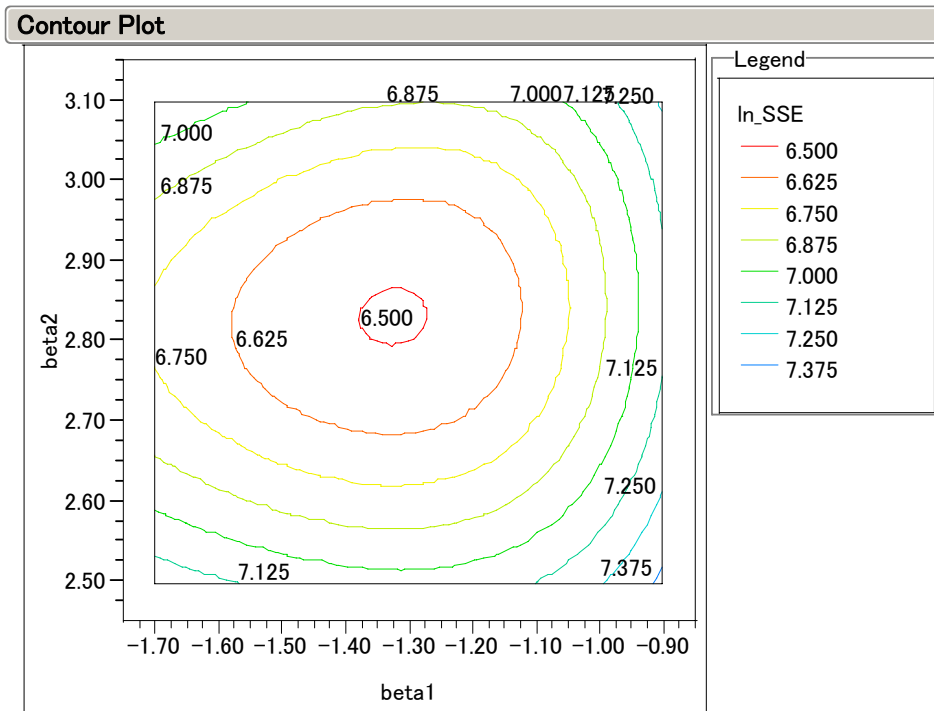
残差 2 乗和のグリッド計算の指定

Specify Grid for Output			
Parameter	Min	Max	N Points
beta1	-1.7	-0.9	50
beta2	2.5	3.1	50

グリッド計算の結果、50×50 ポイント

	beta1	beta2	SSE	ln_SSE
1	-1.70	2.50	1343.69	7.20
2	-1.70	2.51	1306.12	7.17
3	-1.70	2.52	1270.01	7.15
4	-1.70	2.54	1235.35	7.12
5	-1.70	2.55	1202.14	7.09
6	-1.70	2.56	1170.39	7.07
7	-1.70	2.57	1140.10	7.04
8	-1.70	2.59	1111.27	7.01
9	-1.70	2.60	1083.89	6.99
10	-1.70	2.61	1057.97	6.96

残差 2 乗和 S_T の自然対数についての等高線図



SASによる残差2乗和のグリッド計算

```

Title 'photo_TX_S_T 2004-1-20 Y.Takahashi' ;

data d01 ;
  array x{1:16} (0, 0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7) ;
  array y{1:16} (109.9, 88.7, 102.9, 86.8, 65.1, 83.3, 44.2, 41.3,
                18.6, 20.7, 5.8, 4.3, 3.4, 1.2, 6.4, 7.7) ;

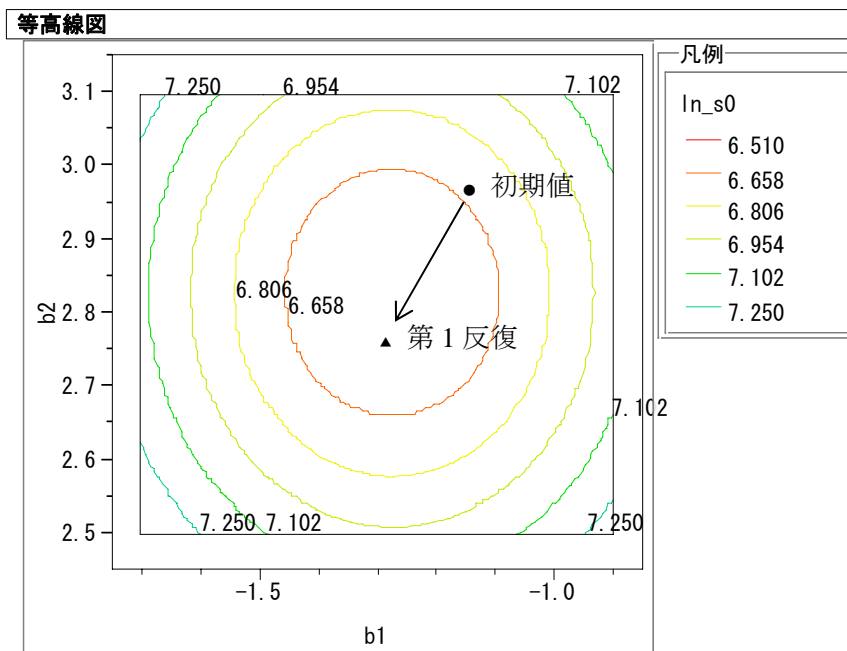
  b11=-1.7 ; b12=-0.9 ; b13=0.01 ;
  b21= 2.5 ; b22= 3.1 ; b23=0.01 ;

  do b1=b11 to b12 by b13 ;
    do b2=b21 to b22 by b23 ;
      s=0 ;
      do i=1 to 16 ;
        fx=100 / (1+exp(b1*(b2-x[i]))) ;
        s=s+(y[i]-fx)**2 ;
      *      output ;
      end ;
      ln_s = log(s) ;
    output ;
  end ;
end ;
keep b1 b2 s ln_s ;

proc print data=d01 ;
run ;

```

6.2. 第1反復での線形化によるパラメータの推定



SASによる線形化による最小2乗解

```
Title 'photo_TX_SS0 2004-2-6 Y.Takahashi' ;

data d01 ;
  array x{1:16} (0, 0, 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7) ;
  array y{1:16} (109.9, 88.7, 102.9, 86.8, 65.1, 83.3, 44.2, 41.3,
                18.6, 20.7, 5.8, 4.3, 3.4, 1.2, 6.4, 7.7) ;

  b11=-1.7 ; b12=-0.9 ; b13=0.01 ;
  b21= 2.5 ; b22= 3.1 ; b23=0.01 ;

  b10=-1.1 ;
  b20= 3 ;

  do b1=b11 to b12 by b13 ;
    do b2=b21 to b22 by b23 ;
      s0=0 ;
      do i=1 to 16 ;
        eta0 = b10 * (b20 - x[i]) ;
        fx0 = 100 / (1 + exp (eta0)) ;
        z10 = (- (b20 - x[i]) * exp(eta0) * fx0 ) / (1 + exp(eta0)) ;
        z20 = - b10 * exp(eta0) * fx0 / (1 + exp(eta0)) ;
        z0 = ((b1-b10)*z10 + (b2-b20)*z20) ;
        s0 = s0 + (y[i] - fx0 - z0)**2 ;
        * output ;
      end ;
      ln_s0 = log(s0) ;
    output ;
  end ;
end ;
keep b1 b2 s0 ln_s0 ;

proc print data=d01 ;
run ;;
```

6.3. パラメータの信頼区間

\hat{S}	n	p	q	F _{1,q}	S
659.58	16	2	0.25	1.53	804.04
659.58	16	2	0.05	3.74	1011.88
659.58	16	2	0.01	6.51	1273.45
659.58	16	2	0.005	7.92	1406.00

