

翌日最大電力需要想定における ニューラルネットワーク構成の最適化

小野田 崇

1. はじめに

翌日最大電力需要想定問題には従来、回帰モデルを中心とした研究がなされてきた。最近では、要求する処理に対して具体例を与えるだけで自動的にネットワーク内の結合荷重を変化させ、教えた信号に出力を調整する学習能力と非線形性を有するような任意関数を近似できる能力を持つニューラルネットワークをこの翌日最大電力需要想定問題に適用する研究が進められている。その際、最大電力需要想定に適用されるニューラルネットワークの構成をいかに選定するかが重要な課題となっている。

本報告では、翌日最大電力需要想定問題に回帰モデルを適用した場合と経験的に構成の決められたニューラルネットワークを適用した場合の予測能力の比較結果について述べる。さらに、ニューラルネットワークを適用する際の、最適なネットワーク構成の設計方法について述べ、提案する設計方法を用いた場合と用いないで経験的にネットワークの構成を決めた場合とでの、予測能力の比較結果について報告する。

2. 翌日最大電力需要想定問題

電力会社の中央給電指令所では、火力発電所などの設備は立ち上がりに時間がかかるため、当日午前中に翌日の最大電力需要を予測し、設備運用計画を作成しなければならない。このことは、最大電力需要予測問題が表1に示されているように、予測前日までの実績値と、予測を行なう当日と予測される翌日との予測値に基づいて行なう予測問題であることを意味する。

おのだ たかし (財)電力中央研究所情報研究所
〒201 狛江市岩戸北 2-11-1

表1: 翌日最大電力需要予測に利用できるデータ

	翌日	当日	前日	...
最大電力需要	?	?	実績値	実績値
天気	予報値	予報値	実績値	実績値
平均気温	予報値	予報値	実績値	実績値
最高気温	予報値	予報値	実績値	実績値
最低気温	予報値	予報値	実績値	実績値
曜日	既知	既知	実績値	実績値

3. 回帰モデルでの予測

本章では、ある電力会社のデータに基づき、1990年に当所で開発した重回帰分析に基づく翌日最大電力需要想定モデル [1] の概要とその問題点について述べる。

モデルは以下の式で表される。

$$\begin{aligned}
 PEAK_t = & a_0 + a_1 \cdot t + a_2 \cdot \overline{PEAK}_{t-1} \\
 & + a_3 \cdot MON_t + a_4 \cdot CD_t \\
 & + a_5 \cdot (CD_t - \overline{CD}_{t-1}) \\
 & + a_6 \cdot HD_t \\
 & + a_7 \cdot (HD_t - \overline{HD}_{t-1}) \\
 & + a_8 \cdot WT_t + e_t
 \end{aligned}$$

各変数の意味は以下のとおりである。

$PEAK_t$: t 日の修正最大電力需要

\overline{PEAK}_{t-1} : 以下の式で計算される $t-1$ 日までの修正最大電力需要の指数平滑値 $\overline{PEAK}_{t-1} =$

$$\begin{cases} 0.4 \cdot PEAK_{t-1} + 0.6 \cdot \overline{PEAK}_{t-2} & t-1 \neq \text{特殊日} \\ \overline{PEAK}_{t-2} & t-1 = \text{特殊日} \end{cases}$$

MON_t : 月曜ダミー

$$MON_t = \begin{cases} 1 & t = \text{月曜} \\ 0 & t \neq \text{月曜} \end{cases}$$

CD_t : 次の式で計算される冷房需要に関する変数

$$CD_t = \max(x_t - 20, 0)$$

ここで x_t は t 日の平均気温を表している。

\overline{CD}_t : 以下の式で計算される上記の変数 CD の t 日までの指数平滑値

$$\overline{CD}_t = \begin{cases} 0.4 \cdot CD_t + 0.6 \cdot \overline{CD}_{t-1} & t \neq \text{特殊日} \\ \overline{CD}_{t-1} & t = \text{特殊日} \end{cases}$$

HD_t : 次の式で計算される暖房需要に関する変数

$$HD_t = \min(x_t - 16, 0)$$

\overline{HD}_t : 以下の式で計算される上記の変数 HD の t 日までの指数平滑値

$$\overline{HD}_t = \begin{cases} 0.4 \cdot HD_t + 0.6 \cdot \overline{HD}_{t-1} & t \neq \text{特殊日} \\ \overline{HD}_{t-1} & t = \text{特殊日} \end{cases}$$

WT_t : t 日の天候を表す変数

$$WT_t = \begin{cases} 4 & \text{雨・雪, かつ } x_t < 18 \\ 2 & \text{曇り, かつ } x_t < 18 \\ 1 & \text{薄曇り, かつ } x_t < 18 \\ 0 & \text{晴れ, または } x_t \geq 18 \end{cases}$$

なお回帰係数の推定には、特殊日を除く至近 180 日のデータを用いる。

このような回帰モデルを用いた予測手法には 2 つの問題点がある。問題点の一つは、モデルを構築する担当者が最大電力需要の変動に通じている必要があるばかりでなく、重回帰分析などの統計解析に関する知識も有している必要があることである。もう一つの問題点は、最大電力需要の振舞いが、提示したモデルで考慮している気象条件などに対して、非線形であるにも関わらず、非線形性を容易に扱えないために予測誤差が大きくなる点である。

4. ニューラルネットワークでの予測

ニューラルネットワークには、階層型、相互結合型など様々な種類がある。翌日最大電力需要想定のような数値を予測する問題では、ニューラルネットワーク

の有する任意の非線形連続関数を近似できる能力 [2] を利用することが有効であると考え、階層型ニューラルネットワークを本問題に適用した。階層型ニューラルネットワークの基本的な構造は、入力層、中間層、出力層からなり、各層は複数のユニットから構成される (図 1)。

ここでは、ニューラルネットワークへ入力した学習データ、適用した学習方法、及び予測結果について説明する。

4.1 学習データ

重回帰モデルとの予測能力の比較を行なうため、ニューラルネットワークの学習に用いるデータの項目としては、できる限り重回帰モデルで採用している項目と同じものを採用した。項目選定のために最初に注意したことは、2 章で述べた通り、翌日最大電力需要予測が殆んど気象についての予報値を基に行なわれるため、予報値の予測精度が極端に悪いものは避けなければならないということである。例えば、夏季における湿度と最大電力需要との関係は相関が非常に高いが、湿度の予測誤差は平均 10% 以上と非常に大きいので、翌日最大電力需要の予測には利用できないと判断した。上述のようにして入力項目を絞り、絞られた項目と最大電力需要の変動との統計分析を基に以下のようなデータを採用した [3]。

- 夏季の場合 (6 月末から 9 月末迄)、学習時の入力データとしては、当日の平均気温、当日の最高気温、当日の最低気温の 3 種類を採用し、その値には実績値を使用した。また、予測時にはこれらのデータの予報値を採用した。

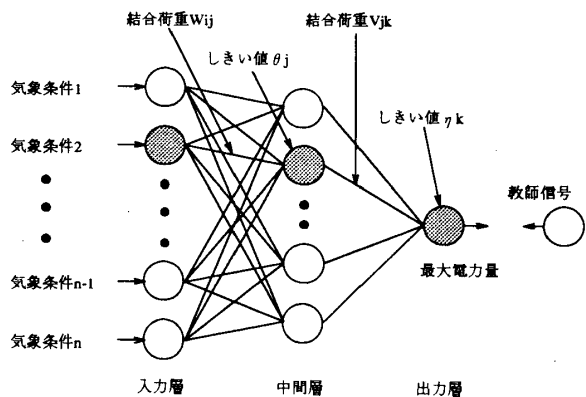


図 1: 階層型ニューラルネットワーク

- 冬季の場合(11月から3月末迄)、学習時の入力データに前々日の最大電力需要、当日の天候、当日の平均気温、当日の最高気温と当日の最低気温の5種類を採用し、その値には実績値を採用した。また、予測時にはこれらのデータの予報値を採用した。
- 学習時の出力ユニットへの教師データとしては、最大電力需要から特約分¹を引き、休振り等減電²を加えた修正最大電力需要の実績値を採用した。

4.2 学習

階層型ニューラルネットワークの学習には一般に誤差逆伝播学習法が用いられるが、この学習法は学習速度が遅い、極小解に捕まるなどの問題点がある。そこで、本研究ではニューラルネットワークの学習として、一般的な誤差逆伝播学習に学習速度を速め、最小解への促進を図るための学習制御パラメータの最適化ルーチンを組み込んだ手法を採用した。この学習制御パラメータの最適化ルーチンには、DSC 探索方法(Davis, Swann, Campey search method)を適用している。

ニューラルネットワークの入力データには修正最大電力需要、平均気温など単位の違うデータを使用するため、実際の学習、予測の際には値がスケールに依存しないように、上述した入力データの各値を[-1,1]間に規格化したものを入力信号とした。また、教師信号としては、修正最大電力需要を同様に[0,1]間に規格化したものを使用した。

学習データのニューラルネットワークへの提示は、夏季は10日分を1組として、冬季は15日分を1組として以下の手順を用いて行なった。

1. ネットワークの結合荷重、しきい値を疑似乱数(が慣用的)を使用して初期化する。
2. 予測日の至近10日(または、15日)分の学習データを学習させる。
3. 学習の収束条件を満足したら学習を止め、予測を行なう。
4. 予測終了後、その日の学習データを作成する。

¹産業用等大口需要家等に適用される特約分の電力需要。現在は、需給調整契約電力と言われる。

²企業グループ等の休業により減る電力需要。

5. 10日(または、15日)分の学習信号より一番古いものを削除し、前段階で作成された学習信号を加える。
6. 予測対象日がなくなったら終了、予測対象日が残っている場合は、1.へ。

上述のように10日(または、15日)単位で学習を行なったのは、ニューラルネットワークにできるだけ至近の最大電力需要の変動を学習させたかったのと、学習時間をできるだけ削減するためである。また、中間ユニット数は実験的に予測誤差の小さくなるものを採用した。夏季の場合は、中間ユニット数を3個とし、冬季の場合は4個とした。

4.3 予測シミュレーション

実際に最大電力需要を予測する際に利用可能な情報のみを適用して予測シミュレーションを行い、出力された予測値と最大電力需要の実績値とを比較した結果を示す。シミュレーションは、夏季(7月初旬から9月末)について1988年、1989年の2ケース、冬季(11月中旬から3月末)について1988年、1989年の2ケースの計4ケースについて行った。また、予測誤差の評価には、平均自乗誤差の平方根(RMSE)と平均絶対パーセント誤差(MAPE)を用いた。

表2にニューラルネットワークを用いた予測シミュレーション結果を、表3に3章で述べた重回帰モデルによる予測結果を示す。

また、1988年度の最大電力需要実績値とニューラルネットワークによる予測値および重回帰モデルによる予測値のグラフを図2に示す。

ニューラルネットワークによる予測結果と重回帰モ

表2: ニューラルネットの予測誤差

	RMSE(MW)	MAPE(%)
1988年度夏季	519.2	2.50
1988年度冬季	298.9	1.59
1989年度夏季	605.5	2.56
1989年度冬季	351.5	1.80

表3: 重回帰モデルによる予測誤差

	RMSE(MW)	MAPE(%)
1988年度夏季	668.4	3.27
1988年度冬季	298.8	1.58
1989年度夏季	665.7	2.90
1989年度冬季	372.4	1.90

5. ニューラルネットワーク構成の最適化

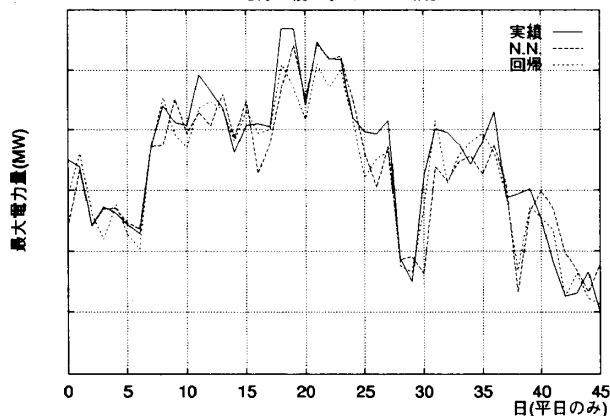
本章では、ニューラルネットワーク構成上の問題点、それを解決するための情報量基準、および、情報量基準を用いて決められたニューラルネットワークによるシミュレーション結果について述べる。

5.1 ニューラルネットワーク構成上の問題点

前章までで、電力需要想定問題にニューラルネットワークを適用することの有効性を示した。しかし、ニューラルネットワークを用いる場合の問題の一つとして、対象とする問題にニューラルネットワークを適用して近似したモデルの質をいかに測定するかがある。この近似の質は、近似しようとする非線形で連続な入出力関係の複雑さばかりでなく、適用するニューラルネットワークの構造にも依存する。この質の測定問題は、ニューラルネットワークの構造の視点から見ると、ニューラルネットワークの最適なパラメータ数をどのように推定したらよいかという問題であり、入出力関係のサンプルだけから、非線形で連続な入出力関係を近似するのに必要なニューラルネットワークの中間ユニット数を推定する問題に帰着する。

この問題の難しさは、パラメータの数を増加させれば教師信号に対する出力誤差は減少するが、未学習のデータに対する誤差は増加するという点にある。一般にこのような現象は「過学習」と呼ばれる。例えば、次の2つのニューラルネットワークを考えてみる。一つは中間ユニット数10個のニューラルネットワークであり、もう一つは中間ユニット数100個のニューラルネットワークを想定する。1000個の教師信号に対して十分な学習が済んだ後では、100個の中間ユニットを有する大きなニューラルネットワークの方が、10個の中間ユニットを有する小さなネットワークより教師信号に対する出力は良いものが得られる。しかし、未知のデータに対しては大きなニューラルネットワークの方が小さなネットワークより良くない出力を導く。つまり、既知の入力データに対して精度の高い出力を生成しようとする事と未知の入力データに対して精度の高い出力を予測することの間にはトレードオフの関係が存在する。このような既知の入力データと未知の入力データの間不均衡は、推定されるパラメータ数の増加とともに大きくなる。

最大電力実績と予測1988年度夏季



最大電力実績と予測1988年度冬季

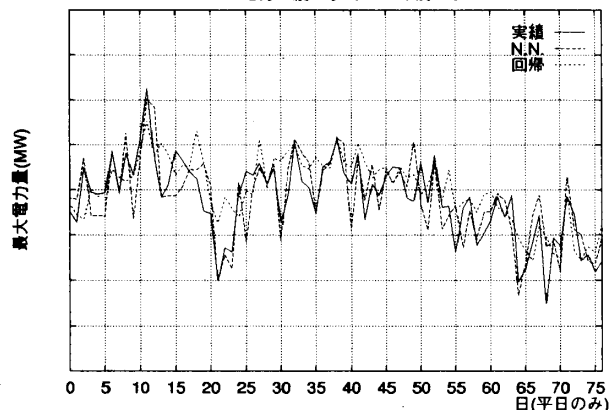


図2: 1988年度最大電力実績値と予測値

デルでの予測結果を比較すると、表2、3より夏季の場合、ニューラルネットワークを用いた予測の方が重回帰モデルより実績値に近い結果が得られていることがわかる。さらに、夏季の場合、ニューラルネットワークを用いた予測のRMSEが重回帰モデルを用いた場合よりも小さく、実績値に対する予測値の大きなはずれがニューラルネットワークを用いた場合の方が小さい。また、図2よりニューラルネットワークを用いた予測が夏季の最大電力需要の大きな変動に十分に対応していることがわかる。

冬季についても同様に、ニューラルネットワークを用いた予測で、重回帰モデルによる予測と同等以上の結果が得られている。そして、図2より冬季の場合についても、ニューラルネットワークを用いた予測の方が重回帰モデルを適用した予測よりも、最大電力需要の変動に十分に追従していることがわかる。

5.2 情報量基準

前節の問題を解決するアプローチとしては、情報量基準を用いたアプローチが代表的である。情報量基準と言えば、重回帰モデルなどのパラメータ数設定などに使われる AIC (*Akaike's Information Criterion*) が有名である [4]。しかし、ニューラルネットワークのパラメータ数設定に AIC を直接適用することは難しい。ニューラルネットワークの場合、最適なパラメータと実際にニューラルネットワークの学習によって推定されるパラメータの間には、次の 2 つのゆらぎが存在する。

- 統計的ゆらぎ

最適パラメータは、学習データが無限に与えられた場合に求められるパラメータである。実際には、有限個の学習データからある種の方法で最適パラメータの近似値を求めているにすぎない。このために起こるゆらぎ。

- 学習に伴うゆらぎ

ニューラルネットワークの学習は一般に非線形連立方程式となるため、特殊な場合を除き解析的に解くことは困難である。そこで反復法を用いた近似計算によって最適パラメータを求めることになる。しかし、一般に数値演算によって推定されたパラメータは初期値に依存し、必ずしも最適パラメータと一致せず、推定されるパラメータは、最適パラメータの周辺にばらつくことになる。このために起こるゆらぎ。

ニューラルネットワークの適切なパラメータ数を決定する情報量基準は、上記 2 つのゆらぎを評価できる必要がある。

情報量基準として有名な AIC では、モデルが faithful であるという条件の下で、統計的ゆらぎを評価している。しかし、学習に伴うゆらぎを評価してはいない。また、AIC は学習データが無限に近い状況において導かれているため、データ数が少ない場合には評価が難しい。

学習に伴うゆらぎを評価するために提案された情報量基準 $IC(K, N)$ [5] は、統計的なゆらぎについて評価していない。また、 $IC(K, N)$ は以下の仮定の下に導出されており、この仮定を満たす学習データを観測することは難しい。

仮定: 十分大きな学習データ N が与えられたとする。 N から l ($l \ll N$) を抽出して学習を行なったとき、これによって得られたパラメータ推定値の平均対数尤度は学習データ N から l を抜いた信号に対する前述で得られたパラメータ推定値での対数尤度の平均で十分よく近似できると期待できる。また、学習データ l より推定される推定パラメータと学習データ N より推定される推定パラメータとが互いに近傍に存在する。

また、AIC で必要とされるモデルが faithful である条件を必要としない情報量基準 NIC [6] も提案されているが、この NIC では、学習に伴うゆらぎが評価されていない。

上述の情報量基準 AIC、NIC は、統計的ゆらぎを評価するために導かれた。また、 $IC(K, N)$ は学習に伴うゆらぎを評価するために提案された。これら 3 つの情報量基準は 2 つのゆらぎのうちの一つを評価しているが、2 つを同時に評価していない。そこで、ここでは文献 [7][8] で提案されている情報量基準 NNIC を採用することとする。この NNIC は、統計的ゆらぎと学習に伴うゆらぎの両方を考慮に入れ、評価する情報量基準である (この NNIC の導出についての詳細は、文献 [7] と [8] を参照のこと)。ここでは NNIC の形式について簡単に述べる。

NNIC は次のように表現される。

$$\begin{aligned} NNIC &= S(q_N^*, \tilde{w}_N) \\ &+ (t-1)S(q_N^*, \tilde{w}_N) \\ &+ (1-t)\langle S(q_N^*, \tilde{w}_l) \rangle \\ &+ \frac{1}{N} \text{tr} V^*(\tilde{w}_N) H^*(\tilde{w}_N)^{-1} \end{aligned} \quad (1)$$

ここで、

$$S(q_N^*, \tilde{w}_N) = \frac{1}{N} \sum_{j=1}^N s(x_j, y_j; \tilde{w}_N)$$

$$t = \frac{\langle S(q_N^*, \tilde{w}_m) \rangle - \langle S(q_N^*, \tilde{w}_l) \rangle}{\langle S(q_m^*, \tilde{w}_m) \rangle - \langle S(q_m^*, \tilde{w}_l) \rangle}$$

であり、 \tilde{w}_m は各々 N 個の学習データから任意に m 個のデータを抽出して作られた学習データに基づき、求められた推定パラメータである。 $s(x_j, y_j; \tilde{w}_N)$ は推定パラメータ \tilde{w}_N でのニューラルネットワークのエネルギー関数であり、通常はニューラルネットワークの出

表4: 最適なニューラルネットワーク構成での予測誤差

	RMSE(MW)	MAPE(%)
1988 年度夏季	507.0	2.33
1988 年度冬季	278.0	1.48
1989 年度夏季	564.4	2.41
1989 年度冬季	344.0	1.72

力値と学習データの出力値との差の自乗を用いる。また、 $\langle \cdot \rangle$ は期待値を表す。 $V^*(\tilde{w}_N)$ と $H^*(\tilde{w}_N)$ は

$$V^*(\tilde{w}_N) = \frac{1}{N} \sum_{j=1}^N \nabla s(x_j, y_j; \tilde{w}_N) \nabla s(x_j, y_j; \tilde{w}_N)^T$$

$$H^*(\tilde{w}_N) = \frac{1}{N} \sum_{j=1}^N \nabla \nabla s(x_j, y_j; \tilde{w}_N)$$

となる。 ∇ は gradient を意味する。

式(1)で表される NNIC は、その右辺第4項で統計的なゆらぎを評価し、右辺第2、3項で学習に伴うゆらぎを評価している。この NNIC の値を最小にするニューラルネットワークのパラメータ数を選定すれば、学習データに基づく最適なニューラルネットワーク構成が得られる。

5.3 シミュレーション

ニューラルネットワークの構造はその中間ユニット数で決まる。そこで、前述 NNIC を最小とする中間ユニット数を有するニューラルネットワークを最適なニューラルネットワーク構成として採用し、4章で述べた予測シミュレーションを行なった。

4章のシミュレーションでは中間ユニット数を夏季は3個に、冬季は4個に固定していたが、ここでは学習データが変わるごとに NNIC の値を最小とする中間ユニット数を選定した。その結果、中間ユニット数はその学習データによって、1988年度夏季については2個から6個まで変化し、1989年度夏季については2個から5個まで、1988年度冬季については2個から5個まで、1989年度冬季については2個から8個まで変化した。中間ユニット数の変化は、季節の変わり目などモデルの構造的変化が必要とされる場合に起こっていた。最適な中間ユニット数をその都度選定した場合の平均予測誤差を表4に示す。

表4を表2、3と比較してみると、表4のRMSEとMAPEが共に、表2、3の値よりも小さく、平均的に

予測精度の良いニューラルネットワークが選定されていることがわかる。

6. まとめ

ニューラルネットワークの有する学習能力による関数近似を適用した翌日最大電力需要想定問題について解説した。本論では気象データに基づき、ニューラルネットワークを利用した最大電力需要想定問題を検討し、従来の重回帰モデルと同程度以上の精度で予測が行なえることを示した。さらに、情報量基準を用いてニューラルネットワークを最適化し、経験的に構成されたネットワークより、安定した予測ができることを示した。

参考文献

- [1] 小野賢治, 所健一: “「通年モデル」による翌日最大電力予測”, 電力中央研究所報告(依頼報告), Y90502, (1990).
- [2] 船橋賢一: “ニューラルネットワークの capability について”, 電子情報通信学会技術研究報告, MBE88, No.52, 127-134, (1988).
- [3] Takashi Onoda: “Next day peak load forecasting using an artificial neural network with modified back propagation learning algorithm”, *Proc. ICNN'94*, 3766-3769, (1994).
- [4] 坂元慶行, 石黒真木夫, 北川源四郎: “情報量統計学”, 共立出版(1983).
- [5] 和田安弘, 川人光男: “新しい情報量基準と cross validation による汎化能力の推定”, 電子情報通信学会誌 D-II, J74-D-II, No.7, 955-965, (1991).
- [6] Noboru Murata, Shunji Yoshizawa, and Shun'ichi Amari: “Network information criterion - determining the number of hidden units for an artificial neural network model”, *IEEE Transactions on Neural Networks*, Vol.5, No.6, 865-872, (1994).
- [7] Takashi Onoda: “Neural network information criterion for the optimal number of hidden units”, *Proc. ICNN'95*, 275-280, (1995).
- [8] 小野田崇: “階層型ニューラルネットワークの情報量基準”, 人工知能学会誌, Vol.11, No.4, 574-584, (1996).