

ニューラルネットワークの工程時間見積もりへの適用

松成文夫

A Neural Network Applied to the Estimation of Processing Time

Fumio Matsunari

要 旨

CAD, CAMさらにCIMなど、工場における情報の電算化に多大の努力が払われている。CIM構築の一課題として、各種の作業を適切にスケジュールする問題があり、そこでは各作業に要する時間を知る必要がある。

本稿で対象としたのは、樹脂金型加工の工程時間の見積もりである。技術進歩の早い樹脂加工では、その時の技術に応じた多種多様な形状の金型が求められる。そのため、加工に要する時間の見積もりは熟練した工程設計者によってなされるが、その見積もり精度は十分でなく、見積もりの電算化も行えなかった。

我々は、この見積もりにニューラルネットワークの適用を検討した。見積もるべき工程時間は、実情に即したものにするために、実際の加工時間に加えて、加

工前の段取り時間および加工後の待ち時間をも含めた。見積もるための入力変数としては金型図面から読みとる穴の個数、長さなどの各種の設計諸元および工程設計時に決められる加工法および加工機種別とした。

従来、ニューラルネットワークの適用にあたっては、ネットワークの中身がわからない、また、場合によっては非常に飛びはなれた値を示すことなどが指摘され、十分に信頼して使用できる状況になかった。本稿では、この問題を解決するために(1)ネットワークの学習の際に用いられる評価基準に減衰項を導入し、(2)ネットワークの構成に選択層を用いた方法を提案した。その結果、定性的にわかりやすく、定量的に十分精度のよい工程時間見積もり法を得ることができた。

Abstract

Scheduling many tasks properly is an important problem to construct CIM. In solving this problem, precise calculation of the time required for each task is essential.

In this article, we take up the problem of estimating the time required for processing resin molds. In this field, rapid technology progress necessitates processing molds by the state-of-the-art technology, resulting in so many kinds of molds needed. Therefore, the estimation of the processing time has been difficult for a computer system and had to be done by processing experts.

We applied a neural network to this estimation. We defined the processing time to include not only the time for processing molds with machines but also the time before and after processing. Also, we introduced 2 types of input

variables for the network; dimensional variables for mold design and processing variables.

In applying a neural network to the solution of such a problem, it has previously been criticized that the network looks like the black box and that the value estimated by the network is sometimes very different from the expected value. Therefore, the network has not been used with full confidence.

To solve this problem, we added the *weight-decay* term to the learning criterion and introduced a *selective layer* into the network configuration method. As a result, we obtained an estimation system whose qualitative meaning can be easily understood and whose quantitative precision satisfies the requirements.

キーワード

ニューラルネットワーク, 金型加工, 工程時間, 見積もり, 減衰項, 選択層, 結合係数分布

1. はじめに

各企業において、CAD、CAMシステムなど、工場における情報の電算化に多大の努力が払われている。これは、一つの製品単位での作業の効率化、高品質化に主眼があるものと考えられる。一方、複数の製品製作の効率化を意図して、工場の生産準備を統合的に行うためのCIMの構築も進められている。CIM構築の一つの技術課題として作業のスケジューリングの問題があり、作業の所要時間の見積もり方法が必要となる。

本稿で対象としたのは樹脂金型の製作の際の各工程の所要時間の見積もりである。樹脂金型加工においては、樹脂加工の技術進歩が早く、その時の技術に応じた各種の金型が設計される。さらに、樹脂加工では多様な製品の製作が可能であるために多様な形状の金型が必要となる。

そのため、金型製作にあたっていくつかの加工工程に分割した時、各工程で過去に行ったのと同じ工程がなく、その工程の所要時間を十分な精度で見積もるのは、熟練した工程設計者にとっても難しい作業であった。

我々は、この見積もり作業の電算化とそれともなう見積もり精度向上の手法として、ニューラルネットワークの適用を検討した。ニューラルネットワークの魅力として、本稿での見積もりのような曖昧な内容の問題にも適用されることとともに、その学習機能により、逐次、実績値を学習データとして用いれば、技術進歩にも対応できることなども考えられる。

一方、ニューラルネットワークの適用にあたっては、学習時に用いたデータについてはよい精度を示すが、その他のデータについての結果は保証されないという汎化能力の問題があり、さらにネットワークの内容もブラックボックスでわからないという問題とも重なって、実用的には十分に信頼して使える状況になかった。

このような状況の中で、工程時間見積もりへニューラルネットワークを適用する際の目標を次のように設定した。

- ネットワークの構成内容が定性的に理解でき、
- ネットワークの出力結果が定量的に満足できる。

構成内容が定性的に理解でき、概念的にイメージしている分析的なモデルに一致すれば汎化能力の問題は解決し、システムの性能として、定量的に満足できる結果が得られることができれば十分であろう。

以下においては、ニューラルネットワークの適用に際しての考え方を主体に述べた。最初、対象としている樹脂金型加工の工程時間について考察し、見積もりのための概念的なモデルについて整理した。次に、そのモデルをニューラルネットワークで表現し、上に述べた目標を達成するために用いた二つの手法についての考え方と効果について述べる。一つは学習の際の評価基準に導入した減衰項であり、一つはネットワークの構成法として用いた選択層についてである。

2. 工程時間の見積もり

2.1 見積もりの必要性

自動車の樹脂部品製作に用いられる金型は一品生産加工であり、次々に型を受注し、製作していく金型加工工場ではいろいろな情報の流れがあるであろう。Fig. 1は図面と納期をもとにした二つの情報の流れおよびニューラルネットワークによる見積もりの位置づけを示している。

一つは樹脂部品加工に用いられる金型の型設計からその加工を行うために必要な直接的な技術情報でありCAD、CAMに関する情報である。この流れでは、型の設計とそれにもとづく情報が工程設計、作業設計の段階を経て、時間とともに詳細な内容に変換され、細分化される。

もう一つは納期までに製作するために必要な生産準備に関する管理情報であり、多くの型を計画的に製作するのに必要な情報であり、CIMの情報といえるであろう。この流れでの情報処理的な内容としては加工法の決定や加工機および作業者の割り付けなどがあり、これも時間とともに詳細が決定されていく性格を持っている。

図の二つの流れは、相互に干渉しあっており、おおまかな対応関係もある。型設計と大日程計画によって型加工の方針、作業の手配関係が決まってくる。小日程計画で割り当てられた作業者がNCテープを作り、実際の加工を行う。中日程計画では工程設計の内容にもとづいて加工機を割り付け、その加工機を用いた作業設計を行うという対応になる。

本稿で対象としているのは、中日程計画の加工機の割り付けの際のデータとして必要な個々の加工機の占有時間の見積もりであり、これを工程設計時点での型図面情報をもとに推定しようとしている。

加工に要する時間の見積もりは、従来、熟練した工

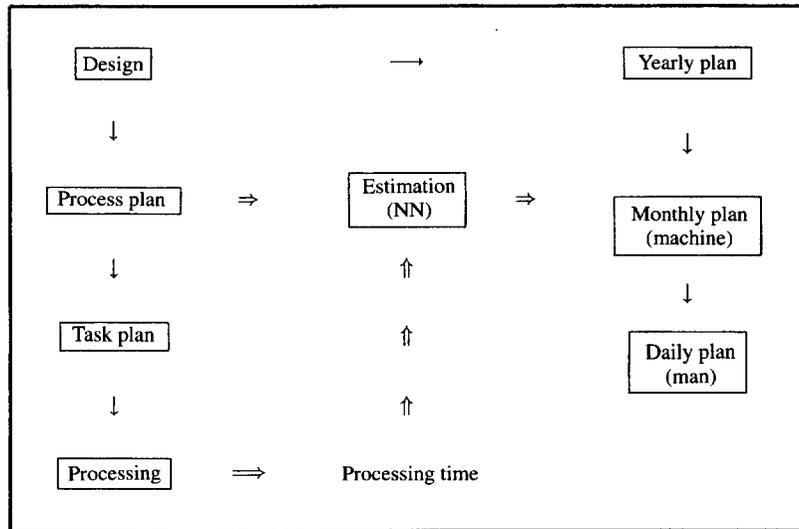


Fig. 1 The flow of information in a factory. Single-arrow (\rightarrow) represents the flow from a global upper task to a detailed lower task. Double-arrow (\Rightarrow) represents the input and/or output data of the estimated task when a neural network is applied to this system.

程設計者によって行われてきた。しかし、樹脂金型では多種多様な形状が要求され、必ずしも精度のよい見積もり結果は得られておらず、中日程計画に悪い影響を与えることも多かった。このような状況で精度のよい見積もり法が要望されており、今回、ニューラルネットワークの適用を検討した。ニューラルネットワークの学習データとしては、実際の加工工程のデータを収集することによってなされる。

なお、Fig. 1に示す工程設計と中日程計画の業務が厳密に二つに分かれているわけではなく、実際の利用にあたっては工程設計者が見積もりをもとに工程の組み直しをするなどの運用になると思われる。さらに、Fig. 1の見積もりの電算化が可能になることによって、それを組み込んだ運用は種々の形態が考えられる。しかし、本報告はあくまでもシステム化に必要な要素技術としてのニューラルネットワークの手法の検討であり、システム自体の検討ではない。

2.2 見積もりのモデル化

金型製作にあたっては、その加工をいくつかの工程に分けて製作するが、見積もりの対象とする工程時間は、その際に各工程で必要とされる時間である。

本報告では工程時間を加工機単位とし、以下のよう

に定義した。

工程時間

$$= \sum_n (\text{段取り時間} + \text{実加工時間} + \text{待ち時間})_n \quad (1)$$

n : 加工部位番号

つまり、金型製作工程を加工機単位の工程に分割し、その加工機で連続して、いくつかの必要な部位を加工する単位を工程と呼び、その工程に必要な時間を工程時間とした。この工程時間には実際の加工時間だけでなく、各部位について加工の前に必要となる段取り時間および加工後に必要となる次の加工までの待ち時間が含まれている。

工程時間を左右する因子は現場の意見を参考に選んだ。このような場合、往々にして、ミクロな現象が強調される場合がある。しかし、あまり詳細に、分析的立場で考えるのはニューラルネットワークの適用を難しくする。具体的な加工方式よりもどのような因子が効くかというのが大事であり、できるだけ多くの因子を入れることとした。もし、不要な因子なら、ネットワークを考察することによって除去できることを期待した。

工程時間見積もりの概念的なモデルとしては、基本的には加工容積で実加工時間が決まり、加工部位個数

で段取り時間、待ち時間が決まると考えられる。具体的には、工程設計の段階で設計図面から読みとれる因子との関係は次のように考えた。

段取り時間 加工部位の設定ごとに必要になる時間であり、加工部位の個数に段取り作業の難易度を係数としてかけたものである。設計図面から同一種類の加工部位の個数を読みとり、入力とする。

実加工時間 実際に加工機が稼働している時間であり、加工部位ごとにその寸法に加工作業の難易度を係数としてかけたものである。設計図面から同一種類の加工部位の長さまたは面積などを読みとり、入力とする。

待ち時間 作業者および加工機の稼働状況に影響される時間であり、設計図面から読みとれる因子はない。しかし、段取り時間の入力因子で代用される。

待ち時間のように、対応する入力因子が不明確なまま、見積もる工程時間にその時間を含めることは結果としての見積もりシステムの信頼性を落とすであろう。しかし、学習データの収集の点で待ち時間を除去するのは難しかった。また、こうすることによってスケジュールの際に直接割り付ける値と直接対応がとれることになる。さらに、楽観的に見ると、精度のよい見積もりシステムができて、的確なスケジュールができ、平準化された稼働が実現できれば待ち時間は段取り時間と同一視しても問題ないと考えた。

設計図面から読みとれる因子を設計変数と呼ぶことにする。例えば、加工穴の長さまたは加工穴の個数などをいう。ニューラルネットワークを用いたモデル化においても、どの程度、詳細に設計変数を選ぶかが問題になるが、基本的にはできるだけ詳細に選ぶことにした。

工程時間は各設計変数に作業の難易度をかけたものであるから、概念的には次式で表される。

$$T = \sum_i W_i X_i \quad (2)$$

T : 工程時間

X : 設計変数

W : 難易度に起因する係数

i : 設計変数の違いを示す添字

基本的には式(2)が今までに述べてきた工程時間の見積もり式である。しかし、加工の難易度に起因する係

数 W は加工法または加工機の違いによって異なってくる。これを加工変数と呼ぶが、通常は、加工変数ごとに係数 W を補正する必要がある。これから工程時間は次のように示される。

$$T_j = \sum_i \eta_{ji} W_i X_i \quad (3)$$

T_j : 加工変数ごとの工程時間

η_{ji} : 加工変数ごとの補正係数

式(3)が本稿で扱う工程時間の概念的な見積もりのモデル式である。しかし、式(3)のままの線形式では十分な精度の見積もり結果が期待できない。そこで、精度の高い、非線形の近似式の生成手法と考えられるニューラルネットワークを適用し、精度の高い見積もり式を得ることが本稿の主題である。

3. ニューラルネットワークの適用

3.1 適用の際の課題と解決策

ニューラルネットワークの適用を考えた場合に、従来から問題とされていたのは学習データについてはよい結果を示すが、学習に用いなかったデータではよい結果を示さないことがあるという汎化能力の問題である。

今回の工程時間見積もりのデータを用いて汎化能力を示さなかった例を示す。Fig. 2は学習の進行にともなう実績工程時間と見積もり時間の誤差である。ここでは、収集されたデータを2種類に分けた。一つは学習に用いたデータであり、図では白丸で示してある。一つはネットワークの性能を見るために、学習に用いなかったデータであり、図では黒丸で示してある。図で見ると、白丸の学習データでは学習とともに誤差は減少していき、よい学習をしているように思われるが、学習に用いなかった黒丸のデータでは8万回程度の学習で誤差が最小となり、それ以上学習すると誤差が増加していく。この現象は過学習と呼ばれる。普通概念では学習をすればするほど、ネットワークの性能はよくなると思われるが、それは学習データに限定した場合であり、学習に用いなかったデータではそうはなっていない。次に、100万回学習した後での実績工程時間と見積もり時間の比較をFig. 3に示している。白丸で示した学習データはほぼ45°の線上に乗っているが、黒丸のデータで検証すると45°の線上から大きく外れているが目だってくる。このような状況が汎化能力がないといわれる。なお、図のデータは[0, 1]の値に正規化して示している。

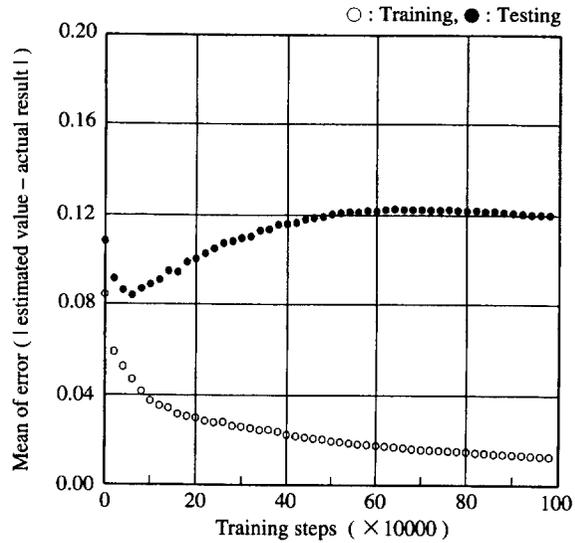


Fig. 2 An example of training curve which shows *over-learning*. The error for the training data (○) decreases as the training proceeds. On the other hand, the error for the testing data (●), which is used to confirm the validity, decreases only at the beginning of the training and increases as the training proceeds. The processing time is normalized to [0, 1], and the same normalization is done in the following figures.

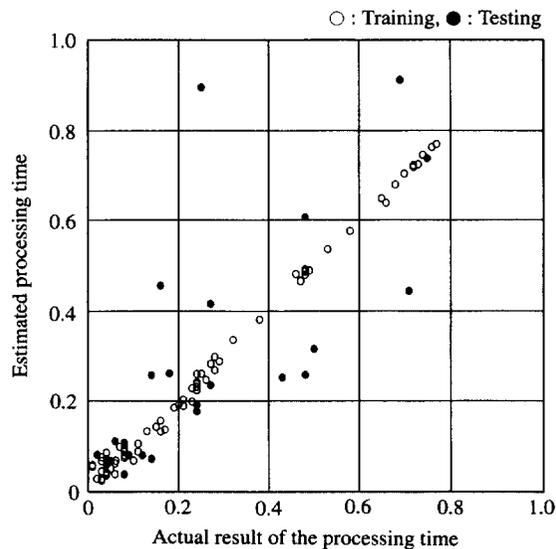


Fig. 3 An example of the effect of *over-learning*. After a million training steps in Fig. 2, the estimated values against actual results lie on a 45° line for the training data (○), but for the testing data (●), no good correlation is obtained.

データをもとに学習によって生成されたニューラルネットワークのモデルが良好な汎化能力を発揮しないのは、端的に言えば、対象の適切なモデルになっていないからである。適切なモデルが得られない理由として考えられるのは次の2点であろう。

- 適切な入出力関係を表すように、学習データが選ばれていない。
- 学習アルゴリズムが不適切なために、定量的に適切なモデルが生成されていない。

最初の問題点について考える。今回の工程時間見積もりの場合の概念的なモデルは式(3)で示されており、式(3)にしたがって入出力因子を選定し、結果として式(3)相当のモデルができれば、定性的には対象のモデルが生成されたといえ、汎化能力についても問題はなくなるであろう。ただ、この場合にも具体的にどの程度まで詳細に入出力因子として取り入れればよいか、学習データをいくつぐらいとればよいかなどというような定量的な問題があるが、これは最適なネットワーク生成の問題であり、本稿では対象外としている。

つぎに学習アルゴリズムについて少し復習しておく。通常、ネットワークはFig. 4の3層階層型の構成で使われており、記号で書くと次のような入出力関係である。

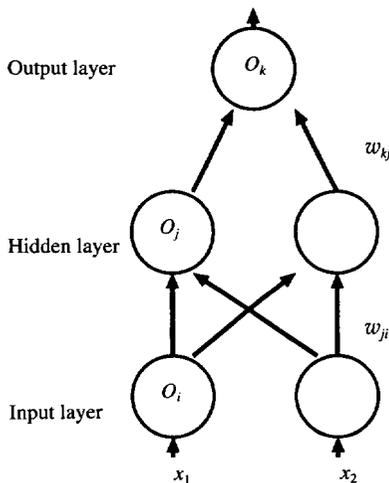


Fig. 4 Neural network with 3 layers. This type of network, which has Input layer, Hidden layer and Output layer, is popularly used.

$$O_k = f(\sum_j w_{kj} O_j) \quad (4)$$

$$O_j = f(\sum_i w_{ji} O_i) \quad (5)$$

O_k : 出力層素子の出力値

O_j : 中間層素子の出力値

O_i : 入力層素子の出力値, 通常は入力値

f : シグモイド関数

w_{kj} : 出力層素子, 中間層素子間の結合係数

w_{ji} : 中間層素子, 入力層素子間の結合係数

各入力層素子の出力値に係数倍した積和が中間層素子の入力となる。同様に、各中間層素子の出力値に係数倍した積和が出力層素子の入力となる。この時利用される係数 w_{kj} , w_{ji} は素子間の結合の強さを表す重み係数であり、学習によって、逐次、適正な値に修正される。本稿ではこれを結合係数と呼んでいる。各積和入力値は出力層および中間層ではシグモイド変換されて出力値となる。シグモイド関数としては出力値が $[0, 1]$ の値をとる $1/(1+\exp(-x))$, または、出力値が $[-1, 1]$ の値をとる $(1-\exp(-x))/(1+\exp(-x))$ がよく用いられている。なお、通常、入力層ではシグモイド変換をしないで入力値をそのまま出力値として使っている。

ニューラルネットワークの学習は次式で示すようにネットワークの希望出力値と実際の出力値の差を小さくするように結合係数を変更することによって行われる。

評価基準として用いられるのは次式の誤差の自乗和である。

$$E = \frac{1}{2} \sum_c \sum_k (D_{c,k} - O_{c,k})^2 \quad (6)$$

E : 誤差の自乗和

D : ネットワークの希望出力値

O : ネットワークの出力値

c : データ番号

k : 出力素子番号

評価基準 E の最小値を求めるのに勾配法が使われる。勾配法のもっとも簡単な方法は、 $\partial E / \partial w$ に比例させて結合係数を変化させることである¹⁾。

$$\Delta w = -\epsilon \partial E / \partial w \quad (7)$$

学習アルゴリズムの不備と考えられるのは式(6)で示される評価基準のとり方である。工程時間見積もりの場合、学習データとしての希望出力値は工程時間の実績値を用いている。しかし、この実績値には待ち時間も含まれており、かなりノイズの多い値である。出

力値をこの値にあまりにも近付ける努力は、当然、学習に用いていないデータの場合に悪い影響を与えるであろう。ある程度の誤差を許容して、できるだけ式(3)のイメージに近いネットワークを生成するのがよいと思われる。

この問題の解決策として、式(6)に減衰項を導入した。その結果、汎化能力も向上し、結合係数の内容もかなり理解でき、加工変数の入らない式(2)段階のモデルとしては十分と考えられた。しかし、式(3)のモデルとしては不十分であった。そこで、入力変数を設計変数と加工変数とで切りはなし、設計変数を通常の入力変数とし、加工変数はネットワークの記憶内容の切り替え信号とみなして、設計変数とは別入力とするネットワーク構成をとった。この加工変数入力のために設けた層を選択層と呼ぶ。

3.2 減衰項の導入とその効果

ニューラルネットワークの学習について、甘利は学習方程式および学習ポテンシャルを用いて理論化している²⁾。簡単に紹介すると、神経細胞の学習方程式は次の微分方程式で表される。

$$\tau \dot{\mathbf{w}} = -\mathbf{w} + cr(t)x(t) \quad (8)$$

\mathbf{w} : シナプス荷重(結合係数)

x : 入力

τ : 学習の時定数

c : 学習の効率を表す定数

r : 学習信号

学習信号は入力 x 、荷重 \mathbf{w} および教師信号 y に依存して決まる信号である。次に、 x 、 \mathbf{w} 、 y の関数 $R(x, \mathbf{w}, y)$ が存在して、これを \mathbf{w} で微分すると、学習方程式の右辺

$$\partial R(x, \mathbf{w}, y) / \partial \mathbf{w} = \mathbf{w} - cr(x, \mathbf{w}, y)x \quad (9)$$

が得られる場合に、ポテンシャル関数を持つ学習といい、 R を学習のポテンシャル関数という。

形式的に式(8)、(9)と式(6)、(7)を比較すると

$$R = \mathbf{w}^2 + E \quad (10)$$

とみることによって、ポテンシャル関数を持った学習と考えることができる。

工程時間見積もりに適用するために式(6)の評価基準を次のように書き直した。

$$E = \frac{1}{2} \sum_c \sum_k (D_{c,k} - O_{c,k})^2 + \lambda \mathbf{w}^2 \quad (11)$$

λ : 減衰率

ここで示した結合係数の自乗和の項は式(8)の微分方程式から減衰項、または学習に関連づけて忘却項と呼ばれる。

式(11)の評価基準を最小にするという観点からは、結合係数の項は罰金項に相当し、結合係数の分散が小さくなり、生成されるネットワークが、あまり変な値をとらないことが期待され、結合係数がわかりやすくなることが期待される。このことは、当時からHinton³⁾も述べていたが、実用的な問題での有効性の評価はあまりはっきりしていなかった。

式(11)を評価基準として工程時間見積もりに適用した。この見積もりに用いた入力変数は正規化して $[0, 1]$ で表される加工穴の長さ、加工穴の個数などの設計変数が14種類であり、さらに、 $\{0, 1\}$ で表される2種類の金型部品の別および14種類の加工機の区別を入力変数とした。以下では、金型部品の種類をも含め、 $\{0, 1\}$ の値をとる変数を加工変数と呼ぶ。ネットワークはFig. 4に示す3層の構成であり、中間層の個数は10個とした。Fig. 5は学習にともなう誤差の収束傾向を示しており、Fig. 6は100万回の学習を終了した後の工程時間実績値と見積もり時間の比較を示している。減衰項の考慮のないFig. 2に見られた過学習の現象はFig. 5では見られず、10万回以上学習しても学習用のデータ、検証用のデータともほぼ一定値を示している。Fig. 6ではFig. 3の場合と比べて、学習データについて実績値と見積もり値の誤差は大きくなっているが、検証用のデータについての誤差は小さくなっており、飛びはずれた値を示すのもなくなっている。なお減衰率の決定はWeigend⁴⁾のアルゴリズムにしたがった。

入力変数のうちで、加工穴の長さ、加工穴の個数などの設計変数に関係している部分のみをとりだした結合係数分布をFig. 7に示す。四角形の大きさが係数の絶対値を示しており、黒塗りが正の値で白抜きが負の値を表している。分布図の下段に入力層と中間層間の結合係数(w_{ji})を示し、上段に中間層と出力層間の結合係数(w_{kj})を示している。なお、出力層素子は一($k=1$)であるが、形式的に w_{kj} で表している。

いま、出力層、中間層素子の入出力変換関数を線形と考えると、 j 番目の中間層素子を通る、 i 番目の入力変数の出力への影響は $w_{kj} \cdot w_{ji}$ で表されると考えられ

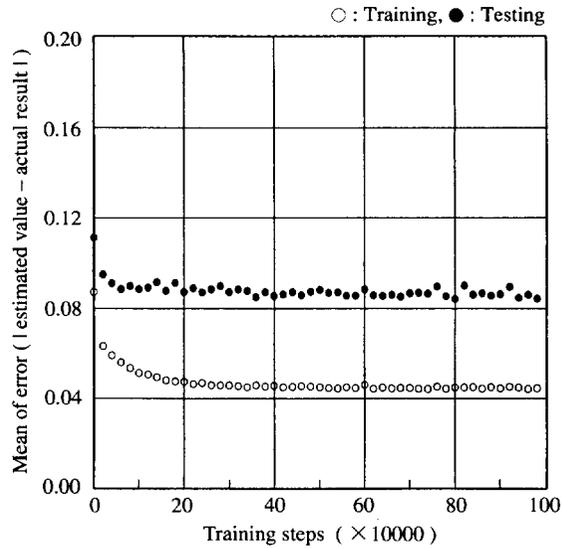


Fig. 5 An example of training curve when the *weight decay* term is applied to the learning criterion of the network. The error becomes constant in the training process after about 200 thousand steps not only for the training data () but also for the testing data ().

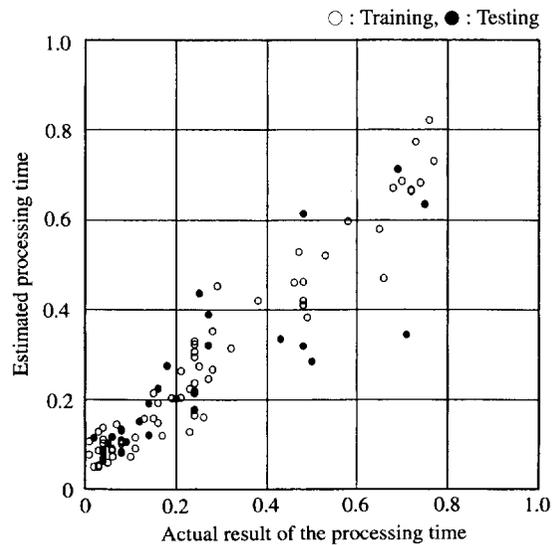


Fig. 6 Estimated values against actual results after a million training steps in Fig. 5. As the effect of using the *weight decay* term the correlation between the two becomes larger for the testing data (), while the correlation becomes smaller for the training data ().

る。Fig. 7ではこの値がほとんど正値を示している。つまり、 j 番目の中間層素子と出力層素子間の結合係数 w_{kj} の正負と同じ極性を j 番目の中間層素子と入力層間の結合係数 w_{ji} は示している。これから、ある程度の定量性を持って、設計変数の工程時間に影響する度合がわかり、入力変数の選定に役立たせることができる。なお、 $\{0, 1\}$ で表される加工変数の場合にはこのような読み方はできず、加工変数に関する結合係数は図に示していない。

減衰項を用いた効果を整理すると次のようになる。

- 検証データにおいて、見積もり時間の実績工程時間からのずれが小さくなる。
- 設計変数の寄与度合が結合係数分布からある程度読めるようになる。

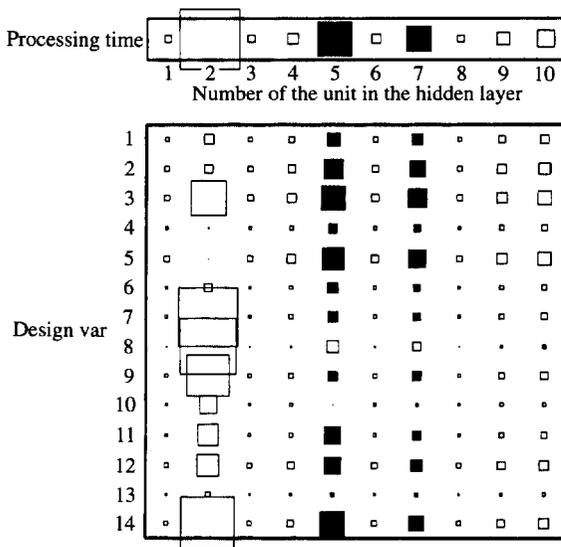


Fig. 7 The distribution of weight coefficients. Black boxes represent positive values while white boxes represent negative values. The size of the boxes represents the absolute value of the coefficients. The distribution becomes so clear by the effect of the weight decay term that the contribution of the input design variable to the output processing time can be presumed from the corresponding value in the lower block multiplied by the corresponding value in the upper block.

3.3 選択層の導入とその効果

通常のニューラルネットワークにおいてのニューロン素子の入出力関係は式(4), (5)に示したように外部素子の出力値と結合係数の積和を入力としシグモイド関数によって変換して出力値を得る。Rumelhartはもっと一般化した素子として、次式のシグマ・パイ素子を示している。

$$O_j = f_j (\sum_i w_{ji} \Pi_k O_{ik}) \quad (12)$$

これは、仮想的な O_i 素子を k 個の O_{ik} 素子の積とみなすことによって通常の積和入力の素子と見ることがができる。この素子の使い方は、例えば $k=2$ として、一方を $\{0, 1\}$ の値にすれば j 素子への i 素子の入力のon-off信号として使える。

工程時間見積もりの場合にこの機能を用いるためにFig. 8に示すように、通常の間層層の上部に切り替え

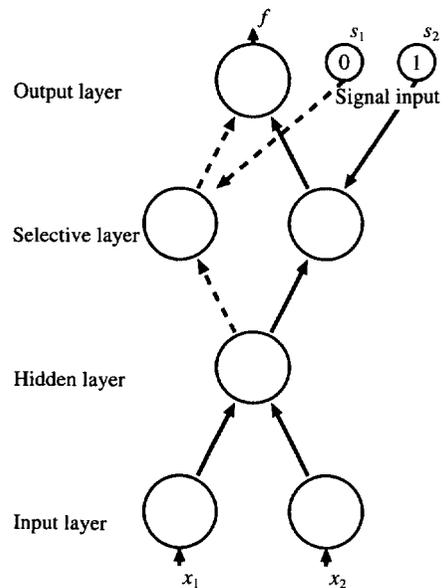


Fig. 8 A network with a *selective layer*. A layer above the hidden layer is added and it receives its input from two sources. One is the hidden layer and the other is the signal input. The summed value of the hidden layer units is multiplied by the signal and becomes the input to the *selective layer* unit. The value of the signal is either 0 or 1, so we can select the path of the flow in the network by controlling the signal. The figure shows the case when the right side path is selected.

のための層を付加し、選択層と呼んだ。図で設計変数 (x_1, x_2) が入力層への入力であり、加工変数 (s_1, s_2) が選択層への出力切り替え指示入力である。図は、 $s_1 = 0, s_2 = 1$ の場合を示し、加工変数 s_2 に対応する経路でネットワークの学習および見積もりの演算が行われる。

選択層は前節の減衰項の導入だけでは加工変数の寄与度とはっきりしなかったのを改良する意図であり、工程時間見積もりのモデルでいえば、式(3)の構造をもっと明確に表現するためである。選択層を用いなくて、加工変数の組合せごとに設計変数を入力としたネットワークを生成するという事も考えられるが、選択層を用いた場合にはFig. 8に示す入力層と中間層の間の結合係数は、加工変数の値に関係なく、常に学習されており、学習データの使用効率がよいと考えられる。

工程見積もりの実際のデータを適用した例として、選択層を用いて3層目で部品金型の種別を指示し、4層目で加工機の種類を指示した5層のネットワークの結合係数分布をFig. 9に示した。なお、中間層の素

子数は10個とした。また、シグモイド関数は出力値が $[-1, 1]$ の関数を用いた。図の結合係数は4つのブロックに分かれているが、それぞれ、左下から時計回りの順で、入力層、中間層、型部品選択層、加工機選択層および出力層の各層間の結合係数を示している。なお、結合係数の学習の初期値を正值にすることによって、結果としての結合係数も正值の結果を得ることができた。Fig. 7に比べて、Fig. 9が複雑であり、読みにくい部分もあるが、次のようなことがもっとはっきりしてきている。なお、本報はニューラルネットワークの適用法に関する報告であり、各設計変数、加工変数の意味づけについては、特に述べない。

- 結合係数が正值であり、入力値が増大すると出力値も増大する関係を示している。
- 工程時間に大きく効く設計変数は、7, 12, 14などである。
- 設計変数7は部品型Aに効き、設計変数12は部品型Bに効く。

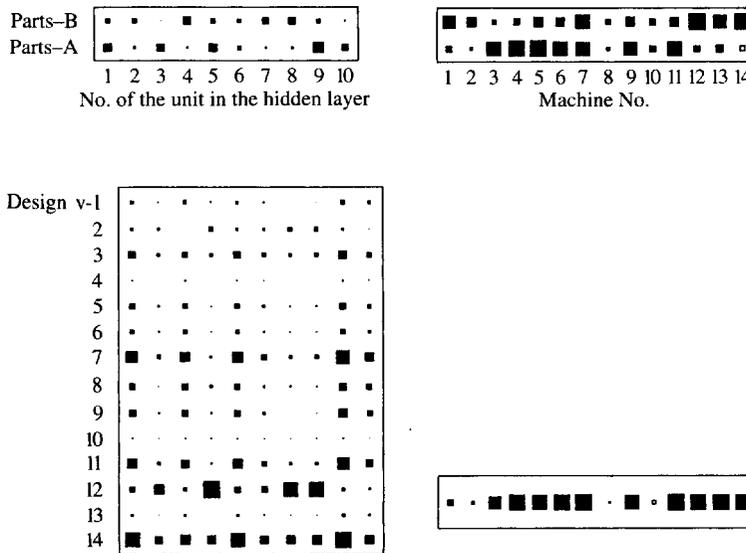


Fig. 9 The distribution of weight coefficients for selective layers. Two selective layers are added; one is a parts-selective layer and the other is a machine-selective layer. The distribution is composed of 4 blocks. The input starts from the left-lower block and the calculation proceeds clockwise to the right-lower block. From this distribution, we can obtain much information about the contribution of not only the input design variables but also parts and machine variables to the output processing time.

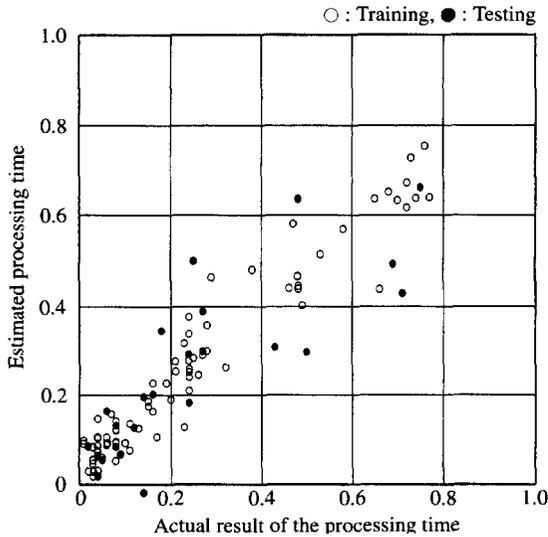


Fig. 10 Estimated values against actual results for selective layers, after a million training steps. The degree of the correlation between the estimated values and the actual results is nearly the same as that shown in Fig. 6.

- 加工機の違いはあまり効いていない。
- 加工機番号の8, 10のデータ個数が少ない。

このネットワークを用いた実績加工時間と見積もり時間の比較をFig. 10に示している。Fig. 6と比較してみて、同程度の精度といえる。今後、全体的な精度向上のためには、学習データの精度向上が必要であろう。本節の結果、選択層を考慮することによって加工変数についても、定性的な合理性程度は読みとれるようになった。

4. まとめ

実績値にノイズの多い工程時間の見積りにニューラルネットワークを適用した。適用にあたって、汎化能力の問題を解決するために(1)ネットワークが定性的に理解できること、(2)出力結果が定量的に満足できることの2点を目標にした。

解決策として導入した減衰項と選択層により、結合係数が定性的にでも理解できるようになり、また過学習の現象も押えることができた。結合係数の理解により、生成されたネットワークが合理的なものであることが確認

でき、また過学習を押えることによって検証データについても十分な精度の見積もりが可能になった。なお、本稿のデータはノイズ成分が大きく、誤差の評価をきちんと行っていないが、別のデータでは選択層を用いた場合の方が精度がよいことも示されている⁵⁾。

今後の問題として、ネットワークの最適な構成の決め方がある。一つは中間層素子の個数をいくらにすればよいかという問題であり、もう一つは選択層の構成をどうすればよいかという問題である。前者の問題については、現在、減衰項の導入はかなり一般的になっており³⁾、さらにモデル選択の立場からもっと精緻な考えの罰金項も提案されている^{4, 6)}。しかし、問題依存の部分が大きく一般論は難しいようである⁷⁾。選択層の構成法については、本稿では型部品の層に加工機選択をおいた。しかし、この順序を逆にするとよい結果を示さなかった。また、加工機による違いがあまりない場合には、同種の加工機と見た方がよいであろう。このようなことを考慮した自動的なネットワーク生成法が必要となると思われる。

最後に、本稿のデータを提供していただいたトヨタ自動車ダイエンジニアリング部の方々に感謝致します。また当所材料1部の大庫主任研究員にはデータの収集、分析の上で多大なるご協力をいただきました。

参 考 文 献

- 1) Rumelhart, D. E., Hinton, G. E. and McClelland, J. L. (eds.): Parallel Distributed Processing; Explorations in the Microstructures of Cognition, 1(1986), MIT Press Cambridge
- 2) 甘利俊一: 神経回路網の数理, (1978), 産業図書
- 3) Hinton, G. E.: "Connectionist Learning Procedures", Artificial Intelligence, 40(1989), 185-234
- 4) Weigend, A. S., Rumelhart, D. E. and Huberman, B. A.: "Generalization by Weight-Elimination with Application to Forecasting", Advances in Neural Information Processing Systems, 3(1991), 875, Morgan Kaufmann
- 5) 松成文夫: "選択層を用いた2値数とアナログ数の分離入力", 神経回路学会第4回全国大会講演論文集, (1993), 314-315
- 6) Nowlan, S. J. and Hinton, G. E.: "Simplifying Neural Networks by Soft Weight-Sharing", Neural Computation, 4(1992), 473-493
- 7) Finnoff, W., Hergert, F. and Zimmermann, H. G.: "Improving Generalization Performance by Nonconvergent Model Selection Method", Neural Networks, 6-6(1993), 771-783

著 者 紹 介



松成文夫 Fumio Matsunari
生年：1939年。
所属：開発部。
分野：ニューラルネットの研究。
学会等：日本機械学会，自動車技術会，
神経回路学会会員。