

TR - H - 118

一般化線形モデルによる発火頻度解析
**Analysis of Neural Firing Frequency
by a Generalized Linear Model**

川人 光男
Mitsuo Kawato

1995. 1. 9

ATR人間情報通信研究所

〒619-02 京都府相楽郡精華町光台2-2 ☎ 0774-95-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-774-95-1011

Facsimile: +81-774-95-1008

© (株)ATR人間情報通信研究所

一般化線形モデルによる発火頻度解析

川人光男

(株) エイ・ティ・アール人間情報通信研究所,
〒619-02 京都府相楽郡精華町光台2-2

Abstract

少ないトレース数や、低い発火頻度のデータの場合にも使える、発火頻度解析の一般化線形モデルを提案する。これは、次のような仮定に基づいている。限られたトレース数 m で、観測される発火スパイク数 Y は事象の生起確率 π の二項分布にしたがっている。生起確率 π は、ある時間の関数の線形結合に、ロジスティック関数の逆関数であるシグモイド関数を施したもので表わす。線形結合のパラメータの最尤推定値は、少ない回数で収束する繰り返しアルゴリズムで求めることができる。推定値の分散（信頼度）も簡単に評価できる。 Y を無限大にした極限では、モデルの信頼性を尤度比（もしくはデバイアンス）で評価することができる。この条件が満たされない場合にも、2つの階層的モデルのどちらが、よりよいモデルかは、2つのモデルの対数尤度の差の2倍が、パラメータ数の差の自由度を持つカイ自乗分布になるので、統計的に決めることができる。

1 はじめに

脳において、何が情報のキャリアーになっているかは、神経系の機能を知る上で、最も基本的で重要な問題の一つである。しかしこの間に対するだれもが納得する答えはいまだ得られていない。ニューロンの発火頻度がキャリアーなのか、一つ一つのインパルスが情報を符号化しているのか [1]、複数のニューロンの発火の同期が大切なのか [9]、あるいは発火頻度の時間波形の主成分分析の係数がキャリアーなのか [6]、さまざまな仮説が提出されてきた。多くの（保守的な？）研究者は、各時刻各時刻の、短時間発火頻度がキャリアーであると考えているようである。しかし、この考えを直接支持するようなデータは中枢神経系、特に脳では得られていなかった。末梢神経系などで、刺激の強度や発生しているトルクと、かなり長い時間巾（数秒のオーダー）で平均をとった発火頻度の間に、ある種の比例関係があるというのが間接的な支持データであった。

神経系の入力や出力からある程度はなれた脳の中枢部では、あるニューロンが純粋に物理的に定義できるような1つの変数だけを符号化していることはありそうにない。すると、ニューロンの瞬時の発火頻度が何か物理的に定式化できる情報を表現していたとしても、それを実験的に証明するために発火頻度の時間波形を再構成するような、時系列のモデルが必要となる。設楽ら [7] の研究はこのような方向の口火を切ったものであり、初めて、各時刻各時刻の、短時間発火頻度がキャリアーであることを直接支持するようなデー

タを与えた。しかし、そこで用いられた解析手法は、自発放電の発火頻度が高く、多数のトレース数が得られる実験系では有効なものであったが、一般に適用するのは難しい。そこで、本稿では、あるニューロンの短時間の発火頻度が何か物理的に定式化できる量を符号化しているか否かを統計的に検定できるような一般的な手法を与えることを目指す。

過去1年間の、ニューロンの発火頻度の解析手法の再検討は、次のような展開を遂げてきたと要約できる。(1) 線形モデルによる最小自乗推定。(2) モデル化誤差に含まれる、確率的な成分の二項分布による評価。つまり、発火スパイク数は、試行回数がトレース数 m で、事象の生起確率 π が線形モデルで与えられる二項分布にしたがっているとす。 (3) 発火スパイク数が上記の二項分布にしたがっているとする仮説の検定。(4) 二項分布の確率を線形モデルで与え、そのパラメータを最尤推定法によって求めること。

発火頻度の時間パターンの解析の技法にはさまざまな要素があげられる。第一に、ともかくも、観測された時間パターンを良く再現できなければならない。第二に、そのために、モデルに含まれる小数個のパラメータ値を、最もフィットがよくなるように決めなければならない。第三に、そのためには、フィットの良さを客観的にはかる数量的な基準を導入しなければならない。第四に、推定されたパラメータにどれほどの信頼性があるかを評価しなければならない。第五に、モデルそのものの信頼性を統計的に評価するか、あるいは少なくとも異なるモデルのどちらがよりよいモデルであるかを評価する統計的基準を導入しなければならない。

サルの小脳傍片葉のプルキンエ細胞の発火頻度解析 [7] に用いてきた、線形モデル、最小自乗推定、決定係数はそれぞれ上の、第1の要素、第2、4の要素、第3、5の要素への解答になっていた。残念ながらこの方法はトレース数が少ないとき、またトレース数が多くても発火頻度が低い場合には使えない。そこで本稿では、このような条件でも使え、しかも従来の方が使える場面では、従来法と比例定数を除いて同じ結果を与える、一般化線形モデルによる解析手法を提案する。

2 これまでの解析手法の検討

2.1 線形モデルによる最小自乗推定

従来の解析手法：線形モデルによる最小自乗推定は次のような考えに基づいていた。0から1の間の値をとる発火頻度を、あたかも負の無限大から、正の無限大までの値をとる実数値であるかのようにみなして、逆ダイナミクス of 式など時系列データのモデル式を用いて再構成する。このとき、モデルのフィットの良さは、決定係数で評価する。

この手法は、発火頻度が0や1からはなれていて、かつ発火頻度のデータにのってくる誤差が正規分布に従うと仮定できる場合には合理的であると考えられる。しかし、トレース数が少ないか、発火頻度が低いか、あるいは両方の理由で、1ミリ秒の時間ビンにはいるスパイク数が小さいときは、上の2つの前提条件が崩れてしまう。つまり、発火確率は0に近くなるし、誤差は正規分布に従わない。

2.2 モデル化誤差の二項分布による評価

目で見たとときには良く波形を再構成しているように見えるのに、トレース数が少ないか、もしくは発火頻度が低いために、決定係数が低いニューロンのデータでは、モデル化誤差の多くの割合が、二項分布によるサンプリングのばらつきによるものであることは1年近く前に認識した。この点を定式化しておこう。

今、時刻 t での、あるニューロンの有限 (m) 個のトレースに含まれるスパイクの数 $y(t)$ のデータから、計算された発火頻度を $f(t) = y(t)/m$ で表わす。 $f(t)$ を、逆ダイナミクスなどのあるモデル式で再構成したものを $r(t)$ で表わす。一方、ニューロンの真の発火頻度の時間変化を $p(t)$ とする。 $p(t)$ は、トレース数を無限大にとったときに観測されるであろう発火頻度と定義できる。われわれが、モデルが良くデータを説明するかどうかを判定したいとき、理想的には $p(t)$ を $r(t)$ が良く近似しているかどうかを知りたいのである。したがって、 $p(t)$ と $r(t)$ の間の決定係数が求められれば都合がいい。しかし実際には $f(t)$ と $r(t)$ の決定係数しか、もとまらないのである。さて $f(t)$ は次式のように表わされる。

$$f(t) = y(t)/m = p(t) + s(t) + n(t)$$

$$\Pr(Y(t) = y) = \binom{m}{y(t)} p(t)^{y(t)} (1 - p(t))^{m - y(t)}, \quad y = 0, 1, 2, \dots, m \quad (1)$$

ここで、 $Y(t)$ は、試行回数がトレース数 m で、事象の生起確率 $p(t)$ の二項分布にしたがう確率変数であるとしている。よく知られているように、 $Y(t)$ の平均値は $E[Y(t)] = mp(t)$ 、分散は $V[Y(t)] = E\{[Y(t) - E[Y(t)]]^2\} = mp(t)\{1 - p(t)\}$ 。さて、確率変数 $Y(t)/m$ の標準偏差は、 $SD[Y(t)/m] = \sqrt{p(t)\{1 - p(t)\}/m}$ となる。 $s(t) = y(t)/m - p(t)$ は実現値の平均からのずれを表わして、これ自身確率変数と考えると平均は0、標準偏差 $SD[s]$ は $Y(t)/m$ の標準偏差と同じになる。 $n(t)$ は二項分布の確率性以外に起因する雑音項である。

さてプルキンエ細胞、下向き眼球運動で典型的な値と思われる1秒間に100パルスの割合で、単純スパイクが生じていて、100トレースのデータを用いて、1ミリ秒のビンで発火パターンを解析する場合を例にとる。このとき $p = 0.1$ 、 $m = 100$ だから $SD[s] = 0.03$ となって、なんとか信号のほうが、二項分布に由来する雑音より大きくなっている。ところが、同じプルキンエ細胞でも、複雑スパイクを、上向き眼球運動で解析しようとするとなると問題が生ずる。1秒間に5パルス複雑スパイクが生じているとすると、 $p = 0.005$ 、 $m = 100$ だから $SD[s] = 0.007$ となって、信号のほうが、二項分布に由来する雑音よりちいさくなってしまふ。そこではと、トレース数を300に増やしたとしても、 $p = 0.005$ 、 $m = 300$ だから $SD[s] = 0.004$ となって、信号と雑音の大きさはほとんど同じで困ってしまう。しかもトレース数をこれ以上増やしても信号雑音比は、トレース数の平方根でしかよくなるらない。

上の例のように信号雑音比が悪い場合でも、1トレースのなかでのビン数が多ければ、モデルのパラメータはかなりよく推定できている可能性がある。しかし困るのは、モデルのフィットのよさを、もはや決定係数で評価できない点にある。 $p(t)$ と $r(t)$ の間の決定係数は1で、モデルが完全だったとしても、 $f(t)$ と $r(t)$ の決定係数はひどく悪くなってしまふということがおこりうるのである。決定係数： $1 - E\{[f(t) - r(t)]^2\} / E\{[f(t) - \bar{f}]^2\}$ の分子に現われる平均自乗誤差は、モデル化そのものの誤差、二項分布の確率性に起因す

る雑音項、それ以外の雑音が互いに独立であるとみなせると次のように表わせる。

$$E[\{f(t) - r(t)\}^2] = E[\{p(t) + s(t) + n(t) - r(t)\}^2] \simeq E[\{p(t) - r(t)\}^2] + E[s(t)^2] + E[n(t)^2] \quad (2)$$

上の式で第1項は小さいのに、第2項が大きくて、結果として決定係数が低くなり、誤ってモデルのフィットが悪いとしてしまっただけではいけない。ここでの結論としては、決定係数はモデルやデータの取捨選択の物差しとしては注意して使わなければいけないということである。同じトレース数、同じような発火頻度であればよいが、たとえ同じプルキンエ細胞でも、単純スパイクと複雑スパイクに対するモデルのフィットを決定係数で比較してはいけないのである。異なるトレース数、異なる発火頻度のデータについても公平に比較できる物差しを探さねばならない。

2.3 発火スパイク数が二項分布にしたがうとする仮説の検定

発火頻度、つまり二項分布の発火確率を線形モデルで推定したあとで、計測データが二項分布にしたがっているかどうかを検定したくなる。これは五味らの本論文 [2] で、実際のデータについて議論される。記録中確率が時間に対して一定であれば、よく知られた二項分布の分散に対するカイ自乗検定を使えばよい。しかし実際には確率が時間とともに変動するので、確率を消したような検定が必要となる。各時点での累積確率が0と1の間で一様分布するだろうという仮定の下に、解析が行えるかもしれない。これに関しても五味らの本論文で実際のデータに関して議論されるはずである。

2.4 二項分布確率の線形パラメータの最尤推定

前節まで述べてきたように、トレース数が少なく、発火頻度が低いときは、スパイク数が二項分布に従うことを陽に考慮した解析が必要と思われる。すると線形パラメータを最小自乗推定しているのが不釣り合いになってくる。確率モデルで行くなら、やはりパラメータも最尤推定する必要があるだろう。

次節以降でも使うので、ここでサンプリングの離散時間を陽に示した記法を導入しておく。トレース数は今まで通り m で表わす。一つのトレースで、解析に用いる離散時刻を $i = 1, 2, \dots, n$ で表わす。小脳、大脳、橋核の追従眼球運動中の神経生理データでは、 $m = 50 \sim 1000$ 、 $n = 200 \sim 400$ 程度を想定している。各時刻のビンにはいる（カウントされた）スパイク数を y_i ($y_i = 0, 1, 2, \dots, m$) で表わす。各時刻 i での、真の発火確率を π_i ($0 \leq \pi_i \leq 1$) で表わす。本稿での最も基本的で重要な仮定は次のようである。

仮定 1 第 i 時刻でのスパイク数は確率変数 Y_i で表わされ、それは試行回数 m 、発火確率 π_i の二項分布に従う。確率変数 Y_i の実現値を y_i で表わす。

$$\Pr(Y_i = y_i) = \binom{m}{y_i} \pi_i^{y_i} (1 - \pi_i)^{m - y_i}, \quad y_i = 0, 1, 2, \dots, m \quad (3)$$

第2の、本節に限った仮定は、発火確率がある時間に依存するデータの線形和になっているということである。

仮定 2 第 i 時刻での発火確率 π_i は次式で与えられる。

$$\pi_i = \sum_{j=1}^p x_{ij} \beta_j \quad (4)$$

データ x_{ij} は、発火確率を再構成できるものなら何でもよくて、たとえば視覚刺激の位置、速度、加速度、レティナルスリップの位置、速度、加速度などが考えられる（それでは線形モデルでは再構成できないと思うが）。ここでは、例として眼球運動データからの逆ダイナミクス表現を示しておこう。眼球回転運動の第 i 時刻での、位置、速度、加速度をそれぞれ θ_i 、 $\dot{\theta}_i$ 、 $\ddot{\theta}_i$ で表わす。ただし、記法を簡単にするために、遅れ時間に対応する時刻のずらしはすでにすんでいるものとして書いている。実際には、さまざまな時間遅れにたいして、以下の解析を行い最も良い結果（最大の尤度）を与える遅れを採用することになる。すると、データとパラメータの具体的な中味は次のようになる。

$$\begin{aligned} x_{i1} &= \ddot{\theta}_i, & x_{i2} &= \dot{\theta}_i, & x_{i3} &= \theta_i, & x_{i4} &= 1 \\ \beta_1 &= M, & \beta_2 &= B, & \beta_3 &= K, & \beta_4 &= f_{bias} \end{aligned} \quad (5)$$

式 4 と、式 5 を、ベクトルと行列の記法で表わすと、次のようになる。

$$\begin{aligned} \pi &= \mathbf{X}\beta & (6) \\ \pi &= \begin{pmatrix} \pi_1 \\ \pi_2 \\ \cdot \\ \cdot \\ \pi_n \end{pmatrix} \\ \mathbf{X} &= \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ x_{n1} & x_{n2} & x_{n3} & x_{n4} \end{pmatrix} = \begin{pmatrix} \ddot{\theta}_1 & \dot{\theta}_1 & \theta_1 & 1 \\ \ddot{\theta}_2 & \dot{\theta}_2 & \theta_2 & 1 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \ddot{\theta}_n & \dot{\theta}_n & \theta_n & 1 \end{pmatrix} \\ \beta &= \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} = \begin{pmatrix} M \\ B \\ K \\ f_{bias} \end{pmatrix} & (7) \end{aligned}$$

一般に、 π は n 次元縦ベクトル、 β は p 次元縦ベクトル、 \mathbf{X} は n 行 p 列の行列である。

本節の線形モデルで、パラメータ β を決めれば、確率 π が定まるので、与えられた実験データ y が逆に考えている確率モデルでどれほど説明できるかを評価する尤度 $L(\pi; y)$ と対数尤度 $l(\pi; y)$ がそれぞれ次式のように計算できる。

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

$$\begin{aligned}
L(\pi; \mathbf{y}) &= \prod_{i=1}^n \Pr(Y_i = y_i) \\
&= \prod_{i=1}^n \binom{m}{y_i} \pi_i^{y_i} (1 - \pi_i)^{m-y_i} \\
&= \prod_{i=1}^n \binom{m}{y_i} \left(\sum_{j=1}^p x_{ij} \beta_j \right)^{y_i} \left\{ 1 - \sum_{j=1}^p x_{ij} \beta_j \right\}^{m-y_i} \\
l(\pi; \mathbf{y}) &= \sum_{i=1}^n \{ y_i \log \pi_i + (m - y_i) \log(1 - \pi_i) \} \\
&= \sum_{i=1}^n \left\{ y_i \log \left(\sum_{j=1}^p x_{ij} \beta_j \right) + (m - y_i) \log \left(1 - \sum_{j=1}^p x_{ij} \beta_j \right) \right\} \quad (8)
\end{aligned}$$

上式 8 の対数尤度の計算で、組み合わせの数から来る定数項は省いてある。さて、 m 、 y_i 、 x_{ij} はすべて実験データから決まっているので、尤度も対数尤度も β だけの関数になっている。尤度あるいは同じことだが対数尤度を最大にするような $\hat{\beta}$ を β の最尤推定量という。最尤推定量を求めるためには式 8 を最大化する β を求めるための非線形最適化を行わなければならない。これは最小自乗推定などに比べるとかなり計算の手間がかかるが本質的な難しさはない。

より本質的な問題は、式 6 から計算した π_i は 0 と 1 の間にはいる保証がないのである。最小自乗推定の場合には π_i が 0 より小さくなったり 1 より大きくなったりしても、気にせず計算を行えるが、確率モデルで尤度を計算するとか、最尤推定をする場合にはそうは行かない。一つの方法は、次式のように図 1 に示した区分線形関数 h を π と $X\beta$ の間に入れることである。

$$\pi_i = h\left(\sum_{j=1}^p x_{ij} \beta_j\right) \quad (9)$$

$$h(q) = \begin{cases} 1 & \text{if } q > 1 \\ q & \text{if } 1 \geq q \geq 0 \\ 0 & \text{if } q < 0 \end{cases} \quad (10)$$

しかしこのようなランプ関数は、その選び方が任意であるという問題点以外に、微分不可能のために数理統計学の有用な結果が使えないという欠点がある。次節では、この難点を解消した一般化線形モデルを用いる。

3 一般化線形モデル

前節までで、一般化線形モデル [5] の導入に必要な記法の定義、動機づけなどはすんでいる。一言で言えば前節最後のランプ関数の代わりに、次式で定義されるロジスティック関数 g の逆関数で、無限回微分可能で、また神経回路モデルでよく使われるシグモイド関数 g^{-1} を使えばよい (図 2 参照)。

$$g(\pi) = \log\left\{\frac{\pi}{1-\pi}\right\} \quad (11)$$

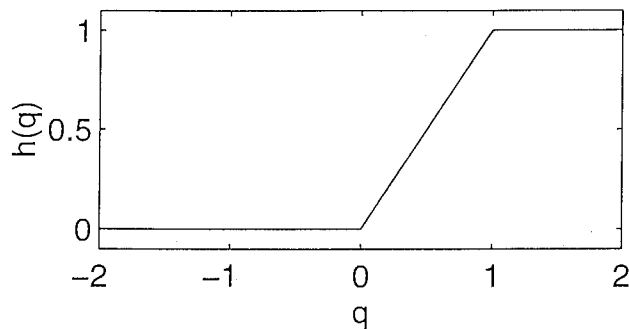


Figure 1: 確率が0と1の間にはいることを保証するランプ関数

$$g(\pi_i) = \eta_i = \sum_{j=1}^p x_{ij}\beta_j \quad (12)$$

$$\begin{aligned} \pi_i &= g^{-1}(\eta_i) \\ &= \frac{\exp \eta_i}{1 + \exp \eta_i} \\ &= \frac{\exp\{\sum_{j=1}^p x_{ij}\beta_j\}}{1 + \exp\{\sum_{j=1}^p x_{ij}\beta_j\}} \end{aligned} \quad (13)$$

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \cdot \\ \cdot \\ \eta_n \end{pmatrix} \quad (14)$$

神経回路モデルのアナロジーで言えば、眼球運動の線形和は、発火頻度そのものではなく、ニューロンの後シナプス電位に相当するような量 η_i をモデル化することになる。発火頻度が比較的高いときには、シグモイド関数の非線形性はほとんど効かず、一般化線形モデルの解析結果と従来の最小自乗法は定数倍を除いてごく近くなる。しかし、発火頻度が低いときには、非線形性が効いてくる。

3.1 最尤推定の漸化式

最尤推定量 $\hat{\beta}$ を求めるための漸化式を与えよう [5]。まず準備として Fischer 情報行列に現われる対角行列 \mathbf{W} を次式で定義しておこう。

$$\begin{aligned} \mathbf{W} &= \text{diag}\left\{m\left(\frac{d\pi_i}{d\eta_i}\right)^2/\pi_i(1-\pi_i)\right\} \\ &= \text{diag}\{m\pi_i(1-\pi_i)\} \end{aligned}$$

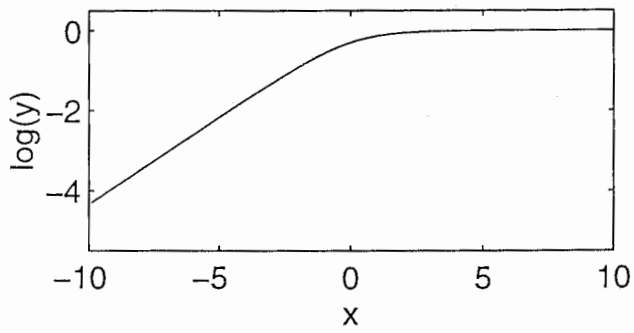
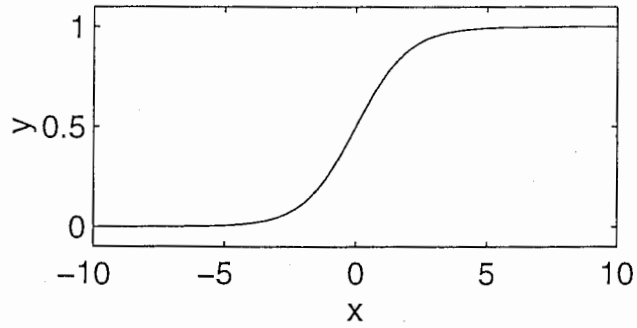


Figure 2: 確率が0と1の間にはいることを保証するロジスティック関数の逆関数：シグモイド関数 $y = g^{-1}(x)$ 。(上) 自然のスケールでプロットしたもの。(下) 縦軸を常用対数でプロットしたもの。

$$= \begin{pmatrix} m\pi_1(1-\pi_1) & 0 & 0 & \cdots & 0 \\ 0 & m\pi_2(1-\pi_2) & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & 0 \\ 0 & 0 & 0 & \cdots & m\pi_n(1-\pi_n) \end{pmatrix} \quad (15)$$

第 k 回目の繰り返しで得られた最尤推定量 $\hat{\beta}$ の k 回目の推定値を $\hat{\beta}^k$ とする。これから 12 式と 13 式を用いて求められた η と π の k 回目での推定値をそれぞれ $\hat{\eta}^k$ 、 $\hat{\pi}^k$ とする。これらの値を用いて次の式で定義する補助変数 z_i^k からなる n 次元ベクトル z^k を計算する。

$$z_i^k = \hat{\eta}_i^k + \frac{y_i - m\hat{\pi}_i^k}{m} \frac{d\hat{\eta}_i^k}{d\hat{\pi}_i^k} = \hat{\eta}_i^k + \frac{y_i - m\hat{\pi}_i^k}{m\hat{\pi}_i^k(1-\hat{\pi}_i^k)}$$

$$z^k = \begin{pmatrix} z_1^k \\ z_2^k \\ \cdot \\ \cdot \\ z_n^k \end{pmatrix} \quad (16)$$

漸化式は次のように与えられる。

$$\hat{\beta}^{k+1} = (\mathbf{X}^T \mathbf{W}^k \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^k z^k \quad (17)$$

ただし、 \mathbf{W}^k は 15 式の π_i に $\hat{\pi}_i^k$ を代入して計算したものである。また T は行列の転置を表わす。小林康による実際のデータへの適用によれば、繰り返しは通常 2、3 回、非常に厳しい終了条件でも、5 回もあれば収束するそうである。

さて初期推定 $\hat{\beta}^0$ は次のようにして求める。

$$\tilde{\mu}_i \equiv (y_i + 1/2)/(m + 1) \quad (0 < \tilde{\mu}_i < 1) \quad (18)$$

$$\tilde{\eta}_i \equiv g(\tilde{\mu}_i) \mid = \sum_{j=1}^p x_{ij} \hat{\beta}_j^0 \quad (19)$$

$$\tilde{\eta} \mid = \mathbf{X} \hat{\beta}^0 \quad (20)$$

$$\hat{\beta}^0 = \mathbf{X}^T (\mathbf{X} \mathbf{X}^T + \epsilon \mathbf{E})^{-1} \tilde{\eta} \quad (21)$$

ただしここで 19 式と、20 式は右辺で左辺を最もよく近似するという意味にとっていただきたい。21 式の \mathbf{E} は n 行 n 列の単位行列であり、 ϵ は小さな正定数で、Moore-Penrose の一般化逆行列を計算するために使われている。この初期推定は、実際の発火数にロジスティック関数を施して $\tilde{\eta}_i$ を求めておき、それを最小自乗推定することである。ただし、ロジスティック関数の引き数は 0 や 1 では困るので、18 式でそれを微調整している。したがってこの初期推定値はかなりいいはずである。

3.2 最尤推定量の信頼性

次式に示すように、最尤推定量は漸近的に不偏であり、その漸近分散は Fischer 情報行列の逆で与えられる [5]。

$$E(\hat{\beta} - \beta) = O(n^{-1}) \quad (22)$$

$$\text{cov}(\hat{\beta}) = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \{1 + O(n^{-1})\} \quad (23)$$

なお、 $O(n^{-1})$ は n^{-1} のラージオーダーを示す。比例係数はわからないが n^{-1} に比例するということである。また分散共分散行列 $\text{cov}(\hat{\beta})$ の第 (i, j) 成分は $E[\{\beta_i - E(\beta_i)\}\{\beta_j - E(\beta_j)\}]$ で与えられる。従って、この行列の対角成分から、推定された各パラメータの信頼性がわかる。

ここである推定量が不偏であるとは次のことを意味する。スパイク数などの観測値が、今考えている確率モデルに従うとする。モデルの真のパラメータを β で表わす。与えられた観測値から前節で説明したような方法で推定された推定値 $\hat{\beta}$ もまた観測値に依存するので、それ自身確率変数となっている。推定値 $\hat{\beta}$ の期待値が真のパラメータ β に等しいとき $E(\hat{\beta} - \beta) = 0$ 、 $\hat{\beta}$ は不偏推定量であるという。22式でわかるように、今の場合は n が無限大になったとき不偏推定量となるので、漸近的に不偏という。

3.3 モデルの尤度比検定

π_i の最尤推定量 $\hat{\pi}_i = y_i/m$ で評価した尤度を L_1 、一方ここで考えているモデルに関する最尤推定量 $\hat{\pi}$ で評価した尤度を L_0 とおく。一般に尤度比 $\lambda = L_0/L_1$ は 0 と 1 の間の値をとり、モデルの仮説が正しければ 1 に近い値、正しくなければ 0 に近い値をとると考えられる。ところで対数尤度比のマイナス 2 倍 (Deviance) $-2 \log \lambda$ はモデルが正しくてかつ以下の条件が成り立てば、 n が無限に大きくなったときには、近似的に自由度が $n - p$ の χ^2 分布に従うことが示せるので、Deviance が χ^2 分布の 100α パーセント点を超えるときにモデルを棄却するという検定方式が得られる。Deviance は次の式で計算できる。

$$\begin{aligned} D(\mathbf{y}; \hat{\pi}) &= 2l(\hat{\pi}; \mathbf{y}) - 2l(\hat{\pi}; \mathbf{y}) \\ &= 2 \sum_{i=1}^n \left\{ y_i \log \frac{y_i}{m \hat{\pi}_i} + (m - y_i) \log \frac{m - y_i}{m - m \hat{\pi}_i} \right\} \end{aligned} \quad (24)$$

条件 1 すべての i について、 n を固定して、 $m \pi_i (1 - \pi_i) \rightarrow \infty$ がなりたつ。

神経生理実験で得られるデータに関しては一般に上の条件は成り立たないものと思われる。特に複雑スパイクなどでは決してありえない。しかしその場合にも次式で定義される Deviance の差をもちいれば、互いに階層的な関係にあって、1 つだけ自由度の違うモデルはどちらが優れているか正しく検定できる。この量は自由度 1 の χ^2 分布に従うことが知られている。そのような 2 つの階層的なモデルの例としては、加速度、速度、位置を用いた標準的なモデルに対して、躍度を付け加えたモデルとか、位置をはずしたモデルなどである。

$$\begin{aligned} D(\mathbf{y}; \hat{\pi}_0) - D(\mathbf{y}; \hat{\pi}_A) &= 2l(\hat{\pi}_A; \mathbf{y}) - 2l(\hat{\pi}_0; \mathbf{y}) \\ &= 2 \sum_{i=1}^n \left\{ y_i \log \frac{\hat{\pi}_i^A}{\hat{\pi}_i^0} + (m - y_i) \log \frac{1 - \hat{\pi}_i^A}{1 - \hat{\pi}_i^0} \right\} \end{aligned} \quad (25)$$

ここで、自由度を増やす前のオリジナルのモデルを0で、自由度を増やしてより複雑にした進んだモデルをAで示した。 $\hat{\pi}_i^0$ と $\hat{\pi}_i^A$ はそれぞれのモデルでの最尤推定値から評価された第*i*時刻での発火確率である。2つの階層的モデルの自由度が*q*違うなら、検定には自由度*q*の χ^2 分布を使えばいい。

4 おわりに

神経生理学の実験では、動物にある課題を課して、刺激や運動などの目印になる出来事を時間の基準に決め、複数回の試行で得られた1つのニューロンの発火を重ね合せるということはよく行われる。このようにして得られたスパイク数の累積データから、そのニューロンが何を符号化しているかを知ろうとすれば、発火頻度の時間パターンを予測したり、再構成したりする数式モデルを使わねばならない。本稿では、発火頻度が低かったり、トレース数が少ないときには、このような数式モデルに付け加えて、シグモイド関数と二項分布モデルを挿入することで、数理統計学のさまざまな結果が有効に使えることを示した。

一般化線形モデルを使うことのデメリットは、モデルの解釈がややこしいということである。発火頻度の確率は線形モデルにシグモイド関数を施したもので表わされるから、非線形モデルになっている。ただしこの非線形性は穏やかなもので、発火頻度があまり激しく変化しないなら、シグモイド関数の使われる部分はほとんど線形といってよいであろう。この時、単純な線形モデルとの最大の違いは、低い発火頻度を非常に小さな線形パラメータで表わすか（線形モデル）、線形パラメータは比較的大きく、バイアス項に対応する量を負にして低い発火頻度を再現するか（一般化線形モデル）の違いとも解釈できる。

たとえば同じプルキンエ細胞の単純スパイクと複雑スパイクの発火頻度のモデルを比較するとき、線形モデルでは、発火頻度の範囲が違うから、線形モデルのパラメータが皆数十倍違ってくる。しかし一般化線形モデルでは、バイアス項だけが大きく（2とか3とか）違って（発火頻度が低いときは自然対数で効くからバイアス項の2とか3とかの違いは発火頻度では7倍とか20倍の違いである）、ほかのパラメータは同程度で比較しやすいということになる可能性もある。

一般化線形モデルの、ニューロンのモデルとしてのより積極的な解釈は、線形部分が後シナプス電位、シグモイド関数は後シナプス電位から発火確率への変換を定めるニューロンの非線形性を表わしていて、バイアス項はニューロンのいきちを表わしているというものである。たとえばのはなしとして、もし複雑スパイクと単純スパイクで、バイアス項だけが、違うのであれば、後シナプス電位は本質的に同じだが、いきちだけ違うなどと推論できる。

これも仮定に基づいての話だが、複雑スパイクの発火頻度の線形モデルではトレース数が多くても決定係数は高くなり、一般化線形モデルでは良いフィッティングが得られたとする。この場合は、複雑スパイクは、プルキンエ細胞が支配している細胞へ情報を伝達するという役目より、細胞内の後シナプス電位がかかわる何らかの過程を制御しているように見える。

あるニューロンを、線形モデルにせよ一般化線形モデルにせよ解析して、その発火頻度が理論的に定式化できる、ある時間とともに変化する量であるとみなせたとする。しかしこの時にも50トレース程度から数百トレースに及ぶデータを重ね合せて初めてそのよ

うな結論が得られる。しかし、1個のニューロンが実時間で、運動制御、認知などを含めた脳の働きに役立っているときは、このような時間の重ね合せはできない。では、上のような解析は意味がないのではないかという疑問が生じる。筆者らの考え方は、時間的な平均が空間的な平均で置き換えられるということである。同じような性質と機能を持ったニューロンが多数個存在すれば、そしてそれらの出力が共通の標的に向かっているなら、時間平均は、標的ニューロンへのシナプスの収束で置き換えられる。したがって、本稿で展開したような発火頻度解析によって得られた結論が脳の情報処理にとって意味を持つかどうかは、似たような性質を持つニューロンが何個程度存在し、またその出力が共通の標的に収束するかどうかにかかっている。追従眼球運動にかかわる小脳傍片葉に関して言えば、粗い試算によれば約5万個のプルキンエ細胞が存在し、その出力は前庭核にしる中位核にしる、1つのニューロンを介して、外眼筋運動ニューロンに収束している。眼球の水平回転、垂直回転、回旋などいくつかの機能別にポピュレーションが別れているだろうが、1つの機能（たとえば垂直回転）に1万個のプルキンエ細胞があるとしてもひどい過大評価にはならないだろう。以前の解析からこのプルキンエ細胞には、線形パラメータの違いという意味で、さらに個性があることがわかっている。しかし、発火頻度がより下流でたしあわされるなら、これらの個性も含めて空間平均がおきるので、数百回程度の時間平均で、つきとめられた情報表現は実際に使われていると考えてよいだろう。

一般論として、本稿にのべた発火頻度解析を行うなら、同様の性質を持つニューロンが少なくともトレース数の数倍はありそうなこと、またそれらの出力が、ある共通の標的、1個のニューロンである必要はないが、少なくとも同様の性質を持った1群のニューロンに収束している必要がある。

本稿の解析が有効であるためのもう1つの条件として、各時刻でのスパイク発火は隣の時刻でのスパイク発火と、発火頻度 $p(t)$ の時間的連続性を除いて相関がないということがあげられる。あるビンで発火が起きたからといって、隣のビンの発火確率が $p(t)$ で与えられるものから上昇しても、減ってもいけないのである。五味は単純スパイクに関しては、発火の自己相関関数が急激に0に近づくことから、この条件が満たされる場合が多いことを確かめている [2]。一般の場合にも、発火頻度解析を行うためには、まずこれを確かめねばならない。

References

- [1] W. Bialek and F. Rieke. "Reliability and information transmission in spiking neurons", *Trends in Neurosciences*, **15**, pp. 428-434, 1992.
- [2] H. Gomi, M. Shidara, A. Takemura, Y. Inoue, K. Kawano and M. Kawato. "Reconstruction of temporal firing patterns of Purkinje-cells in monkey ventral paraflocculus based on inverse dynamics representation", *The Journal of Neuroscience*, to be submitted.
- [3] 広中平祐編 現代数理科学事典 大阪書籍 1991.
- [4] M. Kawato, K. Furukawa and R. Suzuki. "A hierarchical neural-network model for control and learning of voluntary movement", *Biological Cybernetics*, **57**, pp. 169-185, 1987.

- [5] P. McCullagh and J. A. Nelder FRS. *Generalized Linear Models*. Chapman & Hall, London, 1989.
- [6] B.J. Richmond and L.M. Optican. "Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform", *J. Neurophysiol.*, **57**, pp.147-161, 1987.
- [7] M. Shidara, K. Kawano, H. Gomi and M. Kawato. "Inverse-dynamics encoding of eye movement by Purkinje cells in the cerebellum", *Nature*, **365**, pp. 50-52, 1993.
- [8] 竹内啓編 統計学辞典 東洋経済新報社 1989.
- [9] C. von der Malsburg. "Pattern recognition by labeled graph matching", *Neural Networks*, **1**, pp.141-148, 1988.

付録 1 : 最尤推定量を求める漸化式の導出

一般化線形モデルに関する手短な解説に関しては [8] の 1 1 ・ 5 節 : 一般化線形モデル (GLIM) を参照せよ。ここに、最尤推定量を求める漸化式、信頼性、尤度比検定についてもまとめてある。ただし尤度比検定の記述は不完全である。[3] では X[3]-5 : 離散データ解析法 (5 2 9 から 5 3 1 ページ) に関連の記述がある。[5] の 2 ・ 5 節 : 一般化線形モデルのフィッティングアルゴリズム (4 0 から 4 3 ページ) および 4 ・ 4 ・ 2 節 : パラメータ推定 (1 1 5 から 1 1 7 ページ) を参照せよ。

付録 2 : 最尤推定量の信頼性

[5] の 4 ・ 4 ・ 4 節 : 推定のバイアスと正確さ (1 1 9 から 1 2 0 ページ) を参照せよ。

付録 3 : 尤度比検定

[5] の 2 ・ 3 節 : フィットの良さを測る (3 3 から 3 6 ページ) および 4 ・ 4 ・ 3 節 : デバイアンス関数 (1 1 8 から 1 1 9 ページ) を参照せよ。