

////////////////////////////////////

提出したレジメに記載した「局所的位相シフト検出法」について作業を進めるために「情報幾何学」の調査を続けてるのですが、そこで甘利先生の「情報幾何学の新展開 新版」(2019.11.25 発行)に辿り着きました。新版では「深層学習」に関する内容が追加されています。

甘利先生が重視してと思われる項目を目次から抜き出すと、

- 0 1) 「1 5. 3 接神経核 (NTK) 理論」
- 0 2) 「1 5. 4 深層学習に関する驚異の発見」
- 0 3) 「1 5. 4 ランダム回路網の信号伝搬」

となります。「新版への序」には次のように書かれています。

[…

深層学習はこの予想を覆して、やってみればうまく行くという成果を示した。それならばそれを示す理論的な根拠が必要である。近年、 $P$ が $N$ 以上に十分に大きいときの挙動に関する理論が出始めた。このとき極小解問題は生ぜず、 $P$ が十分に大きければ、訓練誤差は0に減ること。また、 $N$ を超えて $P$ を大きくすれば、汎化誤差が再び減少に転じることなどである。

ランダムに結合した回路網を初期値として学習を開始したとき、 $P$ が十分に大きければ最適解はそのごく近くにあるという、驚くべき発見があった。これは Neural Tangent Kernel という構想で発表され、世界に広まっている。

…]

まとめると、大きくは

- 0 4) 「極小解問題は生じない」
- 0 5) 「ランダムに結合した回路網を初期値として学習を開始すると、 $P$ が十分に大きければ最適解はそのごく近くにある」

の二つになると思います。

以上のことを「局所的位相シフト検出法」(特に「局所的位相シフト検出法」の検討過程で出た「白色雑音を搬送波と見立てて行う『変調・復調』」)と関連付けて考え直したのですが、0 4)と0 5)を説明できる「物理的な現象」を少なくとも一つ見つけることができたと考えています。それは

- 0 6) 白色雑音と信号とを非線形特性により相互変調させて、白色雑音が支配的な帯域(だいたい高周波側)をハイパスフィルターで取り出す

というものです。

06)は「局所的位相シフト検出法」の効果確認で行った実験(資料3)に記載で、実験用のデータを作成した時の手順そのものです。06)の手順で作成したデータに「局所的位相シフト検出法」を作用させると元の信号を復調でき、06)の手順は白色雑音を搬送波とした変調そのものと言えます。ですので06)の手順で作成したデータを以下では「変調済み白色雑音」と呼ぶことにします。「変調済み白色雑音」は後で「局所的位相シフト検出法」を用いて復調できますので、元の信号と比較しても情報量の減少などの劣化は起きていないと言えます。

(資料3)に示した実験を行った際に信号に白色雑音を加算したのは、単にノイズフロアを再現しなかったからなのですが、今は白色雑音には04)と05)を実現するための重要な意味があると考えています。

「変調済み白色雑音」は、信号のパターンを白色雑音の局所的な位相の進み遅れ量として埋め込んだものとなりますが、基本的には白色雑音に似ています。また全ての信号に対応して「変調済み白色雑音」が作られますが、それらは全て白色雑音に似ることになり、空間的には全ての「変調済み白色雑音」が白色雑音の近隣に存在することになります。であれば、それらを用いて決められる例えば「分類に用いる超平面(最適解の一例)」も白色雑音の近隣に存在することになります。すなわち05)に示された状況となります。

白色雑音に非線形特性を作用させると次に示す特別な現象が起きると考えてますが、「変調済み白色雑音」についても同様と考えています。

白色雑音は「パワースペクトルの低域から高域までが一律に同じ値、位相スペクトルの低域から高域までが一様分布の乱数」であるようなデータと考えることができます。そのような白色雑音に非線形特性を作用させると、当然、高調波歪と混変調歪とその他の非線形性歪は発生しますが、そのパワースペクトルには、高調波歪と混変調歪とその他の非線形性歪の分だけ、低域から高域まで一律に同じ値が下駄として嵩上げされたような状態になります(元のパワースペクトルが一律な値なので)。つまり白色雑音に非線形特性を作用させても、元の白色雑音の全体にゲインがかかった状態になるだけで線形性は保たれることとなります。ちなみに上では白色雑音と言っていますが、雑音でなくても白色であれば上の現象は起きると考えられます。

元の信号ではなく「変調済み白色雑音」を用いる場合に、かつ白色雑音の近隣範囲を空間的に離れなければ、線形性は保たれることとなります。極小解の原因は非線形性ですから、04)という状況になります。

04)と05)については、現状のNTKに対する検討では「何故、そうなるのか」などの理由は説明されていないようです。06)によって理由が説明できるのならそれは重要です。理由が判ればそれ以降は他者に対して一歩二歩と先行できる可能性が出てきます。

06)を行うということは、「雑音を積極的に利用する」ということとなりますが、「生体内で雑音を積極的に利用してる」とされてる話は既に幾つかあります。例えば、視覚では「眼球の微細振動」、聴覚では「基底膜にある振動(音)を感知する部分である繊毛のゆらぎ」などです。

06)が正しいようなら、神経回路網の動作メカニズムにも「雑音を積極的に利用する」という部分が本質的に含まれることになる、というのは注目すべき点だと思います。

前回のレジメに記載した「局所的位相シフト検出法」の検討では、その特徴である「線形性」と「超解像」では「超解像」にフォーカスしており、信号伝達の媒体は生の物理現象に近い「電波」「音波」「振動」「電圧の変動」「電流の変動」などに限って考えていました。

06) に関しては、「線形性」にフォーカスすることにしまして、入力段の後で人為的に白色雑音を加算することも考えることにします。であれば NTK の応用が考えらる（具体的には画像やその他を含む）分野の全てが、同様に処理の対象になると考えて良いはずです。

白色雑音の特徴を持つ実体としての白色雑音  $N_n$  は無数にあります。そこを含めて「信号」と「変調済みの白色雑音」を空間にプロットする模式図を描き、そこで考察できることを以下に列挙します。（上の図を添付する pdf ファイルに含めます）

- ・ 信号である  $s_1$ ,  $s_2$ 、および何らかの最適解である  $opt$  の周囲は、線形であるという保証がなく極小解が生じる可能性があり、最適解  $opt$  が得られるかはわからない。
- ・ 白色雑音  $N_n$  の近隣には線形性が保たれ極小解が生じない領域がある。06) により得られる変調済み白色雑音  $N_{n_s1}$  は上の領域に含まれる。 $N_{n_s2}$  も同様。上の領域には極小解が無いので、最適解  $opt_n$  が得られる。
- ・ 白色雑音の特徴を持つ雑音の実体は複数存在する。そのそれぞれで上のようなことを行い最適解が得られるとすると、全体で得られる最適解は  $[opt_1, opt_2, opt_3, \dots, opt_n, \dots]$  となる。
- ・ 現在までの NTK に関する検討で、最適解が複数あること、複数ある最適解は「性能的に同等らしい」ということが言われている。
- ・ 現在までの NTK に関する検討で「極小解が生じない」ということが言われてるが、「白色雑音の近隣領域に限られる」という考察は、自分が調べた範囲では見つけられなかった。
- ・ 「白色雑音の近隣領域に限られる」という考察については、「学習の初期値を白色雑音とした時にのみ全体がうまくいく」という NTK の実験事実の裏返しとして説明できると考えている。
- ・ 信号のセット  $[s_1, s_2]$  に対して、同一セット内では白色雑音  $N_n$  を固定する必要があるが（図を見れば明らか）、この固定については、現在までの NTK に関する検討でも、既に言われている。
- ・ 06) の実行により極小解の心配をしなくて済むようにはなるが、その代償として「ノイズ抑制」には気を遣う必要が出てくる、と予想できる。ただし今までの神経回路網の分野では「ノイズ抑制」についてあまり検討されていないように見える。

これらの検討により、深層学習の Neural Tangent Kernel (NTK) について、最前線に立てるものと考えています。

////////////////////////////////////