

多細胞同時記録データの統計解析法

伊藤浩之

京都産業大学工学部情報通信工学科

1. はじめに

複数の細胞のスパイク活動を同時記録し、細胞活動間の関係を統計的に解析する方法は、60年代後半の Perkel らの相互相関ヒストグラム (cross correlogram) の理論的研究に始まり、80年代初頭の外山らのネコ視覚系での実験的研究により大きく注目された。一方、より大規模な多細胞同時測定を用いた研究法の意義と問題点に関して総合的なレビューを試みるものとして、80年代前半に Gerstein らや Krueger の先駆的な論文が提出された (Gerstein et al. 1983, Krueger et al. 1983)。そして、80年代後半に Gray と Singer らによって提示された空間的に離れた細胞間の同期振動現象と情報のバインディング問題 (Gray & Singer, 1989) は大きな注目を集め、多細胞同時測定法は、従来の細胞間の解剖学的な結合形態の解析という目的から、神経ネットワークにおける動的な情報処理メカニズムの解明という新たな観点から、その必要性が強く認識された。また、新たな電極の設計、測定機器の開発、多細胞データのスパイク弁別アルゴリズムなどの進歩により、大規模な多細胞同時測定が現実的に可能な状況と成りつつある。具体的な多細胞同時記録法の最近の進展は、櫻井の講義を参考にさせていただきとして、この講義では、同時に細胞外記録された複数の細胞のスパイク活動データをどのように統計解析していくのかに関して概説を行なう。本ノートは、講義に参加する前に、cross-correlogram までの基礎的な解析法の理解を自習してもらう目的で作成された。

2. これまでの研究

実際の細胞外記録実験においては、微小金属電極などを脳組織に刺入し、電極先端付近に位置するニューロンの発火に伴うスパイク活動電位を記録する。脳組織を直接に電気刺激または化学刺激して活動を測定する場合や自発的な活動を測定する場合も考えられるが、ここでは麻酔下ないし覚醒行動中の動物の脳内においてニューロン活動を記録し、感覚刺激の提示または行動課題の実行との相関を解析する場合を想定する。全く同一の刺激提示下においても、ニューロンの活動には大きな variability が存在するために、通常は同一の刺激提示ないし課題遂行を十分に多くの試行回数にわたり繰り返し、その試行平均をもって統計的な議論を行う。例えば、一次視覚野での方位選択性を持つ単純細胞に最適方位の他にも複数の異なる方位の線刺激を提示して、平均発火率に基づいて方位チューニング曲線を求める場合、最

適方位の刺激提示に対する試行ごとの発火率のばらつきの大きさは、方位チューニング曲線における異なる方位に対する平均発火率の変動より大きい事が知られている。櫻井も議論するように、我々は試行平均などを必要とせず、single shot で信頼度の高い知覚を行っている。このメカニズムにはいくらかの説明が試みられて来たが、未だ決定的なものは無い。とにかく、ニューロンの活動は、脳にとっては十分に relevant であるのだが、我々研究者が論文にまとめるにはあまりに irrelevant であるため、試行平均という操作が不可欠となる。

まず初めに、単一ニューロンのスパイク活動データに対して行われてきた解析法を簡単に概観する。

A. ラスター表示 (Raster Display)

測定したニューロンの活動電位において一定のレベルを設定し、それ以上（またはそれ以下）に電位が変動した場合にスパイクが生じたとして、その時刻のタイムスタンプを記録する。横軸に時間を取り、一つの試行時間内でスパイクが生じた時刻すべてに短い縦線を並べていくと、スパイクの発火パターンが視覚化される。異なる試行ごとに列を下にずらして表示したものをラスター表示とよぶ。試行時間内での発火の変化および試行間での variability の双方が視覚化されるため、定性的な理解にとっては有効である。人間のパターン認識能力はどんな統計解析法より優れている。異なる試行間でのスパイク列の時間の原点をそろえるために、刺激の提示時刻や課題でのキュー信号の提示時刻など外部事象の onset のタイミングを用いる。

B. Peri-Stimulus Time Histogram (PSTH)

上で求めたラスター表示において、時間軸上を一定幅の小区間で分割する。測定する脳部位によって平均の発火レベルが異なるので一概には言えないが、通常 10 ~ 50 msec ぐらいの幅の区間を用いる。各区間に入っているスパイクの数を試行全体に渡って平均し、この平均値を区間の代表値としてヒストグラムを作成したものが、Peri-Stimulus Time Histogram または Post-Stimulus Time Histogram と呼ばれ、一般には PSTH という略名で使われる。刺激の提示時刻というイベント時刻で異なる試行を揃えて、平均を行っている点に注意せよ。これは、ラスター表示とは異なり、小区間内での時間平均および試行平均という統計操作を行っている。これらの人工的な操作によって消失する情報に脳の秘密が隠れているという危険性も存在するのである。解析法というものは、何らかの情報を捨てるという操作の上に常に成り立っている。

C. Inter Spike Interval Histogram (ISI)

一つの試行でのスパイク列に対して、連続して生じる発火の時間間隔をすべて測定し、ヒストグラムを作り、更に全ての試行に対して平均を行う。これを Inter Spike Interval Histogram といい、ISI などと略される。これは、スパイクの間隔パターン

にどのような統計性があるのかを知る簡単な方法である。皮質のニューロンの ISI は Poisson 型であり、moto-neuron では Gaussian 型であることが知られている。ただし、これは時間平均および試行平均を行った一次の統計量であり、全く同じ ISI を持つ、異なるスパイクパターンは幾らでも作れる事に注意せよ。スパイクの発火パターンを更に解析するためには、高次の統計量による記述が必要となる。

D. Auto-Correlogram

ISI では解析できない、スパイクパターンの 2 次の統計性を解析する。一つのスパイク発火が生じた時点で、その前後にどのようなタイミングで他のスパイク発火が生じているのかを統計解析したもの。実際の計算は、(-128 msec, 128 msec) の区間の時間軸をとり（一般には FFT の実行のため、2 のべき乗の区間数となるようにする）、1msec の小区間で分割し、ヒストグラムの計算の作業場を作る。一つの試行のスパイク列を持ってきて、そのコピーを一組作り、時間を揃えて上下に並べる。まず、上のスパイク列の一番目のスパイクに着目する。下のスパイク列のこの時点（単に上のスパイクと同じスパイクがあるだけであるが）が準備してあるヒストグラムの時間軸の原点になるように、下のスパイク列を時間軸に重ねる。このスパイク列に対してヒストグラムの時間軸に切ってある小区間ごとに何発スパイクが含まれるかを計算し、各区間のカウント値として保存する。次に、再び 2 つのスパイク列を並べ、今度は上のスパイク列での 2 番目のスパイク発火時点に着目する。下のスパイク列でのこの時点が原点となるようにヒストグラムの時間軸に重ね、先程と同じく各小区間ごとに入るスパイク数を計算し、先程保存してあったカウント数に加える。この作業を、上のスパイク列のすべてのスパイク発火に対して行うことで、ヒストグラムが作成される。PSTH では刺激提示時刻を原点として発火率のヒストグラムを作成した。Correlogram では (Auto- でも Cross- でも) 上のスパイク列での各スパイク発火時点を原点として、下のスパイク列の PSTH を計算する操作に対応する。すべての試行に関して、ヒストグラムを平均することにより最終的な auto-correlogram が求められる。スパイク列をヒストグラムの時間軸に重ねた時に、時間軸の範囲 (-128 msec, 128 msec) を超えてしまう分に関しては、計算を行わない。このヒストグラムでは、PSTH では平均化されてしまった 1msec 程度の短い時間分解能でのスパイク発火間の相関が視覚化される。しかし、試行時間内での全てのスパイクに対しての時間平均は行われており、非定常性などの問題が生じると信頼性は失われる。

刺激性相関と神経性相関 (Shuffling と Shift Predictor)

上で求めた auto-correlogram での計算法でも理解されるように、どんなスパイク列でもスパイクが存在すれば、ヒストグラムにおいて正のカウントが積み上がり、なんらかの構造が存在する。PSTH の計算では、区間幅が 50msec 程度であったた

めに、この区間内であれば同じ数のスパイクをどのように配置しても同じヒストグラムとなってしまう。ここで、実際のスパイクデータと全く同じ PSTH と ISI (一次統計量) を持つが、PSTH の各区間内は全くランダムなスパイク列を考える。このスパイク列に対しても、auto-correlogram を計算すると、何の構造もない平たんなヒストグラムを得る。これは、一次統計量は等しいが、スパイクの発火タイミングに関しては何の相関もないとする場合 (帰無仮説) でも生じる accidental な相関構造である。実際に関心があるのは、この帰無仮説からのズレである。これは、スパイクデータから直接計算された correlogram (raw-correlogram という) から帰無仮説の場合の correlogram を各小区間ごとに差を取ることで評価される。これは、統計では、スパイク発火の covariance を計算する事に対応する。

実際の実験データに対して、理想的な帰無仮説のスパイクデータをシミュレートする事は大変である。一般には shuffling または shift predictor という二つの方法のどちらかを用いて近似的に correlogram を評価する。Raw-auto-correlogram の計算においては、各試行でのスパイク列の下にそのコピーを並べてヒストグラムを計算した。ここで、上の列のオリジナルのスパイク列 (試行回数分ある) は固定して、それぞれの下にあるコピーのスパイク列の試行の順番をトランプを切るようにランダムに shuffle する事を考えよう。この場合には、オリジナルなスパイク列と他の試行でのスパイク列のコピーとを組み合わせる correlogram を計算することになる。この試行回数分を平均することで、疑似的に帰無仮説の場合の correlogram を評価できる。これは、異なる試行のスパイク発火のタイミング間に統計的な独立性を仮定した場合、試行の順番を shuffle する事により、PSTH および ISI は同じであるが、細かい時間タイミングは独立なスパイク列の組が出来上がるからである。これを shuffled correlogram という。また、もっと簡単には、shuffle ではなく、単に 1 つずれた試行でのスパイク列同士を組み合わせる shift correlogram または shift predictor と呼ばれるものをもって評価する。別に、いくつ分試行を shift させても良いが、1 つの shift で統計的独立性は満たされると信じられている。試行ごとのスパイク列の統計性の variability や異なる試行をまたぐような長時間スケールの変動などがあると、この方法は破綻する可能性があることを覚えて置くべきである。

E. Cross-Correlogram

Auto-correlogram は一つのニューロンのスパイク発火における微細な時間的構造を統計解析したものであったが、神経ネットワークでのニューロン間の発火の時間的関係を解析するためには、それぞれのニューロンのスパイク列の組に対して、cross-correlogram を計算する必要がある。Auto-correlogram で説明した各試行ごとのオリジナルスパイク列とそのコピーのスパイク列の代わりに、同一試行で同時記録された二つのニューロンのスパイク列を用いて同様なヒストグラムを計算すること

で、raw-cross-correlogram が得られる。また、この場合の帰無仮説は、二つのスパイク列での発火時刻が統計的に独立である事であるが、疑似的な帰無仮説の場合の cross-correlogram は auto-correlogram の場合と同様に shuffling または shift によって評価できる。現在、この解析法は最も広く用いられている。

3 . 最近の研究

Cross-correlogram は有効な方法であるが、以下のような問題点が存在することはかなり以前より指摘されていた。1) 発火率が異なるようなスパイクデータ間での相関の強さの比較のための定量的な議論 (規格化の問題) があいまいである。2) 試行時間内で相関構造が非定常性を示す場合に、時間平均操作は不適切な結果を導く。3) あくまでも二つのニューロン間の相関しか視覚化できない (多体相関構造の視覚化の必要性)。これらのどの問題も、スパイク発火タイミングにおける時空間的關係性に着目する情報符号化パラダイムの実験的検証においては、重要となる問題点である。1の相関の強さの規格化の問題に対して、Aertsen らの Joint-PSTH は一つの方向を示したが (Aertsen et al. 1989)、最近伊藤らは規格化の問題には原理的な問題点が存在する事を指摘した (Ito & Tsuji 1999)。2の非定常性の問題は、同じく Joint-PSTH や Unitary Event Analysis (Riehle et al. 1997) で扱われている。3の多体相関に関しては、Unitary Event Analysis および Gravitational Clustering (Gerstein 1998)で考察されている。上に挙げた3つの問題点はすべて難題であり、現在までのところ十分に満足できるものは提案されていない。最近の15年間ぐらいにわたり、世界中で多くの研究者がスパイクデータの更に有効な視覚化法を目指して研究しているが、少なくともヒット作と考えられるものは、すべて Gerstein や Aertsen から出されていることは評価すべきである。彼らのこの問題に対する関心の先見性および本質の理解の的確さは、システム的な神経科学を目指すものにとっては把握しておくべき事である。

当日の講義においては、cross-correlogram を超えようとする試みと、なぜ超えなければいけないかなどの背景を説明する。

4 . おわりに

多細胞同時記録の技術の進歩により、100個レベルのニューロンのスパイク活動の同時記録も不可能では無い状況である。しかし、この膨大なデータの洪水から何の特徴に着目して、脳の情報処理メカニズムのシナリオを作っていったらよいかは、誰も知らない。これは、「ニューロンのスパイク発火の何に、情報が符号化されているのか」という問題に対する答えが明確でないからである。実際には決して知ることのできない脳内ネットワークの活動状態の全貌を、測定された一部のニューロンのスパイク活性のみを用いて逆推定する問題は、明らかにユニークな解が求まらな

い不良設定問題である。この問題に対する唯一のアプローチは、何らかの拘束条件を研究者自らが設定して解の範囲を限定して、可解な逆問題へと変換する事である。この拘束条件こそが、脳の情報処理に関して我々が設定する作業仮説、パラダイムであり、データ解析の本質は研究者が設定したパラダイムという色眼鏡を通してデータに含まれる情報を縮約し、脳内の現象を解釈していくという主観的な作業にほかならない。設定したパラダイムとデータ解析の結果との論理的整合性からパラダイム自体も改良されていく。成熟した科学分野においては近代科学が構成してきた合理主義的オリエンテーションに従った論理の階層構造化により、巧みに研究者の主観はできるだけ対象から遠方に位置するように設定されている。しかし、未だ全く組織化されていない脳のダイナミクスの解析に関しては、本来不可避である研究者の主観が対象のすぐ近くに存在してしまうのである。今後、多細胞同時記録により多細胞データの洪水に溺れ流される中で、我々はいかなる論理システムを構築して脳を理解していけば良いのであろうか？これは、我々の世代に課せられた、困難だが非常に興味ある問題である。

5 . 参考文献

- Aertsen, A.M. H. J., Gerstein. G. L., Habib, M. K., Palm, G. (1989) Dynamics of Neuronal Firing Correlation: Modulation of "Effective Connectivity", *J. Neurophysiol.* 61, 900-918.
- Gerstein, G. L., M. J. Bloom, I. E. Espinosa, S. Evanczuk, and M. R. Turner (1983) Design of a laboratory for multineuron studies: *IEEE Transactions on System, Man, and Cybernetics*, SMC-13: 668-676.
- Gerstein, G. L., P. Bedenbaugh and A. M. H. J. Aertsen (1989) Neuronal Assemblies, *IEEE Trans. Biomed. Engineer.* 36 4-14.
- Gertsein, G. L. (1992) Computers in Study of Simultaneously Recorded Spike Trains: Organization of Neuronal Assemblies, in *Methods in Neurosciences*, Volume 10 Computers and Computations in the Neuroscieneces ed. by P. M. Conn (Academic Press, San Diego) pp.48-61.
- Gerstein, G. L. (1998) Correlation-Based Analysis Methods for Neural Ensemble Data, in *Methods for Neural Ensemble Recordings*, ed. by M. A. L. Nicolelis (CRC Press, Boca Raton).
- Gray, C. M., P. Koenig, A. K. Engel, and W. Singer (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties: *Nature* 338: 334-337.
- Krueger, J. (1983) Simultaneous individual recordings from many cerebral neurons: Techniques and results: *Rev. Physiol. Biochem. Pharmacol.* 98: 177-233.

Ito, H. and S. Tsuji (1999) Model Dependence in Quantification of Spike Interdependemnce by Joint Peri-Stimulus Time Histogram, *Neural Computation*, to be published.

Riehle, A., Gruen, S., Diesmann, M., Aertsen, A. (1997) Spike synchronization and rate modulation differently involved in motor cortical function. *Science*, 278, 1950-1953.

伊藤浩之 (1998) 多細胞活動同時測定法、*医学のあゆみ*、184, 599-605.

櫻井芳雄 (1998) スパイク相関解析法、*医学のあゆみ*、184, 607-612.