

ニューラルネットワークを用いた逆問題解法

電気通信大学情報工学科

竹田辰興、福原誠、馬笑峰、Liaqat Ali

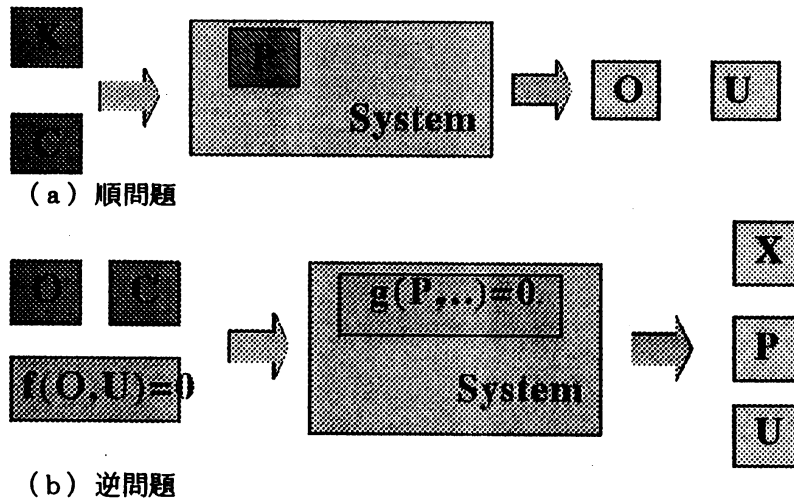
Department of Computer Science, The University of Electro-Communications

Tatsuoki TAKEDA, Makoto FUKUHARA, Xiao Feng MA, Ali LIAQAT

1. 逆問題

逆問題について厳密な定義を与えることは困難である [1]。しかし、多くの場合、逆問題を順問題から区別する事は必ずしも困難なことではない。特に、リアルワールドにおいて遭遇する逆問題を解くという現実的な課題に対しては逆問題の持つ非適切性に基づく困難をいかにして克服するかということに問題が集約されるので考えるべき対象が自ずから明らかになってくる。我々はニューラルネットワーク [2] の応用について研究する過程で、ニューラルネットワーク選点法が広い範囲の逆問題解析に有効であることを見いだした。そこで、このような扱いが可能であるような種類の問題を当面の「逆問題」として設定し、これらについて具体的な解法を提案し、またモデル実験などでその方法の妥当性について検証した。

あるシステムの状態が、初期条件、境界条件、微分方程式によって定まる場合を想定して順問題と逆問題の関係を図示すると第1図のようになる。逆問題では、多くの場合、解を定めるには不十分な情報しか用意できないような事態が生じるために解を求めるのに困難が生じる。



第1図 順問題と逆問題：ここで、XとCは入力データと拘束条件、Pは内部パラメータ
OとUは、順問題において観測されるあるいは観測されない出力データを表す。

2. 多層ニューラルネットワーク

2.1 多層ニューラルネットワークの構造

ニューラルネットワークは、単純な演算機能を持つニューロン（ユニット）を多数ネットワーク状に結合して複雑な演算機能を実現できるようにしたシステムで、もともと脳の機能を模擬しその働きを明らかにするために考案された。現在では本来の目的をはずれてニューラルネットワークの優れた性質を利用した応用についての研究も多数行われるようになった。生物のニューラルネットワークと区別する必要がある場合には「人工ニューラルネットワーク」（ANN）と呼ぶ。

一つのニューロンは複数の入力チャンネルと一つの出力チャンネルを持ち、各チャンネルに割り当てられた重み（ウェイト）を使って全入力データの重み付き和を計算しその後何らかの非線形変換を施して出力する。この際、固定値のデータを入力するバイアス（オフセット）入力チャンネルが存在するのが普通である。ここでは、ニューロンを層状に並べて各層間のニューロンを完全に結合させて構成する多層ニュー

ラルネットワークの応用について考える。

2. 2 多層ニューラルネットワークの学習法

多層ニューラルネットワークの基本的な機能は入力データの組から出力データの組への写像を行うことである。1組の入力データのことを1「パターン」と呼ぶ。あるパターンに対して出力データになる変数の正しい値が既知であるときこの値を教師データと呼ぶ。入力データと教師データから構成される多数の組を用意して、これらの組に対して正しい入力・出力関係が構成されるようにネットワークの重みを調整することを「学習」と言う。

学習は、通常、出力データと教師データの差の二乗和が最小になるように重みを調整して行われる。出力データと教師データの差の二乗和は、誤差評価関数と呼ばれ次のように表される。

$$E = \frac{1}{2|X|} \sum_{\bar{x} \in X} c(x) [\bar{y}(\bar{x}) - \bar{y}_i(\bar{x})]^2$$

ここで、 X は入力パターンの集合を表す。誤差評価関数は重みの関数であるので、次式に従って重みの値を「時間 (t) 的に」変化させることが可能である。これは、重み空間における重み変化の「速度ベクトル」の向きが誤差評価関数の値が最も急に減少している向きになるように選ぶことに相当しており、「最大傾斜法」あるいは「勾配法」と呼ばれている非線形最適化法である。

$$\frac{dw_j^{(m)}}{dt} = -\frac{\partial E}{\partial w_j^{(m)}}$$

このように、重みの値が時間的に変化すれば、次式から明らかなように、誤差評価関数の値は時間に関して常に減少する。誤差逆伝播法はこの原理に基づく学習方法である。

$$\begin{aligned} \frac{dE}{dt} &= \sum_m \sum_i \sum_j \frac{\partial E}{\partial w_{ij}^{(m)}} \frac{dw_{ij}^{(m)}}{dt} \\ &= -\sum_m \sum_i \sum_j \left(\frac{\partial E}{\partial w_{ij}^{(m)}} \right)^2 \leq 0 \end{aligned}$$

誤差逆伝播法によるニューラルネットワークの学習は次のような順序で行われる。

- (1) ネットワークに入力データを与えて、入力層から出力層に向かって順方向の計算を行う。この際、各ニューロンにおいて計算された活性化関数の微分値は次の逆方向計算のために保存しておく。
- (2) ネットワークの出力データと教師データの差から構成された目的関数に基づいて、出力層から入力層の方向に向かって目的関数の重み微分値を順次計算する。この過程の計算は次式に基づいて行う。

- 1) 最終層の入力チャンネルへの結合路の重み

$$\frac{\partial E}{\partial w_j^{(M-1)}} = \frac{\partial E}{\partial x_j^{(M)}} \frac{df_M}{dX_j^{(M)}} x_i^{(M-1)}$$

- 2) 一般層の入力チャンネルへの結合路の重み

$$\begin{aligned} \frac{\partial E}{\partial w_j^{(m-1)}} &= \frac{\partial E}{\partial x_j^{(m)}} \frac{df_m}{dX_j^{(m)}} x_i^{(m-1)} \\ \frac{\partial E}{\partial x_j^{(m)}} &= \sum_{k=1}^{l^{(m)}} \frac{\partial E}{\partial x_k^{(m+1)}} \frac{df_{m+1}}{dX_k^{(m+1)}} w_k^{(m)} \end{aligned}$$

2. 3 多層ニューラルネットワークの特徴

逆問題解法を考えると多層ニューラルネットワークの持つ以下の特徴が利用される。

- (1) ある仮定のもとに、任意の関数を任意の精度で近似できる [3, 4]。
- (2) 構造、反復回数次第で汎化能力が選べる。内挿、平滑化を陽に意識しなくてもよい。
- (3) 入力データ (測定値) に含まれるノイズに強い。
- (4) 誤差評価関数の工夫次第で、種々の複合的非線形問題が解ける。
- (5) 多層構造を利用しているので非線形最適化計算のための計算量が著しく少なくなっている。未知パ

ラメータ (ウエイト) が多いにも拘わらずロバストな勾配法が実用的に使える。(誤差逆伝播法)

3. ニューラルネットワークによる逆問題解法

3. 1 教師データ内挿

ニューラルネットワークを用いた最も単純な逆問題解法は「教師データ内挿」を基本とする「単純関数当てはめ法」である。この方法は、順問題の計算が容易にできる場合や実験、観測によって入力データと出力データの関係が容易に多数得られる場合に用いられ種々の分野で実用化されている。ニューラルネットワークの学習法の立場からは、これは狭義の「教師あり学習法」に基づく方法である。

例えば、

順問題 $X \rightarrow O$, U : 計算可能、

逆問題 $O \rightarrow X$: 計算不可能、あるいは困難

のような場合には次の手順で学習させれば O を入力して X を出力する「写像装置」ができあがる。

- (1) $\{X_1, X_2, \dots, X_N\}$ について
 $\{O_1, O_2, \dots, O_N\}$ を計算する。
- (2) 入力を O 、出力を X とするニューラルネットワークを構成
- (3) 入力、 O_1, O_2, \dots, O_N に対して
 X_1, X_2, \dots, X_N を出力するように学習させる

単純関数当てはめ法による逆問題解析の例は多数報告されている。以下に、いくつかの例を示す。

(1) 磁気プローブデータからプラズマ位置形状解析 [5, 6]

トカマクプラズマ等の閉じこめ実験において最も基本的な量はトーラスプラズマの断面の形状と真空容器内におけるその位置である。これらの情報は、プラズマの周囲に設置された多数の磁気プローブ (コイル) で測定される磁場の強さに含まれている。プラズマの位置形状他のパラメーターを定めて磁気プローブの位置における磁場を計算するのは容易に実行可能であるが、この逆の問題を解くのは困難である。従来は狭い範囲のパラメータについて線形化を行い最小二乗フィッティングを行って解を求めるのが主流であった。この方法ではパラメーターが大きく変化する場合などには特別な対策が必要である。そこで、ニューラルネットワークを使った単純関数当てはめ法によって解を求める方法が考案された。いくつかの核融合プラズマ実験で成果を上げている。

(2) 静電プローブデータからプラズマ密度温度解析 [7]

プラズマの温度が余り高くない場合には、プラズマ中に直接固体電極 (静電プローブ) を挿入して得られる電流・電圧特性からプラズマの温度と密度を求めることができる。この場合には、プラズマ密度・温度を与えて電流・電圧特性の関係を導くことは容易であるが、逆の問題は一種のパターンマッチングの問題であってニューラルネットワークの利用が有効である。

(3) 放射光スペクトルから粒子の位相空間分布 [8]

シンクロトロン等の高エネルギー加速器中を荷電粒子が高速で回転するとき軌道の接線方向に光が発生する。これを、放射光あるいはシンクロトロン軌道放射等と呼ぶ。かつては、放射光は加速エネルギーを損失する有害な過程と考えられていたが現在ではこの放射光の優れた性質を利用していろいろな分野に応用がなされている。放射光のスペクトルは発生源である荷電粒子群の種々の情報を含んでいる。これらは、例えば、粒子数、粒子運動の相空間における分布形状等である。これらを放射光スペクトルの関数として陽に表現することはできないが単純関数当てはめ法を用いれば数値的に求めることが可能になる。

(4) 合成ガウス分布型スペクトルの単一ガウス分布への分離 [9]

物理学における実験や自然観測では光のスペクトルや粒子の速度分布のスペクトルなどを測定して解析する場合がしばしばある。これらのスペクトルは、多くの場合、ガウス分布をしているが、複数のガウス分布が接近しているときには、これらを合成したものが観測されるので単一のガウス分布に分離する必要がある。一つのガウス分布を特徴づけるパラメータは、スペクトル分布の中心位置、分布の幅、強さの3つであるので n 個のガウス分布が重なっていることがわかっている場合には $3n$ パラメータにいろいろな数値を設定してスペクトルを計算し、単純関数当てはめ法を適用して解析を行うことが可能である。

(5) 一般的な曲線当てはめ [10]

このほかにも、単純関数当てはめ法は、原理的には、非常に広い範囲の逆問題解析に適用可能である。しかし、計算時間、精度、その他の条件により、実際に有効に利用できる応用は限られている。

単純関数当てはめ法による逆問題解析の具体例として合成ガウス分布の分解について以下に整理しておく。この問題は、選点法によっても容易に解析できるので、比較のため、選点法を適用した場合の方法についても後述する。一般的な解法は次の通りである。

$$y(\lambda) = \sum_{m=1}^M A_m \exp\left(-\frac{(\lambda - B_m)^2}{W_m^2}\right)$$

$$o_j^{(n)} = y^{(n)}(\lambda_j), \quad u^{(n)} = y^{(n)}(\lambda)$$

* 教師データの作成：J 波長点、N パターン

$$A_m^{(n)}, B_m^{(n)}, W_m^{(n)} \longrightarrow o_j^{(n)} = y^{(n)}(\lambda_j)$$

* 教師データによる学習：J 入力・M 出力ネットワーク

$$o_j^{(n)} = y^{(n)}(\lambda_j) \longleftarrow A_m^{(n)}, B_m^{(n)}, W_m^{(n)}$$

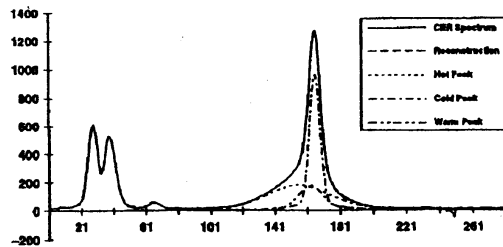
具体的な問題に適用した例として、DIIID トカマク装置における荷電交換粒子スペクトルの分離が報告されている [9] (第 2 図)。ニューラルネットワークに関するデータは次の通りである。

ネットワーク構造：280+34+15+6 (or 16)

入力データ：280 の波長における強度

出力データ：ピーク位置、強度、幅、(+13 データ)

学習データ：3000~4000 の解析済みデータ



第 2 図 DIIID トカマクの電荷交換粒子スペクトル解析

3. 2 残差最小化

単純関数当てはめ法では、教師データが存在する場合について、ニューラルネットワークを用いて逆写像を構成した (教師あり学習法) が、残差最小化法では順方向の写像関数 (条件式) が与えられているときに適当な選点集合を用意して選点における順写像の条件式の残差の絶対値がなるべく小さくなるようにニューラルネットワークを訓練して逆写像を構成する。教師データは不要である。

次のような関係を考える。

$$L(x) - y = 0 \quad x \in \Omega$$

ここで、L は微分、積分、そのほか具体的に計算手順を表すものなら何でもよい。ただし、L(x) は計算可能であるが、L の逆を計算することは困難である。具体的には、計測値 y に対応する物理量 x を求めたい。入力 y に対する出力を x とし、この x を使って L(x) を計算し y と比較して学習を進める。

4. 残差最小化法と評価関数

4. 1 誤差評価関数

ニューラルネットワークの誤差評価関数として方程式の残差の二乗を使う場合には下記のようないろいろな形の数式が利用可能である。

(1) 関数値誤差

$$E(\vec{p}) = \sum_{\vec{x}} (\bar{y}(\vec{x}, \vec{p}) - \bar{y}_i(\vec{x}))^2$$

(2) 微分値誤差 (微分方程式残差)

$$Dy(\vec{x}) = f(\vec{x})$$

$$E(\vec{p}) = \sum_{\vec{x}} (Dy(\vec{x}, \vec{p}) - f(\vec{x}))^2$$

(3) 積分値誤差 (積分方程式残差)

$$\int_{(q)} y(\vec{x}) d\vec{x} = f(q)$$

$$\int y(\vec{x}) C(\vec{x}, q) d\vec{x} = f(q)$$

$$E = \sum_q \left(\int_{(q)} y(\vec{x}, \vec{p}) d\vec{x} - f(q) \right)^2$$

4. 2 多層ネットワークの出力に対する微分積分演算

前節で述べたように誤差評価関数に微分値誤差や積分値誤差を用いる場合には多層ネットワークの出力に対する微分積分演算が必要である。これらは次のように表される。

(1) 微分

微分は解析的な表現が与えられる。簡単のため、3層ネットワークの場合について微分の解析的表現を求めると次のようになる。

$$y = \sum_j w_j^{(2)} \sigma \left(\sum_i w_i^{(1)} x_i \right) \Rightarrow \frac{\partial y}{\partial x_k} = \sum_j w_j^{(2)} \sigma' w_j^{(1)}$$

(2) 積分

積分は解析的な表現が与えられないので数値積分を行う必要がある。しかし、微分の場合と違いノイズによる誤差の累積効果はないので十分に実用に耐える演算が可能である。

$$I = \int_{\Omega} y(x) d\Omega \Rightarrow I \approx \sum_i a_i \bar{y}(\vec{x}_i, \vec{w})$$

4. 3 残差最小化法による微分方程式解法

◇次のような2階の微分方程式を解くものとする。

$$F(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}) = 0$$

ある選点における誤差評価関数は次のように表される。

$$E = \frac{1}{2} \left| F(x, y, \frac{dy}{dx}, \frac{d^2y}{dx^2}) \right|^2$$

この誤差評価関数の重み微分は次のように表される。

$$\frac{\partial E}{\partial w_{\alpha\beta}^{(m)}} = 2F \left(\frac{\partial F}{\partial y} \frac{\partial \bar{y}}{\partial w_{\alpha\beta}^{(m)}} + \frac{\partial F}{\partial y'} \frac{\partial \bar{y}'}{\partial w_{\alpha\beta}^{(m)}} + \frac{\partial F}{\partial y''} \frac{\partial \bar{y}''}{\partial w_{\alpha\beta}^{(m)}} \right)$$

◇この値を評価するために出力 y の入力 x 微分の重み微分が必要となる。

(1) ここでは、次に示すような出力 y の入力 x 1階微分の重み微分についてのみ考える。

$$\frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial y_{i_M}}{\partial x_i} \right)$$

(2) 関数値は次のように表される。

$$y_{i_M} = f_M(X_{i_M}^{(M)}), \quad X_{i_M}^{(M)} = \sum_{i_{M-1}=0}^{I_{M-1}} w_{i_{M-1}i_M}^{(M-1)} x_{i_{M-1}}^{(M-1)}$$

$$x_{i_m}^{(m)} = f_m(X_{i_m}^{(m)}), \quad X_{i_m}^{(m)} = \sum_{i_{m-1}=0}^{I_{m-1}} w_{i_{m-1}i_m}^{(m-1)} x_{i_{m-1}}^{(m-1)}$$

$$x_2^{(2)} = f_2(X_2^{(2)}), \quad X_1^{(2)} = \sum_{i=0}^1 w_{i1}^{(1)} x_i$$

(3) 導関数値は次のように表される。

$$\frac{\partial y_{iM}}{\partial x_i} = f'_M \sum_{i_{M-1}=0}^{i_{M-1}-1} w_{i_{M-1}iM}^{(M-1)} \frac{\partial x_{i_{M-1}}^{(M-1)}}{\partial x_i}$$

$$\frac{\partial x_{im}}{\partial x_i} = f'_m \sum_{i_{m-1}=0}^{i_{m-1}-1} w_{i_{m-1}im}^{(m-1)} \frac{\partial x_{i_{m-1}}^{(m-1)}}{\partial x_i}$$

$$\frac{\partial x_{i2}^{(2)}}{\partial x_i} = w_{i2}^{(1)}$$

(4) これらにより、出力の導関数値の重み微分は次のように表される。

まず、 $m=M-1$ の場合については、

$$\frac{\partial}{\partial w_{\alpha\beta}^{(M-1)}} \left(\frac{\partial y_{iM}}{\partial x_i} \right) = \left(f'_M \cdot x_{\alpha}^{(M-1)} + f'_M \cdot \frac{\partial x_{i_{M-1}}^{(M-1)}}{\partial x_i} \right) \delta_{\beta, i_{M-1}}$$

次に $m < M-1$ の場合については、

$$\frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial y_{iM}}{\partial x_i} \right) = f'_M \cdot \frac{\partial X_{iM}^{(M)}}{\partial w_{\alpha\beta}^{(m)}} \sum_{i_{m-1}=0}^{i_{m-1}-1} w_{i_{m-1}im}^{(m-1)} \frac{\partial x_{i_{m-1}}^{(m-1)}}{\partial x_i} + f'_M \cdot \sum_{i_{m-1}=0}^{i_{m-1}-1} w_{i_{m-1}im}^{(m-1)} \frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial x_{i_{m-1}}^{(m-1)}}{\partial x_i} \right)$$

(5) ここで、 X の重み微分は次のように計算される。

① $\frac{\partial x_{i_{n-1}}^{(n-1)}}{\partial w_{\alpha\beta}^{(n)}} = f'_{n-1} \cdot \frac{\partial X_{i_{n-1}}^{(n-1)}}{\partial w_{\alpha\beta}^{(n)}}$

$n=m$ から M について、次の②、③を繰り返す。

② $\frac{\partial x_{i_{n-1}}^{(n-1)}}{\partial w_{\alpha\beta}^{(m)}} = f'_{n-1} \cdot \frac{\partial X_{i_{n-1}}^{(n-1)}}{\partial w_{\alpha\beta}^{(m)}}$

③ $\frac{\partial X_{i_n}^{(n)}}{\partial w_{\alpha\beta}^{(m)}} = \sum_{i_{n-1}=0}^{i_{n-1}-1} w_{i_{n-1}in}^{(n-1)} \frac{\partial x_{i_{n-1}}^{(n-1)}}{\partial w_{\alpha\beta}^{(m)}}$

(6) 一方、 x の x 微分は次のように計算される。

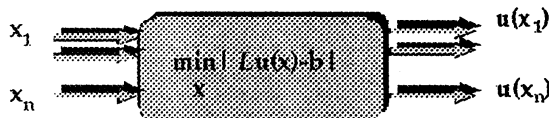
① $\frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial x_{i_{m+1}}^{(m+1)}}{\partial x_i} \right) = f'_{m+1} \cdot \left(\frac{\partial x_{\beta}^{(m)}}{\partial x_i} \right)$

$n=m$ から $M-1$ について、次の②を繰り返す。

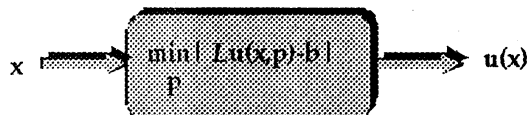
② $\frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial x_{i_n}^{(n)}}{\partial x_i} \right) = f'_n \cdot \frac{\partial X_{i_n}^{(n)}}{\partial w_{\alpha\beta}^{(m)}} \sum_{i_{n-1}=0}^{i_{n-1}-1} w_{i_{n-1}in}^{(n-1)} \frac{\partial x_{i_{n-1}}^{(n-1)}}{\partial x_i} + f'_n \cdot \left(\delta_{m,n-1} \delta_{\alpha, i_{n-1}} \delta_{\beta, i_n} \frac{\partial x_{\alpha}^{(n-1)}}{\partial x_i} + \sum_{i_{n-1}=0}^{i_{n-1}-1} w_{i_{n-1}in}^{(n-1)} \frac{\partial}{\partial w_{\alpha\beta}^{(m)}} \left(\frac{\partial x_{i_{n-1}}^{(n-1)}}{\partial x_i} \right) \right)$

◇通常の微分方程式解法との比較

微分方程式： $Lu(x) = b$ (or $\min |Lu(x)-b|$) を解く場合、通常の方法 (第3図) とニューラルネットワーク選点法 (第4図) の違いを示す。ニューラルネットワーク選点法では微分方程式の解が解析的な関数として得られるので計算結果が得られた後、選点以外の点における関数値が簡単に求まる。



第3図 通常法による微分方程式の解法



第4図 ニューラルネットワーク選点法による微分方程式解法

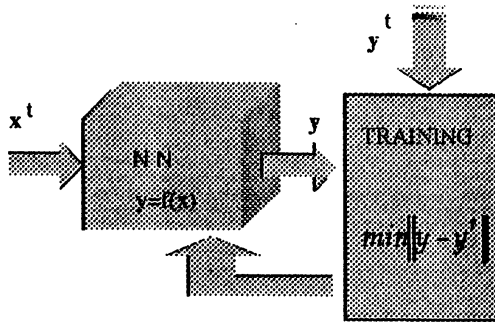
5. 逆問題パターン

ニューラルネットワークを用いて実行できる逆問題解法は次のように、教師データ内挿を基本とする「タイプ1」と残差最小化を基本とする「タイプ2～5」に分類することができる。これらについて以下の小節で簡単に述べるとともに次節において具体的な問題解決への応用例について詳しく述べる。

- タイプ1：単純関数当てはめ型
- タイプ2：透過CT型
- タイプ3：非対称アーベル変換型
- タイプ4：データ同化問題型
- タイプ5：係数決定問題型

5. 1 単純関数当てはめ型

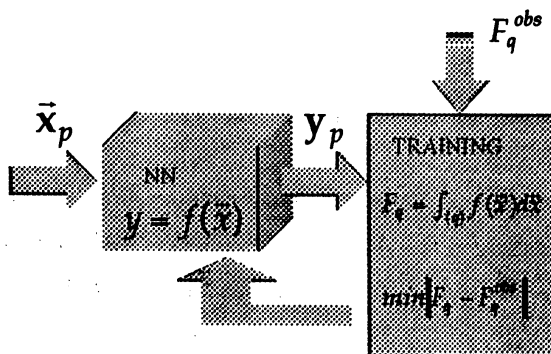
教師データのあるパターンが十分な数だけ用意できる場合には適用可能である。しかし、学習のために用意したパターンからはずれたところの計算や学習パターンの密度の低いところの計算は信頼性が低い。ネットワークの学習は第5図のように行われる。



第5図 ニューラルネットワーク単純関数当てはめ法による逆問題解法

5. 2 透過CT型

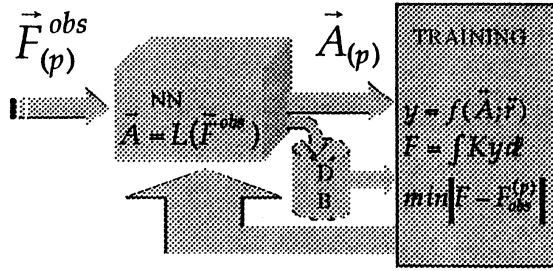
医療用X線CT画像解析で代表される問題である。ニューラルネットワークの優れた関数近似機能を利用して、座標の関数としての像強度を表す関数をニューラルネットワークで構成する（第6図）。



第6図 透過CT型問題のニューラルネットワークによる解法

5. 3 非対称アーベル逆変換型

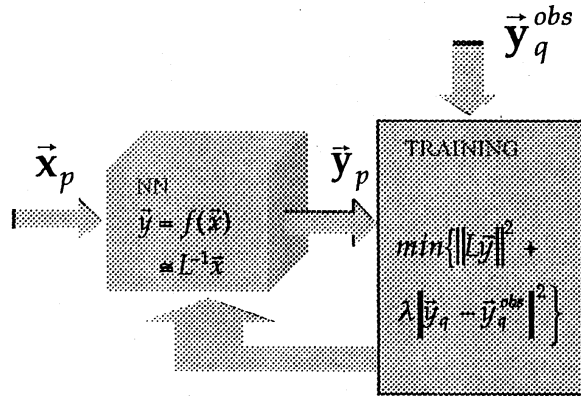
円柱対称のシステムについて断面を横切る投影経路に沿った線積分値の組が与えられたときに局所強度分布を計算するアーベル逆変換の拡張である。円柱対称の場合には像強度の等高線の形が同心円になるのに対して非対称な場合には等高線の形に関する情報を取り込むことが課題である（第7図）。



第7図 非対称アーベル逆変換型問題のニューラルネットワークによる解法

5.4 データ同化問題型

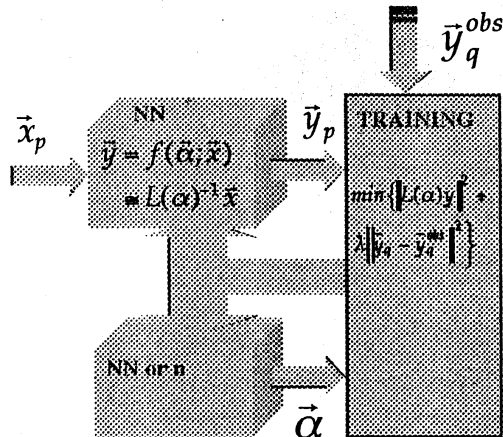
データ同化問題は、後に詳しく述べるように、支配方程式が既知でノイズを含む観測データがある場合に観測データが支配方程式に矛盾しないように整形する問題である。大規模な気象・気候シミュレーションで重要性が高まっている。支配方程式の残差と観測データの誤差を適当なペナルティ係数を使って結合しニューラルネットワークで解を求める（第8図）。



第8図 データ同化問題のニューラルネットワークによる解法

5.5 パラメータ評価問題型

物理・工学の問題や自然観察の問題等ではしばしば支配方程式はだいたいわかっているが含まれているパラメータの一部がわかっていないような問題に遭遇する。このとき利用可能な測定データと支配方程式から未知のパラメータを評価する問題がパラメータ決定問題である。この問題は上記データ同化問題の変形として取り扱うことが可能である（第9図）。



第9図 パラメータ決定問題のニューラルネットワークによる解法

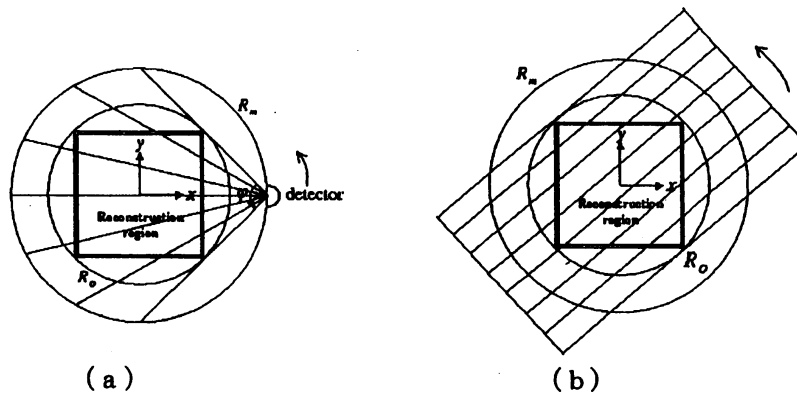
6. ニューラルネットワークによる逆問題解析の例

6. 1 少数データCT

透過CT型の逆問題の代表的な方法はフィルター補正逆投影法である。医療用のX線CTに関してはきわめて高精度で効率の良い方法が開発されている。しかし、これらは測定データである線積分値を与える投影経路の数が十分大きくかつ一様に分布している場合である。物理・工学の実験や自然観察では投影経路の数が必ずしも十分多くなく、また経路の分布も一様ではない。このような問題についても数多くの研究があるがニューラルネットワークを利用した方法は良い結果を与えることがモデル数値実験で確認できた [11, 12]。

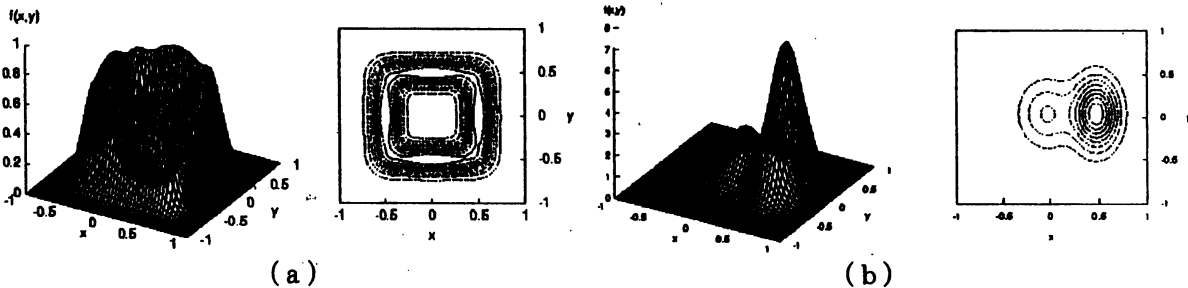
6. 1. 1 少数データCTモデル数値実験

第10図に示すような正方形の領域内の局所値 $f(x, y)$ の投影経路にその線積分値が測定データとして得られる場合について局所値 $f(x, y)$ を求める問題を考える。



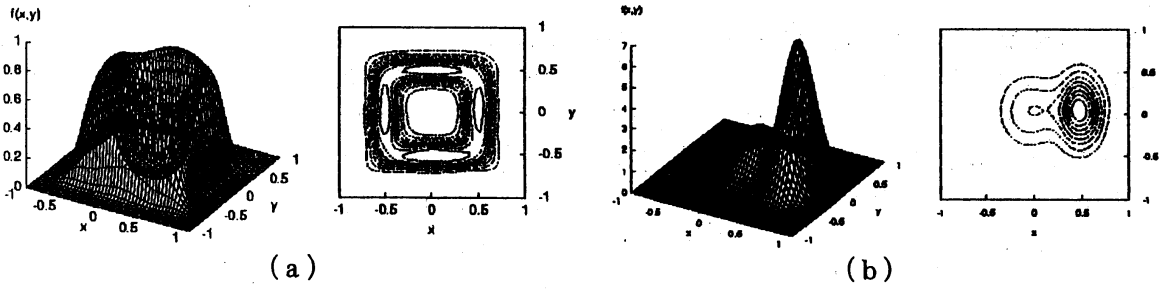
第10図 投影経路と幾何学的形状 (a) ファンビーム型、(b) 平行ビーム型投影経路

再現すべきモデル分布としては、第11図に示す井戸型分布と2重ピーク型分布を採用し、投影経路に沿う線積分値を計算しモデル測定値を作成した。標準的には、ファンビームの時には一つのファンにつき30経路、平衡ビームと時には一つの方向につき30経路を用意しこれを30点あるいは30方向用意して「測定データ」を用意した。しかし、極端に測定経路の少ない場合の数値実験では経路の数は全体で30本である。



第11図 モデル分布形状 (a) 井戸型分布、(b) 2重ピーク型分布

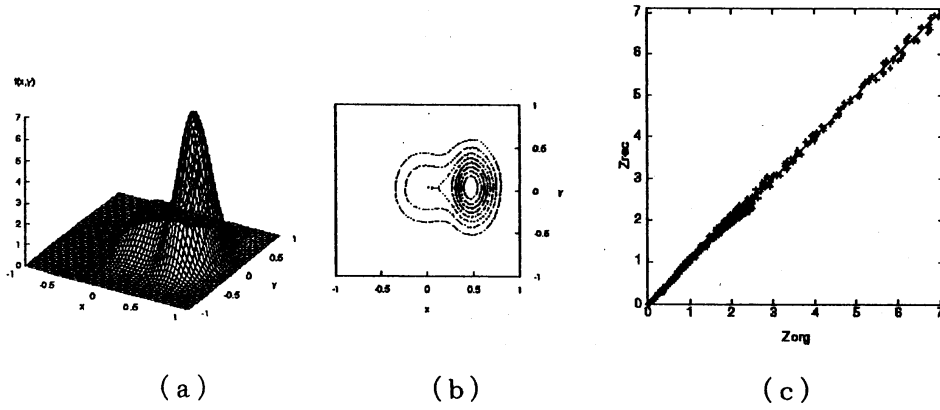
これらについて再構成例を以下に示す。第12図は測定データにノイズが含まれない場合についてファンビーム投影経路のデータを用いて再構成した。第13図は、平行ビームの投影データについての再構成例で、同じデータからFBP法(フィルター補正逆投影法)を用いた結果(第14図)と比較できる。このように投影経路の数が少ない場合にはニューラルネットワークを用いた方法がきわめて有利である。また、投影経路の数が極端に少ない場合の再構成例を第15図に示した。このようにデータ量が少なくともおおまかな再構成ができるのがこの方法の特徴である。



第 1 2 図 ノイズのない場合の再構成図

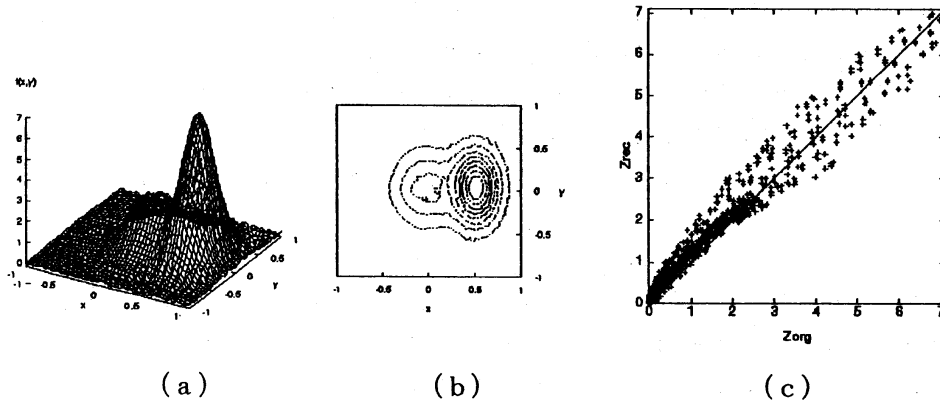
(a) 井戸型分布：積分値誤差=0.0042、関数値誤差=0.0135、

(b) 2重ピーク型分布：積分値誤差=0.0167、関数値誤差=0.0191



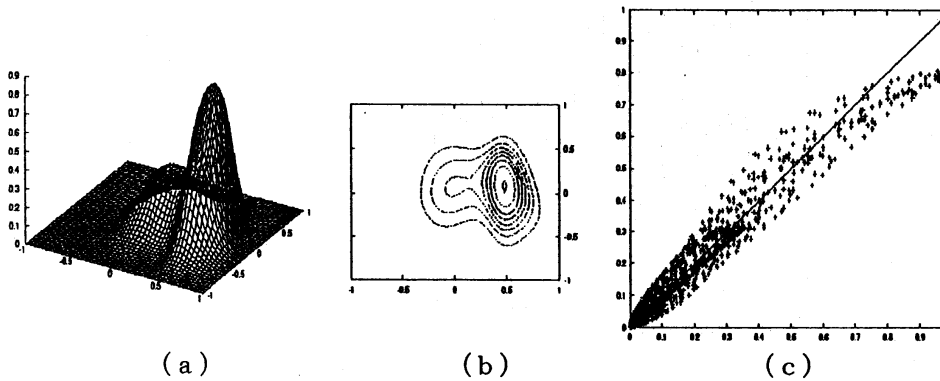
第 1 3 図 平行ビーム型投影経路のデータについての再構成例

(a) 鳥瞰図、(b) 等高線図、(c) 相関図



第 1 4 図 平行ビーム型投影経路のデータについての再構成例 (FBP法)

(a) 鳥瞰図、(b) 等高線図、(c) 相関図

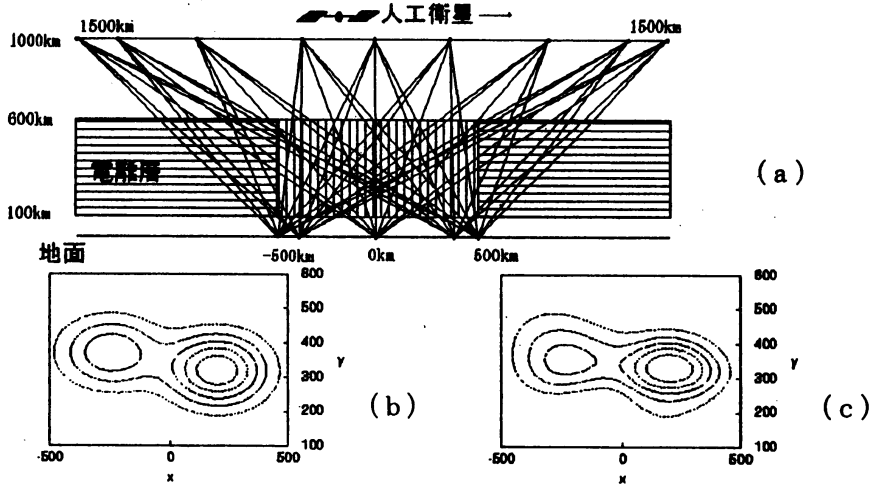


第 1 5 図 投影経路の数が極端に少ない場合の再構成例 (3 x 10ファンビーム)

(a) 鳥瞰図、(b) 等高線図、(c) 相関図

6. 1. 2 電離層プラズマ密度解析のモデル数値実験

人工衛星からの電波の位相遅れから電波の経路にそったプラズマ密度の線積分値が測定されるので、このデータが十分多く得られれば上記の少数データCTの手法でプラズマ密度の局所分布が再現可能である。ただし、この場合の問題点は水平に近い角度の経路が欠如していることである。モデルデータによる数値実験で、このような解析が困難であるにしても不可能ではないことが判明した(第16図)。



第16図 電離層プラズマ密度CTのモデル数値実験例
 (a) 電離層プラズマ密度CTの概念図 (b) モデル分布等高線図、(c) 再構成分布等高線図

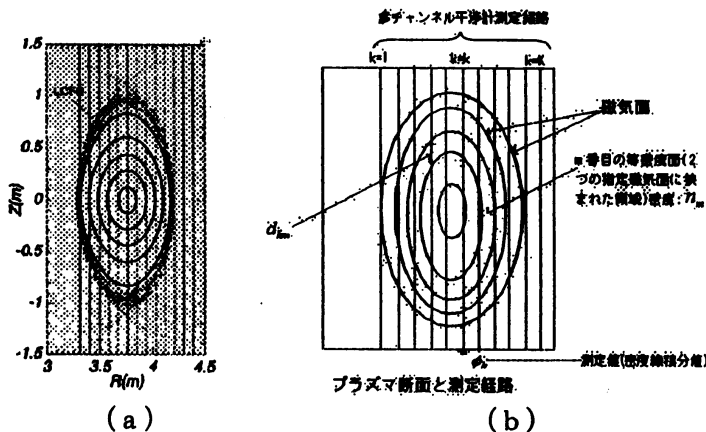
6. 2 非対称アーベル逆変換

6. 2. 1 アーベル逆変換

上記の電離層プラズマの密度解析は、プラズマ中に電磁波を入射するときその電磁波の周波数がプラズマ周波数に比べて十分大きいときには電磁波の位相変化が電磁波経路にあるプラズマ密度の線積分値に比例することに基づいている。この事実を利用すれば測定平面内でプラズマ密度分布の等高線が同心円状であるときにはこの問題はアーベル逆変換を行うことで線積分値の組から局所密度分布を再構成できる。核融合プラズマの密度分布解析にはこの方法が用いられるが、核融合のためのプラズマ閉じこめ装置の中では一般にはプラズマ密度の分布形状は同心円状ではない。このような逆変換を非対称アーベル変換と呼ぶ。

6. 2. 2 非対称アーベル逆変換

普通のアーベル変換と違って、非対称アーベル変換を行うには密度分布の等高線の形で定める必要がある。いろいろな方法が提案されており、実験条件に応じて使い分けられている。等高線の形の決定方法について何らかの仮定がなされた場合にも密度の決定にはいろいろの方法が使われている。比較的好まれる方法は、タマネギの皮をむくように密度を決めていくオニオンピーリング法である(第17図)。

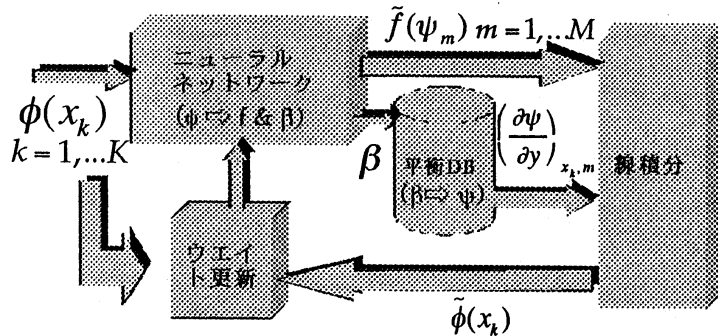


第17図 オニオンピーリング法による密度分布再構成
 (a) プラズマ断面の磁気面図(密度等高線に相当) (b) 密度再構成概念図

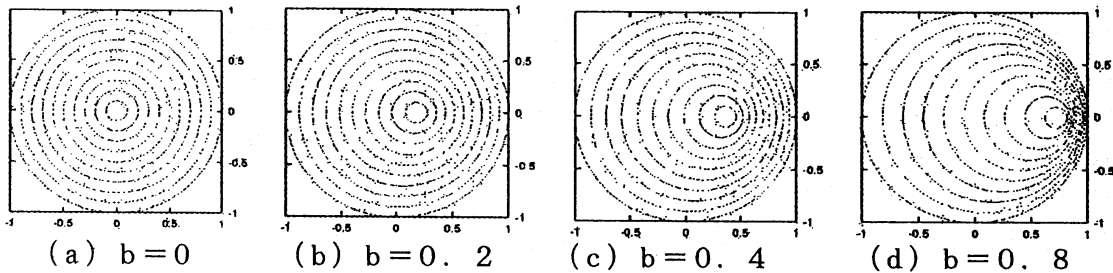
最近の大型核融合閉じこめ装置での実験はよく制御されているので、一連の実験におけるMHD平衡の変化が少ないパラメータの値の時間変化で表せる場合が多い。MHD平衡の磁気面は、ほぼ密度等高線を見てもよいので、MHD平衡シリーズを組み込むことによって非対称アーベル変換が矛盾なく実行できる。しかし、MHD平衡から得られるデータと密度線積分値の関係は複雑でこの解析を実際に行うことは従来の方法では極めて困難である。ニューラルネットワーク選点法を使うとデータベースの組み込みが極めて簡単である（第18図）。この方法をモデルデータについて数値実験を行ったところよい結果を得た〔第19～21図〕[13, 14]。また、現実のプラズマの密度解析に適用したところ、別の方法で求めた密度分布とよい一致を示した。さらに、MHD平衡パラメータとして使ったベータ値も別の測定によるベータ値の変化とよい一致を示した（第22図）。このことは密度等高線の形も正しく再現されていることを示している。

MHD平衡データベースと結合するとき重要な量は各線積分経路が磁気面を切る長さが採用したパラメータの関数として与えられることである。我々の方法では、この部分に関してはスプライン補間を用いた。なお、線積分経路が磁気面を切る長さは次式のメトリック量を差分表示したものである。

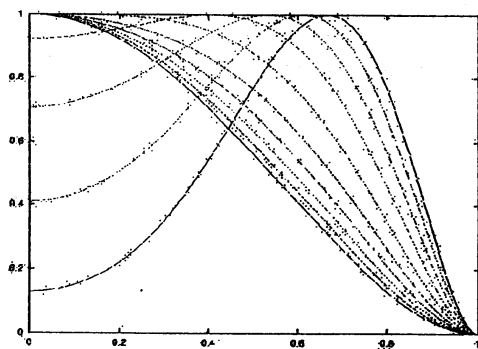
$$\phi(x_k) = \int_{x_k} f(\psi) dy = \int_{x_k} \frac{f(\psi)}{\left(\frac{\partial \psi}{\partial y}\right)_{x_k}} d\psi \quad \phi(x_k) \approx \sum_{m=1}^M \frac{\tilde{f}(\psi_m)}{\left(\frac{\partial \psi}{\partial y}\right)_{x_{k,m}}} \Delta\psi_m$$



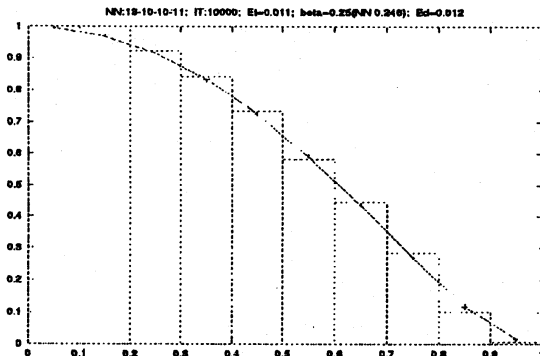
第18図 ニューラルネットワークを用いた非対称アーベル変換



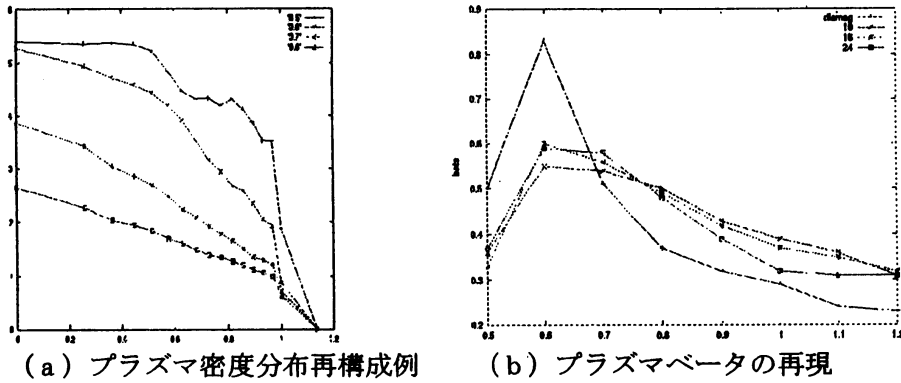
第19図 非対称アーベル変換モデル数値計算のためのモデル磁気面



第20図 モデル密度分布例



第21図 モデル計算例



(a) プラズマ密度分布再構成例 (b) プラズマベータの再現
第2.2図 実際の実験データを用いた密度分布再現

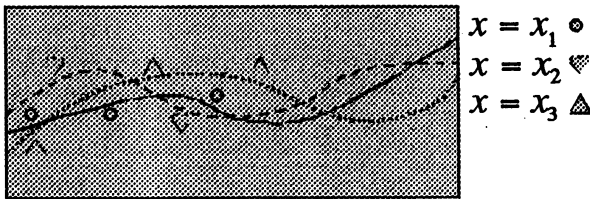
6.3 データ同化問題

すでに記したように、データ同化問題は支配方程式とノイズを含む観測データの組からより適切なデータの組を探し出すことが目的でいろいろな方法が提案され、また実用化されている。次のような条件の下でデータの整形を行うことが要求される(第2.4図)。

支配方程式に従う

- * 拘束条件を与える時刻が場所によって変わる
- * 測定誤差が含まれる

⇒ 同一時点で誤差のない条件に整形



第2.4図 データ同化の概念図

この問題を解く方法の一つは支配方程式の残差と観測値と真値の差がともに小さくなるように目的関数を構成し最適化問題を解くことである。ニューラルネットワークを用いる方法もこの方針で定式化された。

目的関数の中で拘束条件を組み合わせる方法は色々ある。

微分方程式+単純拘束条件

単純拘束条件：測定値そのまま

初期値、境界値 (含：不規則配置、誤差)

微分方程式+積分型拘束条件

積分型拘束条件：線積分、面積分、体積分

我々は、ローレンツ方程式に従う系についてこの方法の適用可能性を確認した。次に示すようなローレンツ方程式をルンゲ・クッタ法で時これを真値とした。これに、ランダムノイズを加えて擬似的実験データを作り、このデータとローレンツ方程式を組み合わせでデータ同化の方法の妥当性を確認した [15]。

$$\frac{dx}{dt} = \sigma(y - x)$$

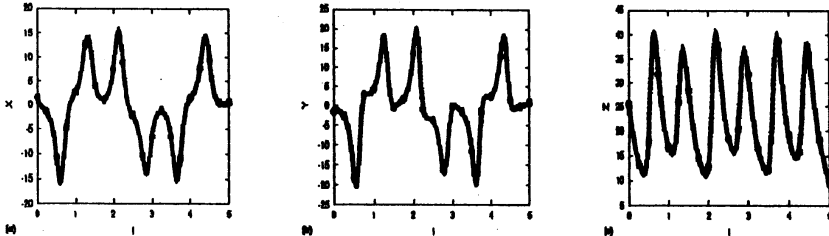
$$\frac{dy}{dt} = rx - y - xz$$

$$\frac{dz}{dt} = xy - bz$$

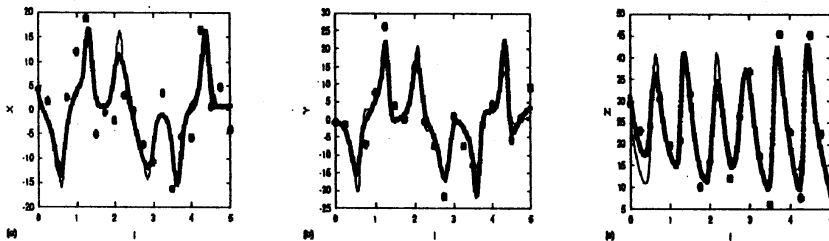
$$\sigma = 10, r = 28, b = 8/3$$

数値実験は、ノイズレスの場合、低ノイズレベルの場合(第2.5図)、高ノイズレベルの場合(第2.6図)について行った。また、データ同化領域の幅を増大すると、一般に、計算は困難になることを鑑み、

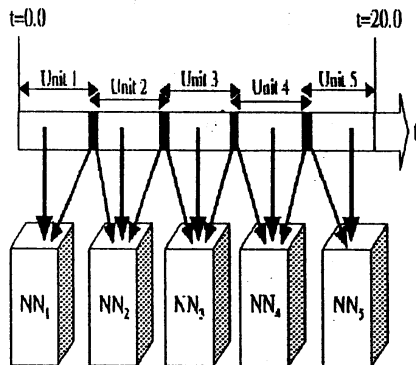
我々はデータ同化領域を領域分割しそれぞれの領域に別々のニューラルネットワークを割り当てて計算することを試みた（第27図）。この方法によれば原理的にいくらでも広い領域を対象とできることを確認した（第28図）。



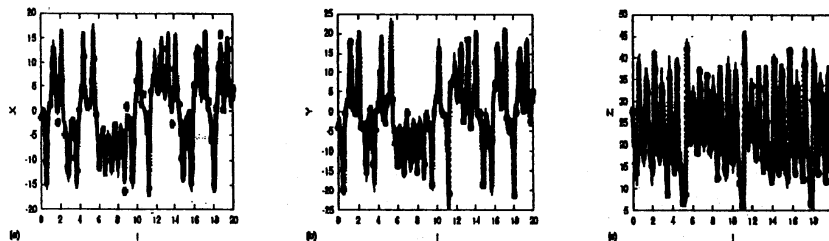
第27図 ローレンツモデルのデータ同化例（低ノイズレベル）



第26図 ローレンツモデルのデータ同化例（高ノイズレベル）



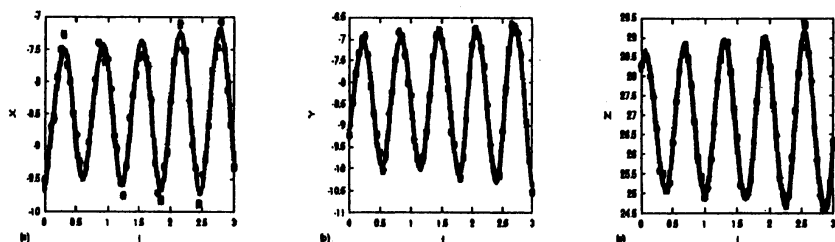
第27図 同化ウィンドウが広い場合の並列計算法



第28図 長区間同化の例

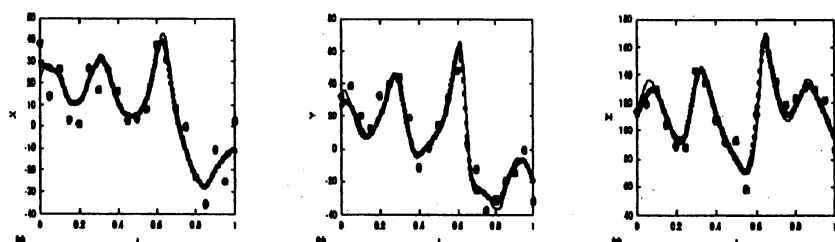
6. 4 パラメータ評価問題

前節のデータ同化の方法を拡張するとニューラルネットワークはパラメータ評価問題にも適用可能である。ローレンツモデルを使い、いくつかの条件の下でパラメータ評価を試みた〔第29～32図）[16]。



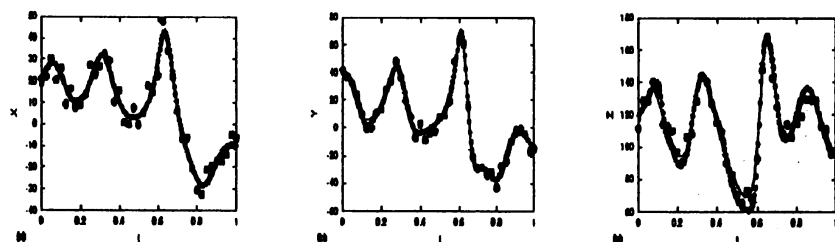
第29図 パラメータ評価計算の例 (弱非線形性、低ノイズレベル)

$$\left. \begin{array}{l} \sigma = 10.0 \\ r = 28.0 \\ b = 8/3 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \sigma = 9.960 \\ r = 27.970 \\ b = 2.666 \end{array} \right.$$



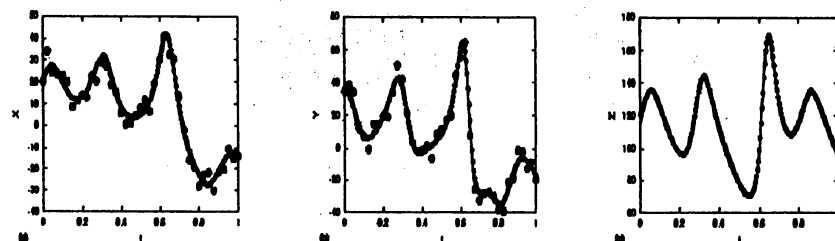
第30図 パラメータ評価計算の例 (強非線形性、高ノイズレベル)

$$\left. \begin{array}{l} \sigma = 16.0 \\ r = 120.1 \\ b = 4.0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \sigma = 15.682 \\ r = 119.229 \\ b = 3.851 \end{array} \right.$$



第31図 パラメータ評価計算の例 (強非線形性、中ノイズレベル、第3式欠損)

$$\left. \begin{array}{l} \sigma = 16.0 \\ r = 120.1 \\ b = 4.0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \sigma = 15.735 \\ r = 119.564 \\ b = *** \end{array} \right.$$



第32図 パラメータ評価計算の例 (強非線形性、中ノイズレベル、z測定欠損、r既知)

$$\left. \begin{array}{l} \sigma = 16.0 \\ r = 120.1 \\ b = 4.0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \sigma = 15.548 \\ r = (120.1) \\ b = 3.839 \end{array} \right.$$

7. まとめ

ニューラルネットワークを用いることによっていろいろなタイプの逆問題計算が可能であることを示した。まず、医療用CTに代表される透過型CTについて小雨数とウエイデートの場合に有効な方法を提案し実証した。拘束条件付きの逆問題の例としてプラズマ密度解析のための非対称アーベル変換に応用して満足すべき結果を得た。パラメータ評価問題の例として「誤差を含んだ観測」データからのローレンツ方程式の係数決定を行い満足すべき結果を得た。ニューラルネットワークを用いた解法は色々な種類の数値演算を統合的に扱えるという特徴を持っており実際のデータ同化問題の解決にも有効であると考えられる。ニューラルネットワークの学習課程のさらなる高速化を追求すると共に、本方法の理論的（数理的）解析を進める必要がある。

参考文献

- [1] Charles W. Groetsch, *Inverse Problems in Mathematical Sciences* (Friedr. Vieweg & Sons, 1993, Braunschweig)
- [2] (for example) R. Beale, T. Jackson, *Neural Computing: An Introduction* (IOP Publishing Ltd., Bristol, 1990)
- [3] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks* 2, 359 (1989).
- [4] K. Funahashi, On the approximate realization of continuous mappings by neural networks, *Neural Networks* 2, 183 (1989).
- [5] J.B. Lister, H. Schnurrenberger, *Nuclear Fusion* 31, 1291 (1991).
- [6] R. Yoshino, J. Koga, T. Takeda, *Fusion Technology* 30, 237 (1996).
- [7] L. Lagin, D. Manos, M. Diesso, *Bull. APS Vol. 36, No. 9 7T13* (1991).
- [8] J. Koga, T. Takeda, *Nucl. Instr. Meth. in Phys. Res. A363*, 580 (1995).
- [9] D.R. Baker, et al., *Plasma Phys. Control. Fusion* 36, 109 (1994).
- [10] C.M. Bishop, *Rev. Sci. Instr.* 63, 4450 (1992).
- [11] 馬笑峰、竹田辰興、ニューラルネットワークによるCT像再構成法、*日本応用数理学会論文誌*、10 (2000) 145-161.
- [12] X.F. Ma, M. Fukuhara, T. Takeda, Neural network CT image reconstruction method for small amount of projection data, *Nuclear Instruments and Methods in Physics Research A* 449 (2000) 366-377.
- [13] Xiao Feng Ma, Tatsuoki Takeda, Asymmetric Abel inversion by neural network for reconstruction of plasma density distribution, *Nucl. Instr. Methods in Phys Res. A* (submitted).
- [14] Xiao Feng Ma, Makoto Fukuhara, Tatsuoki Takeda, Kenji Tanaka, Kazuo Kawahata, Tokihiko Tokuzawa, Shigeki Okajima, Asymmetric Abel inversion by neural network for density profile reconstruction of LHD plasma, in preparation.
- [15] Ali Liaqat, Makoto Fukuhara, Tatsuoki Takeda, Application of neural network collocation method to data assimilation, *Computer Physics Communications* 14 1/3 350 (2002).
- [16] Ali Liaqat, Makoto Fukuhara, Tatsuoki Takeda, Optimal estimation of parameters of dynamical system by neural network collocation method, *Computer Physics Communications* (submitted).