

情報処理技術の応用に関する調査研究

地域経済予測サポートシステムに関する 調査研究報告書

平成 5 年 3 月



財団法人 日本情報処理開発協会

財団法人 関西情報センター

この報告書は、日本自転車振興会から競輪収益の一部である
機械工業振興資金の補助を受けて平成4年度に実施した「情報
処理技術の応用に関する調査研究」の成果をとりまとめたもの
です。

はじめに

高度情報化社会の円滑な実現のためには、地域間の情報化格差を是正しつつ、全国的にバランスのとれた情報化を図ることが不可欠となっている。そのため、各所において情報基盤の整備、情報利用の高度化が促進されている。

財団法人日本情報処理開発協会は、昭和54年度から昭和58年度の5ヶ年にわたり、地域情報拠点の育成事業を実施し、地域データの整備、クリアリング機能の形成、データ交換利用を中心に地域情報化の基盤整備を推進してきた。昭和59年度から昭和61年度の3ヶ年にわたり、地域活性化と情報の有効利用を目的に、地域内オンライン・ネットワークによる情報流通システムに関する調査研究事業を実施し、地域内情報流通システムのビジョンの作成を行った。また、昭和62年度から平成元年度の3ヶ年にわたり、地域の実態に即したかたちでニーズに対応した情報利用の高度化を図ることを目的に、「情報処理技術の応用に関する調査研究」を実施し、中部地域をモデルとした「オンラインデータベースの統合検索技術に関する調査研究」および九州地域をモデルとした「マルチメディアに対応した地域開発支援型情報システムに関する調査研究」を行った。

この間における情報化の進展は目ざましく、データベースの構築やサービス化の進展、ネットワークシステムの拡大と通信技術の向上、人工知能技術の実用化等の情報処理技術の進展は、産業社会から国民生活に至るまでの情報利用の裾野を拡大してきている。また、地域においても産業、社会、生活のあらゆる分野で情報化が進展してきている。しかしながら、各地域それぞれのニーズへの対応という点では、必ずしも十分な状況とはいえないことも事実である。

こうした状況を踏まえ、当協会では平成2年度から3ヶ年計画で地域ニーズに対応した情報利用の高度化を図ることを目的に「情報処理技術の応用に関する調査研究」事業を継続実施し、中部地域においては「オンラインネットワークによる地域産業政策コンサルティングシステムに関する調査研究」、近畿地域においては「地域経済予測サポートシステムに関する調査研究」、九州地域においては「複合通信システムによる中小企業リモート診断システムに関する調査研究」の三つをモデルケースとして個々のテーマに基づき調査研究を実施している。

本年度は、これら3地域における調査研究の最終年度にあたり、本報告書は、このうち、近畿地域の「地域経済予測サポートシステムに関する調査研究」について、財団法人関西情報センターに委託、実施した内容をまとめたものである。

最後に、本調査研究にあたって、ご指導、ご協力をいただいた委員をはじめ、関係各位に対し、深く感謝の意を表する次第です。

平成5年3月

地域経済予測サポートシステムに関する調査研究委員会委員名簿

<委員長>

中村 良平 岡山大学

経済学部 助教授

<委員会委員>

新井 京子	財団法人関西経済研究センター	研究部長
荒川 宜三	株式会社大和銀総合研究所	常務取締役
有海 三夫	住友電工ワークステーション株式会社	大阪支店長兼大阪営業部長
栗津 征二郎	株式会社京都新聞社	政経部部長代理
飯田 洋	株式会社大丸	経営計画室部長
大崎 幹雄	シャープ株式会社	情報技術開発センター副所長
神吉 正和	株式会社日本総合研究所	地域開発研究部次長
中村 信彦	株式会社さくら銀行	調査部長兼神戸調査部長
柴山 康生	南海道総合研究所	常務理事
鈴間 能成	大阪ガス株式会社	総合企画部部長補佐
滝川 孜郎	株式会社福井経済経営研究所	経済調査部次長
武田 壽夫	関西電力株式会社	企画部次長
津田 盛之	大阪府立産業開発研究所	産業経済部長
西上 和雄	財団法人南都経済センター	常務理事
松下 滋	株式会社三和総合研究所	参与
大西 康一	近畿通商産業局	総務企画部調査課長
石田 尚夫	近畿通商産業局	総務企画部情報管理室長
井上 道典	財団法人関西情報センター	常務理事
松本 和美	財団法人関西情報センター	情報企画部部長

<作業部会委員>

徳田 裕平	株式会社日本総合研究所	地域開発研究部副主任研究員
小柴 光永	財団法人南都経済センター	事務局次長
井上 隆	関西電力株式会社	企画部副長
丹戸 裕之	大阪ガス株式会社	総合企画部
松尾 泰秀	株式会社大丸	経営計画室 専任課長
空閑 茂起	シャープ株式会社	情報技術開発センター第2開発室長
深瀬 澄	大阪府立産業開発研究所	産業経済部研究員
武田 至弘	近畿通商産業局	総務企画部情報管理室
広井 健一	近畿通商産業局	総務企画部情報管理室
福井 健二	近畿通商産業局	総務企画部情報管理室
坂本慎一郎	近畿通商産業局	総務企画部調査課
桑原 昭翁	近畿通商産業局	総務企画部調査課
橋本 恵子	財団法人関西情報センター	情報企画部主任研究員

(順不同 敬称略 役職は委員当時)

目 次

1. 調査の概要	1
2. 非線型テクノロジーの発展と応用	5
(1) はじめに	5
(2) 非線型の現象と理論	6
(3) 非線型系による情報処理の例	13
(4) まとめ	14
3. ニューラルネットワークによる経済予測	17
(1) 現在までの経緯	17
(2) ニューラルネットワークの基本仕様について	17
(3) ニューラルネットの学習データについて	17
(4) 予測手法について	55
(5) ニューラルネットを用いた経済予測手法のまとめ	66
4. コネクショニズムの経済学——ニューラルネットワークが示唆する新しい経済学—— ...	67
(1) はじめに	67
(2) 均衡の可能性と「直感する経済人」	68
(3) 逐次記号処理と並列分散処理	70
(4) 経済構造の自己組織化	72
(5) 新しい経済学へ	75
5. 時系列データと経済予測	77
(1) はじめに	77
(2) 時系列データの特徴	77
(3) 伝統的な時系列分析	80
(4) 最近の時系列分析—VAR、単位根、共和分—	86
(5) おわりに	93

1. 調査の概要

1. 調査の概要

過去2年間の調査において、ニューラルネットの経済予測への適用に一定の成果を得ることができた。

今年度はこの成果を、より確かなものとするため、適用方法に幾通りかのケースを設定し最適な方法を探るとともに、本調査の理論背景を整理し、その中での本調査の位置づけを試みた。

はじめの事項については、ニューラルネットによる経済予測には試行錯誤が避けられないが、今年度はハード的に数値演算コプロセッサを付加することができたため、演算速度の驚異的な高速化（約60倍）に助けられ、余裕のある学習ケースの設定の中から、適切な方法を探ることができた。

すなわち、昨年度の年次データに対し、今年度は四半期データを用い、学習の時期と期間、中間層のニューロン数、予測のための変数設定方法と期間などのケース設定によるシミュレーション、さらに公定歩合の操作を例にニューラルネットによる予測の、経済理論との整合性の確認などを行った。

2番目の事項については、今回のニューラルネットによる予測が、通常経済予測に用いられる最小自乗法による推計と大きく異なる点の1つは、非線型理論の応用であるということである。

線型に比べて、非線型現象の研究への取り組みは、数学的取扱いが難しいこともあって、これまで遅れていた。しかし近年、コンピュータを数値解析に活用することにより、ニューロやカオスなど非線型理論を応用したテクノロジーの発展がみられ、その適用分野も、力学や電気工学など技術面ばかりでなく、金融・経済、娯楽・芸術などの分野へも拡大してきている。今回の調査も、大きくはこの潮流の中に含まれるものである。

一方、計量経済学においては、計量モデルの有効性をめぐる議論、マクロ計量モデルと時系列モデルの論争などの課題を抱えている。定量的に捉えられない定性的な要因の予測への反映などもこの調査の課題であった。このような課題に対して、数式でなくパターンで予測行動を捉えようとするニューラルネットの考え方は、計量経済学あるいは広く近代経済学の考え方に1つの示唆になりうるものだと考えられる。

今回の調査は、これら、非線型テクノロジーの発展と応用分野の広がり、計量経済学の課題という、2つの事項の接点に位置するものである、と考えられる。

また、今後の本調査の展開の方向を考えると、前年度や今年度のニューラルネットによる経済予測は、変数の選択などマクロ計量モデルとの比較を念頭に置いたものであった。しかしその考え方や手法は時系列モデルにより近いものだといえる。今後ニューラルネットによる予測の経済学的な意味づけを深めていく上でも、時系列モデルとの比較検討は今回の調査に残された課題の1つである。

以上のような考え方から、今年度の報告書は、

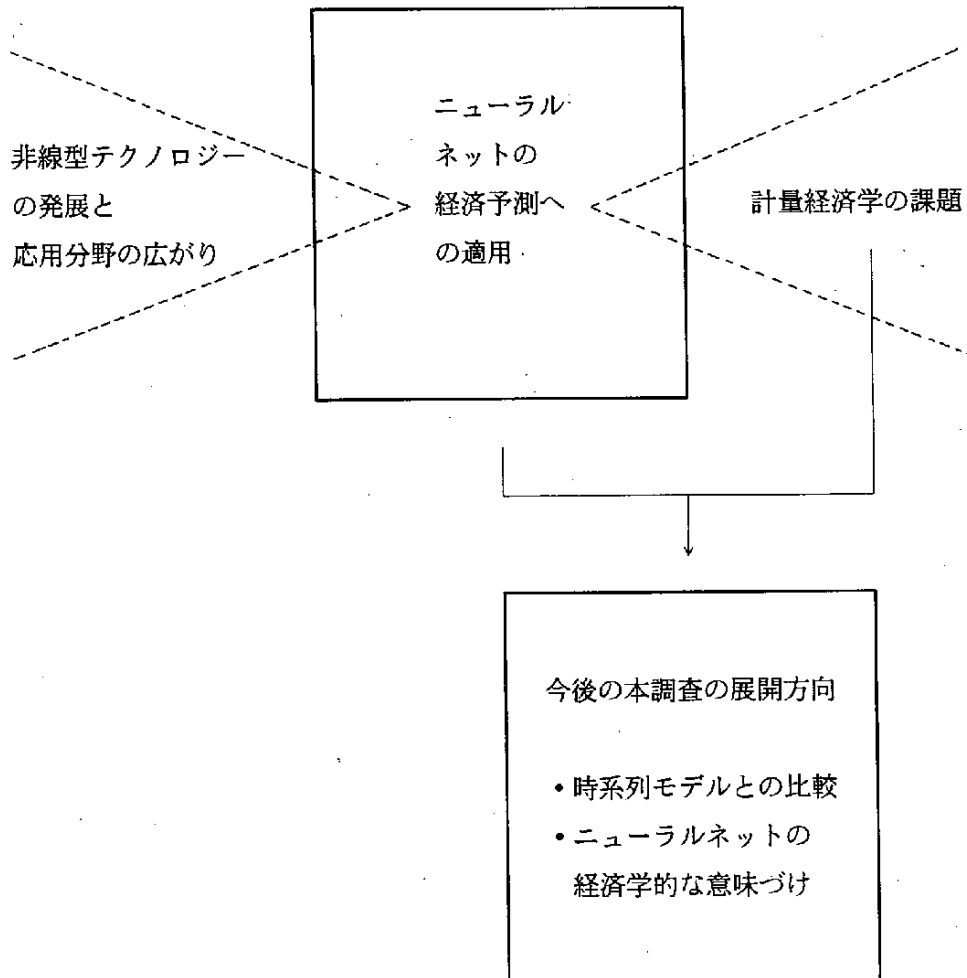
2章において、ニューラルネットなどの非線型テクノロジーの発展と応用について

3章において、ニューラルネットによる経済予測シミュレーション結果について

4章において、ニューラルネットが示唆する新しい経済学の考え方について、

5章において、時系列モデルの動向について本調査の今後の展開の方向の示唆を含め、それぞれ、とりまとめを行った。

これらの今年度の本調査の概要を整理すると、次の図の通りである。



以上のように、3カ年にわたる調査の結果、試行錯誤を交えながらではあったが、経済予測における1つの新しい考え方や手法を検討することができ、その残された課題も明きらかとなった。今後は、機会を捉えて、残された課題への取り組みを進め、本調査の成果の拡大深化を図っていく所存である。

なお本報告書は、主に、委員長をはじめとする本研究会メンバーで執筆したものであるが、2章と4章については、本研究会メンバーではないが特に専門の立場から、石井俊直氏（三菱電機株式会社中央研究所）、鷺田豊明氏（和歌山大学経済学部）に執筆協力をいただいた。

2. 非線型テクノロジーの発展と応用

2. 非線型テクノロジーの発展と応用

(1) はじめに

「非線型テクノロジー」という言葉はここでは、非線型理論を応用したテクノロジーという意味である。したがってこの言葉には非常に広範な内容が含まれ得るのであるが、この報告の目的はそれ全体を議論することではない。昨今のコンピュータの発展などに伴って「ニューロ」や「カオス」などのキーワードが登場しテクノロジーを一つの新しいパラダイムへ導くものとして期待されているが、それらにとって非線型系に特徴的な現象が非常に重要な役割を果たしている。そこで非線型現象やその理論を、情報処理技術への応用という観点から見た側面について議論してみる。

さて、「非線型テクノロジー」を大きく二つに分けてみることにする。

1. 対象とする問題に生じる非線型現象を“避ける”ためのテクノロジー
2. 対象とする問題に生じる非線型現象を“利用する”ためのテクノロジー

そもそも非線型に取り組むのは、テクノロジーが対象とする現実的な問題には広く非線型現象が見られることに起因している。ここに上げた“避ける”ためのテクノロジーは、強力で体系化された道具である線型理論を適用するためにいわば問題の非線型性を剥ぎ取るためのテクノロジーである。したがって現象を線型に近似するといったものはこちらに属すると考えて良い。また非線型制御などに於ける可制御性に関わる理論（テクノロジー）もこちらに分類される。カオスの研究に於いてもどのようにしてカオスの発生を抑えるかとか、カオティックな系での予測問題などもこちらのものである。一方2に含まれるテクノロジーを考えることは生体システムや社会・経済システム抜きには難しい。生体システムの神経系を例にとると、まずこのシステムは明らかに高度な情報処理を実現するものであって、しかも非線型の素子が多数結合してできている。このことから非線型システムによって実現される新しい情報処理原理が存在すると信じられている。2のテクノロジーはそのような原理から成るものであって、神経系に限らず免疫系、生態系、発生、進化などの研究を基として進められている。

二つに分けては見たが、これはテクノロジーの指向がどちらに向いているかによるものであって、とくに理論では不可分のものであることに留意していただきたい。（局所線型かによってシステムを近似する手法などは非線型システムの解析における常套手段である。）筆者は主にニューラルネットワークの研究に携わっていることもあり、また生物に代表される複雑なシステムに関する研究を概観することがこの報告に与えられた重要な目的と考え、2に話題を絞ることにする。

(2) 非線型の現象と理論

① パターン形成——非線形現象概観

さて「非線型テクノロジー」の議論に移る前に非線型系に特徴的な現象——パターン形成——について触れることにする。「非線型テクノロジー」を考える上でこれらは欠くことのできないバックグラウンドになるが、同時にパターン形成やその理論は、この若い分野における、「テクノロジー」の重要な部分でもある。

毎日、テレビの天気予報で見ることのできる衛星から取った地表の雲の写真には様々なパターンを見つけることができる。台風は大きな渦を形成しているし冬になると大陸からの季節風によって日本海上空に筋状の雲が現れる。冬から春にかけて季節風が吹くと九州の西に位置する済州島の風下に巨大な（～100km）渦の列が観測されることがある。これはカルマン渦列と呼ばれているものである。カルマン渦列は一定の向きに流れている流体の中に円柱を流れに垂直な向きに入れた時、円柱の下流に生じる。実験室で作られるカルマン渦列のサイズは普通1 m以下の小さなものであるが、済州島付近に発生するカルマン渦列はとてつもなく大きなものである。これはただ単に流体の流れる速度の違いから生じているのではないことが分かっている。流れを特徴つけるレイノルズ数を計算するとこの二つのカルマン渦列では10°程度のオーダの違いがある。カルマン渦列に階層構造があって、より小さなカルマン渦列が発達し大きなスケールのカルマン渦列を形成するのである。この例はパターンがこの入れ子状の構造を持って形成される点で興味深い。

パターン形成に共通していることは外部から系にエネルギーが供給されていてパターンがその一部をその維持や発展のために使い、残りを散逸しているということである。パターン形成を解析する場合まず微分方程式などで現象をモデル化する。この方程式は普通非線型になる。また系に供給されるエネルギーの大きさに相当するものは方程式のパラメータで表される。一つのパターンはその微分方程式の一つの解なのだが、問題は、一般解を求めるのではなく（それはしばしば不可能である）、解全体がパラメータによってどのように振舞うかを調べることである。多くのパターンは対象で自明な解がパラメータの変化によって不安定化しそれによって系の状態（解）が対称性の低いものに遷移することによって生じる。この解の不安定化や状態の遷移がどのような構造で起こるのかといった定性的性質が問題になるのである。そこでロジスティックマップといわれる離散力学系を用いてもうすこし具体的に見てみることにする。

② ロジスティックマップ——無限の分岐とカオス

離散時間での状態発展

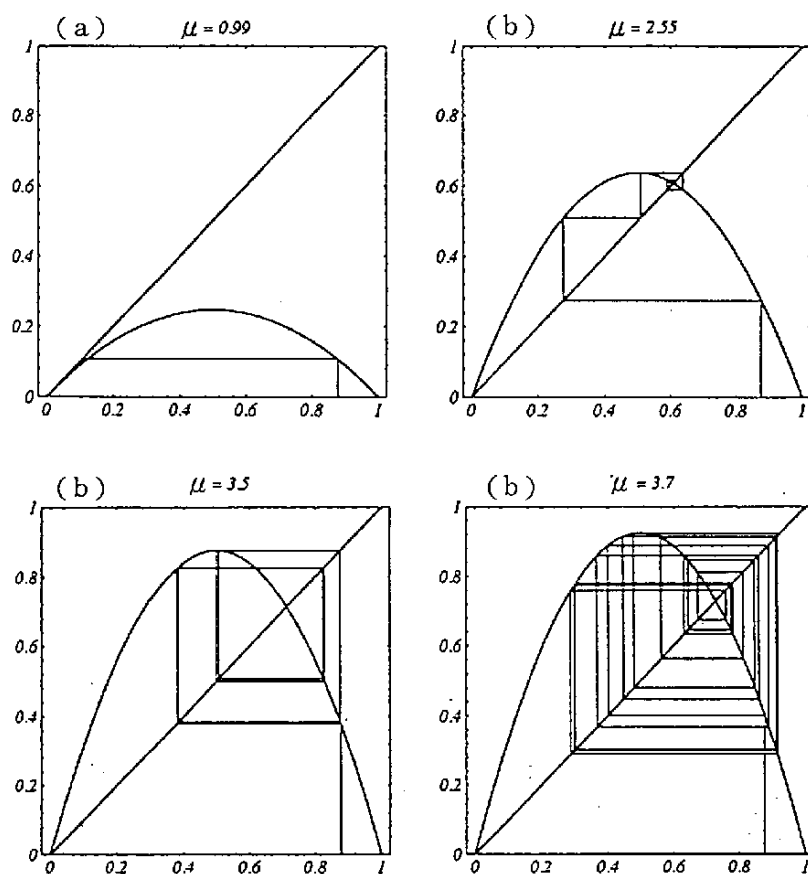
$$x_{n+1} = f_{\mu}(x_n)$$

すなわち時刻 $n + 1$ の状態 x_{n+1} が一つ前の時刻の状態 x_n のみによって決まる系を一般にマップと呼ぶ。 μ はパラメータである。マップの周期倍化分岐現象をロジスティックマップ

$$x_{n+1} = L_\mu(x_n) := \mu x_n (1 - x_n) \quad (\mu > 0)$$

を例に取って議論する。ロジスティックマップによる状態の時間発展の様子を縦軸に x_{n+1} 横軸に x_n をとってグラフで表すと2-1図になる。状態が x_n であるとする、 $x_{n+1} = L_\mu(x_n)$ はグラフの x 軸上の点 x_n から y 軸に平行に移動して曲線 $y = L_\mu(x)$ と交わる点での y の値である。 x_{n+1} をさらにマップによって遷移するには得られた交点から x 軸に平行に移動し直線 $y = x$ と交わる点での x 座標から、先程 $x_n \rightarrow x_{n+1}$ を得たのと同様の移動を行えば良い。2-1図の直線はこの時間発展の様子を示すものである。

2-1図(a)は $0 < \mu < 1$ の場合であるが、グラフからわかるように $L_\mu(x)$ の不動点 ($x = L_\mu(x)$ を満たす点であり、 $y = L_\mu(x)$ と $y = x$ の交点の x 座標) は $x = x_0 = 0$ のみである。また x_0 の近くを初期値とする軌道は x_0 に収束する。この時 x_0 を安定不動点と呼ぶ。これに対し $\mu > 1$ の場合 (2-1図(b))、 x_0 が不動点であることには変わりないが x_0 の近傍を初期値とする軌道は一般に x_0 から遠ざかってゆく。このような不動点を不安定な不動点と呼ぶ。一般に不安定な不動点は観測されることはないが、解全体の性質を考えるとときには重要になる。 x_0 の性質が質的に変わる (安定 $\leftrightarrow$ 不安定) のはパラメータ $\mu = 1$ のときである。このように解が質的に変化することを分岐現象という。ロジスティックマップの分岐の様子をもう少し詳しく見てみよ



2-1図 ロジスティックマップによる時間発展。(a)は原点 $x_0 = 0$ が安定な場合でパラメータ $0 < \mu < 1$ 。(b)は x_0 が不安定でパラメータ $\mu > 1$ の場合。

う。不動点の安定性は不動点でのマップの微係数を調べることによって明らかになる。上の例からわかるように

$$\left| \frac{dL_\mu}{dx} \right| < 1 \rightarrow \text{安定}, \quad \left| \frac{dL_\mu}{dx} \right| > 1 \rightarrow \text{不安定}$$

である。(微分方程式による時間発展では不動点の安定性は微係数の符号で決まり、それが負ならば安定で正ならば不安定である。) 多次元系では各成分について同様の議論により安定性の議論をすれば良いことになる。

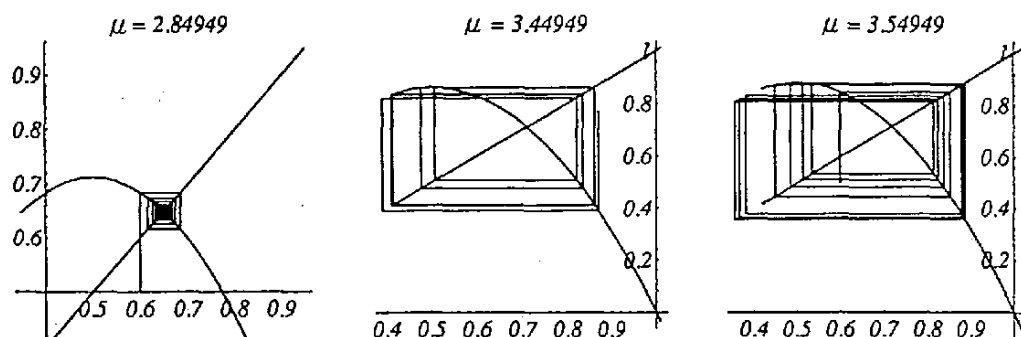
ロジスティックマップにおいて $\mu = 1$ を境に起きる分岐は x_0 に関したただけではない。二次方程式 $x = L_\mu(x)$ の解の判別式からわかるように $\mu > 0$ では新しい不動点 $x_1 = 1 - \frac{1}{\mu}$ が出現する。この不動点の安定性は $\left. \frac{dL_\mu}{dx} \right|_{x=1-\frac{1}{\mu}} = 2 - \mu$ から、

$$1 < \mu < 3 \rightarrow x_1 \text{ は安定}$$

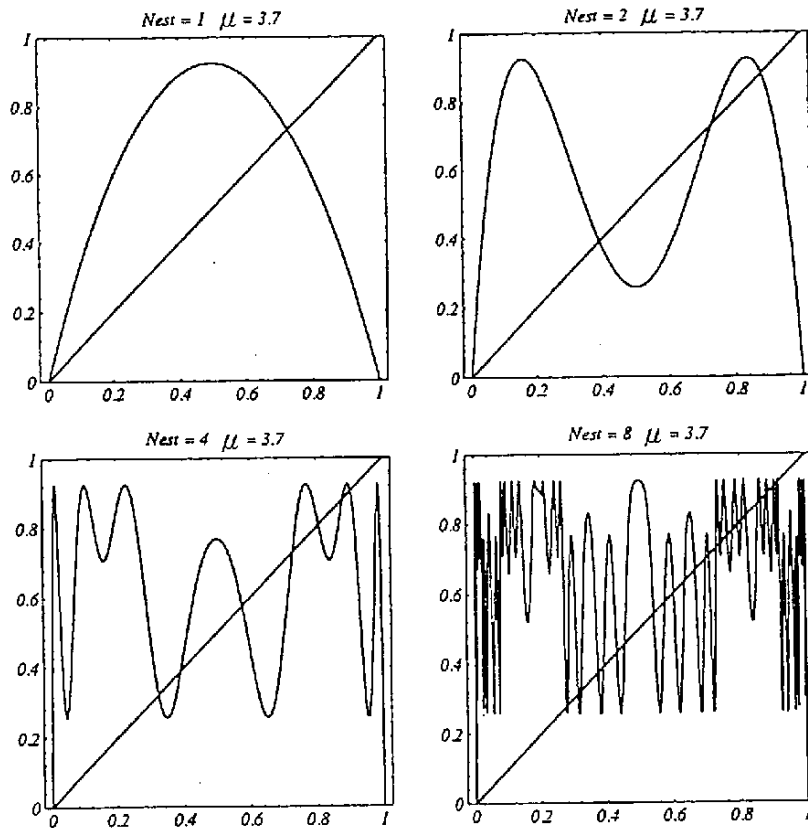
$$3 < \mu \rightarrow x_1 \text{ は不安定}$$

であることがわかる。これは x_0 で起こったのとは異なる場所での分岐である。さらに x_1 が不安定化した直後に x_1 の周りに (x_2^+ と x_2^- の値を取る) 2 周期解が出現することがわかる。

この周期解 2-2 図の安定性はどのようにして調べられるであろうか。 x_2^+ が 2 周期であるので、これは $L_\mu^2(x) := L_\mu(L_\mu(x))$ の不動点であることを利用する。つまり、 $L_\mu^2(x)$ の x_2^+ の微係数の x_2^+ での値からこの 2 周期解はパラメータが $3 < \mu < 1 + \sqrt{6}$ で安定、それより大きい場合には不安定になることがわかる。この分岐はここで止まらず x_2^+ が不安定化するとそれぞれの周りに安定な 4 周期解が出現しその後それらが不安定化する。このような振舞からこれらの分岐を周期倍化分岐(period doubling bifurcation)と呼ぶ。ロジスティックマップの周期倍化分岐は周期 $T=4$ で終わらず、実は無限に続いて行くことが知られている。2-3 図の $y = L_\mu^k(x)$ ($k=1, 2, 4$) を見るとわかるように $L_\mu^k(x)$ は所々で $L_\mu(x)$ が折り畳まれた形をしているため 2 周期解が出現するメカニズムと同じことが時間発展させて行くとあちらこちらで起きるのである。このように周期倍化分岐は無限に続いて行く。この分岐が無限に発展した時の状態が 1 つのカオスである。



2-2図 ロジスティックマップに於いて x_1 が不安定化しその周りに周期解ができる。



2-3図 $y = L_{\mu}^k(x)$ $k=1, 2, 4, 8$ $\mu=3.7$

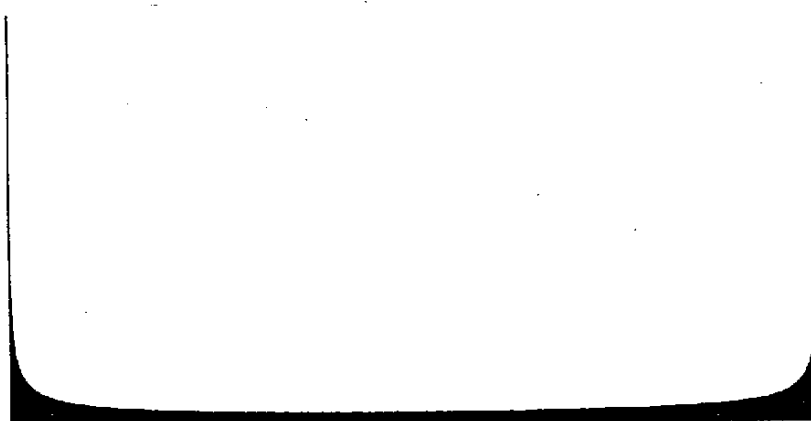
③ カオスの性質

次にカオスの性質について簡単に議論してみよう。まずカオスにおける系の状態の遷移（軌道）は複雑な振舞を示す。「複雑」とはなにかについては多くの議論があるが、ここではあまり厳密なことをいわずにランダム性を持っていると考えていただきたい。ランダム性は確率的な事象に付随したものと考えられているが、カオスが発生するランダム性は決定論的な法則から生じている。2-4図は $\mu=4.0$ のロジスティックマップによって生成される軌道の分布を示したものである。これからわかるように軌道はアトラクタの中をくまなく動きまわっていることがわかる。次に、この複雑な軌道がロジスティックマップのような簡単な仕掛けによって発生していることに注意していただきたい。（「簡単」とは何かという定義ははっきりしないが、それをシミュレートするためのプログラムの行数が少ないとか、デヴァイスでの実現が用意であるとかを考えてもらえば充分である。）想像をたくましくするならば、知的情報処理が示すべき複雑なシステムの振舞を実現するのに必ずしも巨大なプログラムは必要無いかも知れないという可能性がカオス系によって暗示されている。またカオスの発生する軌道はランダム性を持っているに関わらず、ロジスティックマップの定義からわかる通り (x_n, x_{n+1}) をプロットすれば2次曲線が現れる。一般にカオス軌道にはストレンジアトラクタと呼ばれる（隠された）構造（パターン）を持っていて、状態空間内で局在した部分空間内に存在する。ストレンジアトラクタの構造

(パターン) はカオスの特徴つける上で重要であり、乱数のようなものとは異なった性質を示す。この点をどのように応用するかがカオステクノロジーの重要な課題になると思われる。カオス軌道の特徴の最後に、軌道の初期値に対する不安定性をあげておく。どのように小さい初期値の誤差であっても軌道の振舞いには大きな差異をもたらすのである。軌道の安定性は微係数によって決まることを述べたが、マップのストレンジアトラクタ内においては、微係数の平均(リャプノフexponent)で1より大きくなっている成分がある。この不安定性がカオス系の制御を難しくしている原因の一つであるが、これを情報の発生と考えることもできる。何故なら、ある観測によって系の状態がわかるのだが、観測は必ず有限の精度で行われるため、たとえ一度観測しても、カオス系で時間発展した場合にはその軌道について予測(特定)することは再び観測すること無しには不可能だからである。

以上のカオス軌道の特徴をまとめておく。

- 複雑性(ランダム性)
- 簡単な仕掛けによって発生する
- 構造を持っている(ストレンジアトラクタのトポロジ)
- 初期値不安定性



2-4図 ロジスティックマップによる軌道の分布。 $\mu = 4.0$ 、 10^7 回繰り返し
横軸が状態—縦軸が出現確率

④ 振動子の結合系

今まで1次元のマップについて考えて来たが、神経系、免疫系、生態系、進化、社会・経済システムなどはロジスティックマップに相当するそれ単独で一定のダイナミクスを持っている“振動子”が多数結合したシステムであると考えられる。例えば神経系では神経細胞がそれに相当し、社会・経済システムでは個人や会社、国など振動子に相当する。したがって低次元マップが結合系をなした場合にどのような振舞を示すかを議論する必要がある。しかしながらこの問題には多くの困難が伴い現在理論的にも多くのことが解明されずにある。すなわち様々なシステムの数理モデルが個別に研究されているのだが、個別のモデルの性質の解析に手いっぱいであるとい

うのが一般的な傾向である。ここでは、振動子の結合系の比較的単純なモデルに於ける集団引き込みと振動子のクラスタ形成について紹介する。

まず引き込み現象について述べる。初夏の夜に果敢なくも幽玄に光る蛍の群れは、良く観察すると全ての蛍が同じタイミングで光を放っている。物理的な用語を使えば位相を揃えて点滅しているのである。たとえ個々の蛍の点滅のサイクルが同じであったとしても光を放ち始める時刻はばらばらになるはずで、まして点滅のサイクルに関する蛍の個体差があると考えたとこの現象は蛍同士が何らかのコミュニケーションを行っていると考えないと説明がつかない。このような現象を相互引き込みとか集団引き込みなどと呼んでいる。すなわち各振動子の“位相”が振動子間の結合（通信）によってcoherentになる現象が引き込み現象である。

次に簡単なモデルについて少し詳しく見てみよう [1]。振動子として($f_\mu(x)=1-\mu x^2$) というダイナミクスを持つモデルを考える。この関数は先程まで議論して来たマップを平行移動して対称になる位置に持って来たもので、本質的に変わりはない。この振動子が N 個あってそれら全てが一様な強さで結合している系を考える。すなわち以下の方程式で記述されるシステムを考える。

$$x_{i+1}(n) = f_\mu \left((1-\varepsilon) x_i(n) + \frac{\varepsilon}{N} \sum_{j=1}^N x_i(j) \right)$$

各振動子の初期状態をランダムに選んでからこのシステムを時間発展させると、引き込み現象が観察される。その振舞いはパラメータによって以下の4種類に分類される。

- 全ての振動子が同一の位相（状態）に引き込まれる。
- 小数のクラスタが形成され、同じクラスタに属した振動子は同期して振動し、異なるクラスタに属した振動子は異なるパターンで振動する。クラスタの数はクラスタに属す振動指数に比べて小さい。
- 小数の振動子からなる多数のクラスタが形成される。
- クラスタの形成が見られず振動子がほとんどばらばらに振動する。

ただし、ここでいう“同期”あるいは“引き込み”は、複数の振動子状態が完全に一致しているものではなくその差が他のクラスタの振動子の状態と比べると小さいということを持って定義している。パラメータとこれらの振舞いの関係をみると、 μ を固定して考えたとき、結合 ε が強い程クラスタが大きくなり（クラスタ数が減る）、結合強度を固定したとき、 μ が大きくなりロジスティックマップが尖って来る程、振動のcoherencyが減少する傾向がある。

それではこのシステムの情報処理系としての側面を眺めてみよう。一つのクラスタに属す振動子の状態は、他のクラスタのものと比べてみると同期しているが、細かく見てみると僅かに異なったパターンを形成している。そこでクラスタ内の振動パターンの違いを横軸に観測の精度、

縦軸に固定したある時刻の振動子の状態を取って¹グラフにすると、非一様な木構造が現れる。もしクラスタ内の振動子の位相の僅かな差異がランダムなものであるとすると木構造は現れるがそれは一様なものになるはずである。しかしこのシステムではそうはなっていない。すなわちクラスタ内に精度に応じてより細かく分裂するサブクラスタが存在しているのである。

この事実は情報処理系として次のように解釈することが可能である。すなわち一つのクラスタに属す振動子がある大まかな情報を表現しているとする、クラスタ内のそれぞれの精度のサブクラスタはクラスタ全体が示す情報の詳細に相当する。一般に知識は木構造で表現されることが多いが²この非線型振動子の結合系はそのような情報を表現できる可能性があるのである。

振動子の位相を強制的にその振動子の属すクラスタ以外の振動の位相に変更してみると（つまりクラスタ間の振動子の交換）、小数の交換に対してはそれぞれのクラスタは安定であってその全体的な振舞いに変化は起こらないが多くの振動子を交換するとどちらかのクラスタが不安定となりそのクラスタ内の位相の同期が乱れ始める。この乱れは系全体に波及し新しいクラスタ状態が出現する。つまり振動子の交換によってクラスタの状態が遷移するのである。あるクラスタ状態が遷移した後どのようなクラスタの状態になるかを調べると、クラスタの状態遷移前後で確かに異なっているのであるが、全く違った状態が現れることは稀でむしろ状態の一部が変更される場合が多い。これを上に述べた情報処理的なイメージで言えば、振動子の交換によって関連する情報が検索されていると見ることができる。ここでの検索という言葉の意味は現在の実用のデータベースシステムなどでの検索とは異なっていて、どのような検索が行われるかは情報処理系に多くが任されていて、従来の検索に比べると不確かであると言えるが、検索対象がはっきりしないような不良設定問題的性質のある検索により向いていると思われる。

ここでは、具体的なモデルに少々立ち入って議論したが非線型系の発生するダイナミクスには、③で述べたようにランダム性とそれと相反する構構性が同居している。時にランダム性と言う側面はその制御を困難にするが、この2つの要素が同居しているということをどのように利用するかが非線型系を用いた新しい情報処理技術に取っての課題であると考えられる。

¹基本的にはある時刻でのクラスタ内の振動子の状態をプロットしたものであるが、精度 2^{-n} では状態が $n \leq x < n+1$ (n は正数) ならばそれを $n+0.5$ としてプロットし、精度 2^{-1} の時には $n+2^{-1}k \leq x < n+2^{-1}(k+1)$ ならばそれを $n+2^{-1}k+2^{-2}$ としてプロットする。グラフは横軸に精度を取ってそれぞれの精度で振動子の状態をプロットしたもの。

²例えば、生物が植物と動物に分類され、動物が脊椎動物と無脊椎動物に分類され、…という木構造。

(3) 非線型系による情報処理の例

今まで非線型現象を眺めて来たが、ここではより応用に近いカオスネットワークによる最適化〔2〕について簡単に触れることにする。今まで述べて来たような非線型系に特徴的な現象を積極的に採り入れたテクノロジーに関する研究は最近になって盛んになり始めた分野であって、いまだに ambitious な課題である。従ってこれから述べる 2 つの例に関してもすぐに実際の産業に役に立つかどうかははっきりしているとは言えない。しかしながらテクノロジーが既に直面している難問の数々に対処するのにこのような研究を無視することはできない。

Nozawa は、連続時間のホップフィールド型の相互結合ニューラルネットワークで各ニューロンが自己抑制型のフィードバックを持つ場合のモデルを Euler 型に差分化することによって、ニューロンのダイナミクスがそれぞれでカオスを発生することのできるマップであるようなネットワークモデルを導き、その性質を調べた。モデルを方程式で書くと以下のようになる。

$$\begin{aligned}p_i(n+1) &= F_{q_i(n)}(p_i(n)) \\ q_i(n) &= \frac{1}{T} \left(\sum_{j \neq i}^M T_{ji} p_j(n) + I_i \right) \\ F_q(p) &= rp + (1-r) \left(1 - \frac{1}{2} \left(1 + \tanh \left(\frac{p-q}{2\beta} \right) \right) \right)\end{aligned}$$

ただし、 T_{ji} は第 i ニューロンから第 j ニューロンへの結合強度、 r と β はパラメータである。 $F_q(p)$ で決まるニューロンのダイナミクスの特徴はパラメータによって異なるのであるが、およそ以下のようなものである。

- q が 0 または 1 付近の値の場合、ただ一つの安定不動点が存在する。
- q が 0.5 に近付くと、ただ一つの不動点は不安定化しカオスが発生する。

このようなニューロンのダイナミクスを持つネットワークの結合強度をホップフィールドとタンクが巡回セールスマン問題をニューラルネットワークによって解いた時と同じ値に設定してみる。

(都市数は 10。) ランダムに初期値を設定しネットワークが収束した状態をこの問題の解としたとき、Nozawa の計算機実験では 94.7% が最短経路になっていた。ホップフィールドとタンクの例では最短経路に収束したのは 14.3% であった。

この高い正解率が得られる仕組みを理解するには、このネットワークが収束するまでのダイナミクスの様子を調べなくてはならない。まずエネルギー関数がどのように変化するかを見るとおおよそ激しく振動しているのであるが良く見ると比較的最適解に近いエネルギーを持つ状態にしばらく滞在していることがわかる。つまりこのネットワークもシミュレーティドアニーリングの場合と同じようにローカルミニマに落ち込んでいる。ところが、状態がローカルミニマであると幾つかのニューロンへの入力 q が 0 と 1 の間の値となりそれらのニューロンのダイナミクスがカオス的に

なってしまう激しい振動が起きることがあり、この振動が他のニューロンへも波及するためネットワークの状態がローカルミニマから飛び出る仕組みを提供していると考えられる。ここで特に注意したいことは、このカオスを使ったネットワークのパフォーマンスがランダムノイズを使ってローカルミニマを脱出させる方法と比べた時のパフォーマンスに量的な差が見られるということである。Nozawaの計算機実験では各種のパラメータのチューニングをあまり行っていないのでどちらが良いかという問題に答えることはまだできないが、この違いはカオスが生成するランダムな軌道と確率的な軌道とは明らかに違うものであることを示している。したがって、繰り返しになるがカオスのランダム性と非ランダム性を共に利用することがカオスを積極的に利用したテクノロジー（延いては「非線型テクノロジー」）に取っての課題であると考えられる。

(4) まとめ

ミクロな世界の線型性にも関わらず我々は非線型現象に囲まれている。そしてテクノロジーはこの非線型の世界に対応したものである必要がある。この報告では神経系、免疫系、生態系、進化、社会システム、経済システムなどが非線型の特徴を示す合理的なシステムであると言う事実から非線型現象を積極的に取り込むような形のテクノロジーのパラダイムが存在するのではないかと言う動機に基づいて、非線型システムの理論やその応用を概観して来た。非線型現象は「安定かあるいは不安定か」といった単純なものではなかった。いままで非線型現象はそれが示す不安定性のために工学的にどちらかと言うと敬遠されがちであり、解析の困難さのために理論的な研究もなかなか進まなかったという背景がある。しかしながら現代のコンピュータの進歩によってテクノロジーは非線型の世界に取り組みざるをえない状況になって来ている。例えば画像認識や音声認識においては線型理論を中心とした統計学的手法の有効性と共に限界が広く認識されつつある。またコンピュータプログラムの巨大化に伴ってオブジェクト指向プログラミングパラダイムなどの大規模システム開発のための新たな手法が実用の分野で注目されている一方、プログラムは簡単であったとしてもコンピュータ（機械）が学習し、自己組織することによる、従来は巨大なプログラミングが必要であるようなシステムの実現のための研究がなされている。またチューリングの理論を基にした現在のコンピュータが原理的にシリアル処理であるためその処理速度には原理的な壁が存在する。これに対して並列プロセッサによるコンピューティングの必要性が叫ばれているがこの分野においても先にあげた結合系の研究は重要な役割を果たすと考えられる。

この差し迫った必要性に対して実際の「非線型テクノロジー」というと決して順調に育っているとは言えないかも知れない。80年代にニューラルネットブームが起こったがそれが本当の意味で定着するにはこれからの成果に負うところが大きいと思われる。

ニューラルネットの研究も非線型力学の研究も非常に学際的な傾向がある。宮沢りえの写真集で日本では有名になったSanta FeにあるSanta Fe Instituteはこれらの（正確にはcomplex systemの）分野の研究のメッカであるが、そこには物理（それも核物理、統計物理など様々）、生物、数

理工学、経済学など様々の分野の研究者が集っている。彼らの研究は非常に理論的なものから実用的なものまでその範囲は広い。（セルラーオートマトンの進化の研究などで有名なN.H. Packardは、カオスの理論を用いてある種の予測システムを作りこれに基づいてビジネスを行っている。）

思うに理論から実用へという向きの流れと共に実用から理論へのフィードバックが必要な分野なのである。日本でもこの二つの動力が噛みあった研究が発展して行くことを期待したい。³

参考文献

- [1] K.Kaneko. Clustering, coding, switching, hierarchical ordering, and control in a network of chaotic elements. *Physica D*, 41 : 137 - 172, 1990.
- [2] H.Nozawa. A neural network model as a globally coupled map and applications based on chaos. *CHAOS*, 2 (3) :377 - 386, 1992.

³著者自身も一人の研究者として、「期待したい」と言った他人ごとで済ませようとは思っていないということを付け加えておきたい。

3. ニューラルネットワークによる経済予測

3. ニューラルネットワークによる経済予測

(1) 現在までの経緯

昨年度は、実際にニューラルネットを用い、わが国の一般的なマクロ経済指標を対象として「経済活動はパターンである。」という仮定に基づき、実際におきた現象と次におきる現象の組み合わせをニューラルネットに学習させた、その結果を用いて、現在の現象をニューラルネットに与えることにより次におきるであろう現象が予測可能かどうかの検討を行った。この結果、短期間の予測であれば可能であるとの結果を得たが、以下のような課題が残った

(今後の課題)

- ・ソフトウェアを経済予測に使用しやすい様に改良する必要がある。
- ・データの与え方について再考する必要がある。
- ・年次データを四半期データとすることにより学習データを増やした場合の学習期間及び予測精度の変化について調査する必要がある。

本年度は以上のような課題をうけて、昨年度に引き続き、我が国の国民総生産を構成する各要素について、四半期データでの予測手法について検討を行った。

(2) ニューラルネットワークの基本仕様について

ソフトウェアは、昨年開発したものを用了。今回はCPU (486-16MHZ)に数値演算コプロセッサを付加したことにより、演算速度が昨年の約60倍に高速化され、学習に時間がかかるという問題はほとんど解消された。

今年度、用了ニューラルネットワークの仕様を以下に示す。

基 本 仕 様	
ネットワークタイプ	階層型
学習方法	バックプロパゲーション (ウェイト、しきい値変化)
階層数	3層
伝達関数	シグモイド
層間結合タイプ	全結合
入・出力値	0～1 (たとえば0.521 等の中間値も使用可能)

(3) ニューラルネットの学習データについて

① 昨年度のまとめ

使用したソフトウェアでは、入力値は-1から1の間に制限されている。また、出力値 y は第3層において

$$y = 1 / (1 + e^{-x}) \quad [x : \text{第2層から第3層への入力信号}]$$

となるため、出力値も0から1の間に制限される。よって、ニューラルネットに用いる入力・出力データは、ともに0～1の間にある方が学習データを作成する上で効率的である。

また、ニューラルネットでは未学習領域でのパターン処理は、学習済の領域に比べ精度が下がることがわかっている。このため昨年度の研究では、当該年度の入・出力データは、時系列データの増減が最大でも±100%の範囲内におさまると仮定（シグモイド変換後の値は0.3～0.7）し、予測のための入力パターンが学習領域の中に収まるであろうと考えられる前期比をシグモイド変換したものをを用いた。

今年度は、四半期単位での予測を行うに必要な入・出力データについて検討を行った。なお、入・出力データについて分かっていることは以下のとおりである。

- a. 昨年開発したソフトウェアを用いたため、入・出力データの制限値は昨年と同じである。このため、入出力データの値は0～1の間とした
- b. 四半期データでは、経験により入力データは3期分必要である。

② 学習データの作成

昨年同様、前期比（前期からの伸び）等をシグモイド変換した場合には、入力データとして前期比、前々期比、前々々期比を用いる。この場合、各期ごとに計算する必要があるため、学習期間を10期分とすると30のデータを用意する必要がある。また、出力値も来期／今期の計算を各期毎に行う必要があるため、比を用いると合計で40のデータが必要であり多くの計算を行う必要がある。

さらに、出力値から原数値（予測値）に戻す場合、対数変換（シグモイド変換の逆変換）を行い前期の原数値に掛けてやる必要がある。

このように比を用いる場合には非常にたくさんの計算を行わねばならず、昨年の課題にあるようなソフトウェアを経済予測に使用しやすい様に改良する必要がある。

一方、指数を用いる場合は、学習期間の最大値以上の値を1として指数化すれば、入力データは0～1の範囲におさまる。また、学習期間を10期分とすると13（学習する最初の期の前2期分のデータ及び最後の期の次のデータを含む）のデータを用意し、今期を含め前2期分を入力パターンとして、次期を出力パターンとして用いればよい。さらに、出力値はそのまま原数値（予測値）として読むことができる。

このように、四半期データの様に学習に多くのデータが必要な場合、指数化の方がデータを作成し、予測を行う作業は簡単かつ効率よくできる。

次に指数を用いた場合と比を用いた場合の予測結果を比較するため、入力指標として昨年同様9指標（C:民間消費支出、IH:民間住宅投資、IE:民間企業設備投資、J:在庫（在庫品増加を累計したもの）、M:輸入品等、E:輸出等、INT:公定歩合、CG:政府最終消費支出、IG:公的投資）、出力指標（予測を行うもの）として4指標（C:民間消費支出、IH:民間住宅投資、IE:民間

企業設備投資、J:在庫（在庫品増加を累計したもの））を用い、入力データは9種×3期分=27個、出力データは4種×1期分で4個、中間層のニューロン数は11であるニューラルネットで予測を行った。なお、結果は以下のとおりである。

地域経済予測サポートシステムに用いた指標

入力データ

記号	変数名	単位	データの範囲	出 所
C	民間最終消費支出	10億円	全国	経済企画庁国民経済計算
I H	民間住宅投資	10億円	全国	経済企画庁国民経済計算
I E	民間企業設備投資	10億円	全国	経済企画庁国民経済計算
J	在庫品（在庫品増加の累計）	10億円	全国	経済企画庁国民経済計算
M	輸入等	10億円	全国	経済企画庁国民経済計算
C G	政府最終消費支出	10億円	全国	経済企画庁国民経済計算
I G	公的投資	10億円	全国	経済企画庁国民経済計算
E	輸出等	10億円	全国	経済企画庁国民経済計算
INT	公定歩合	%	全国	日本銀行

出力データ

記号	変数名	単位	データの範囲	出 所
C	民間最終消費支出	10億円	全国	経済企画庁国民経済計算
I H	民間住宅投資	10億円	全国	経済企画庁国民経済計算
I E	民間企業設備投資	10億円	全国	経済企画庁国民経済計算
J	在庫品（在庫品増加の累計）	10億円	全国	経済企画庁国民経済計算

a. 指数による予測

予測に用いたデータの一部を3-1表に示す。1972年1四半期の民間最終消費支出(C)についてみると10,951（原数値）/100,000（最大値）=0.11として指数化される。なお、公定歩合については比とデータを統一するため、 $1 / (1 + e^{- (公定歩合の原数値 / 10 + 1)})$ で計算している。このようにして指数化した学習データを3-2表に示す。

3-1表 予測に用いたデータ

C	I H	I E	J	M	E	INT	CG	I G	
69	7358	818	2671	10063	1354	1421	5.84	962	1365
	7914	969	2871	10342	1428	1646	5.84	1141	763
	8255	1113	3365	10566	1555	1810	5.97	890	1225
	9773	1068	3662	11938	1652	1941	6.25	1566	1552
70	8569	1041	3663	12056	1735	1776	6.25	1128	1569
	9033	1205	3545	12653	1850	2002	6.25	1321	940
	9589	1290	4043	13179	1931	2164	6.25	1032	1525
	11142	1211	4155	14512	1973	2331	6.07	1974	1857
71	9693	1138	3925	14494	1975	2160	5.8	1319	1850
	10247	1271	3571	14728	1995	2469	5.6	1618	1212
	10745	1363	3907	14823	1901	2653	5.32	1208	1844
	12545	1307	3956	15727	1935	2613	5.23	2276	2294
72	10951	1292	3863	15578	1930	2266	4.75	1537	2476
	11760	1560	3616	15829	1947	2381	4.71	1939	1504
	12545	1785	4140	16040	2088	2734	4.25	1404	2194
	14645	1813	4582	17026	2272	2997	4.25	2657	2699
73	13106	1870	4556	16945	2482	2623	4.25	1795	3059
	13925	2250	4433	17900	2848	2774	5.17	2308	1932
	15149	2516	5495	18221	3072	3115	6.35	1684	2637
	18129	2658	6331	18910	3680	3620	7.22	3550	3203
74	15604	2425	6172	19341	4615	3668	9	2138	2780
	17030	2299	5629	20704	5172	4539	9	3148	2217
	18631	2728	6281	21470	5359	5338	9	2197	3141
	21647	2502	6579	22306	5548	5902	9	4757	3943
75	19082	2087	6063	21536	5024	4765	9	3042	3519
	20069	2628	5482	21777	4948	4906	8.45	3874	2441
	21247	2964	6239	21830	5022	5035	7.73	2668	3390
	24365	2749	6507	22783	5356	5548	6.75	5305	4068

3-2表 指数化した学習データ

1970CQ1	0.086	0.104	0.073	0.121	0.087	0.089	0.651	0.038	0.078
1970CQ2	0.09	0.121	0.071	0.127	0.093	0.1	0.651	0.044	0.047
1970CQ3	0.096	0.129	0.081	0.132	0.097	0.108	0.651	0.034	0.076
1970CQ4	0.111	0.121	0.083	0.145	0.099	0.117	0.647	0.066	0.093
1971CQ1	0.097	0.114	0.079	0.145	0.099	0.108	0.641	0.044	0.093
1971CQ2	0.102	0.127	0.071	0.147	0.1	0.123	0.636	0.054	0.061
1971CQ3	0.107	0.136	0.078	0.148	0.095	0.133	0.63	0.04	0.092
1971CQ4	0.125	0.131	0.079	0.157	0.097	0.131	0.628	0.076	0.115
1972CQ1	0.11	0.129	0.077	0.158	0.097	0.113	0.617	0.051	0.124
1972CQ2	0.118	0.156	0.072	0.158	0.097	0.119	0.616	0.065	0.075
1972CQ3	0.125	0.179	0.083	0.16	0.104	0.137	0.605	0.047	0.11
1972CQ4	0.146	0.181	0.092	0.17	0.114	0.15	0.605	0.089	0.135
1973CQ1	0.131	0.187	0.091	0.169	0.124	0.131	0.605	0.06	0.153
1973CQ2	0.139	0.225	0.089	0.179	0.142	0.139	0.626	0.077	0.097
1973CQ3	0.151	0.252	0.11	0.182	0.154	0.156	0.654	0.056	0.132
1973CQ4	0.181	0.266	0.127	0.189	0.184	0.181	0.673	0.118	0.16
1974CQ1	0.156	0.243	0.123	0.193	0.231	0.183	0.711	0.071	0.139
1974CQ2	0.17	0.23	0.113	0.207	0.259	0.227	0.711	0.105	0.111
1974CQ3	0.186	0.273	0.126	0.215	0.268	0.267	0.711	0.073	0.157
1974CQ4	0.216	0.25	0.132	0.223	0.277	0.295	0.711	0.159	0.197
1975CQ1	0.191	0.209	0.121	0.215	0.251	0.238	0.711	0.101	0.176
1975CQ2	0.201	0.263	0.11	0.218	0.247	0.245	0.7	0.129	0.122
1975CQ3	0.212	0.296	0.125	0.218	0.251	0.252	0.684	0.089	0.17
1975CQ4	0.244	0.275	0.13	0.228	0.268	0.277	0.663	0.177	0.203
1976CQ1	0.213	0.276	0.123	0.219	0.266	0.265	0.657	0.114	0.204

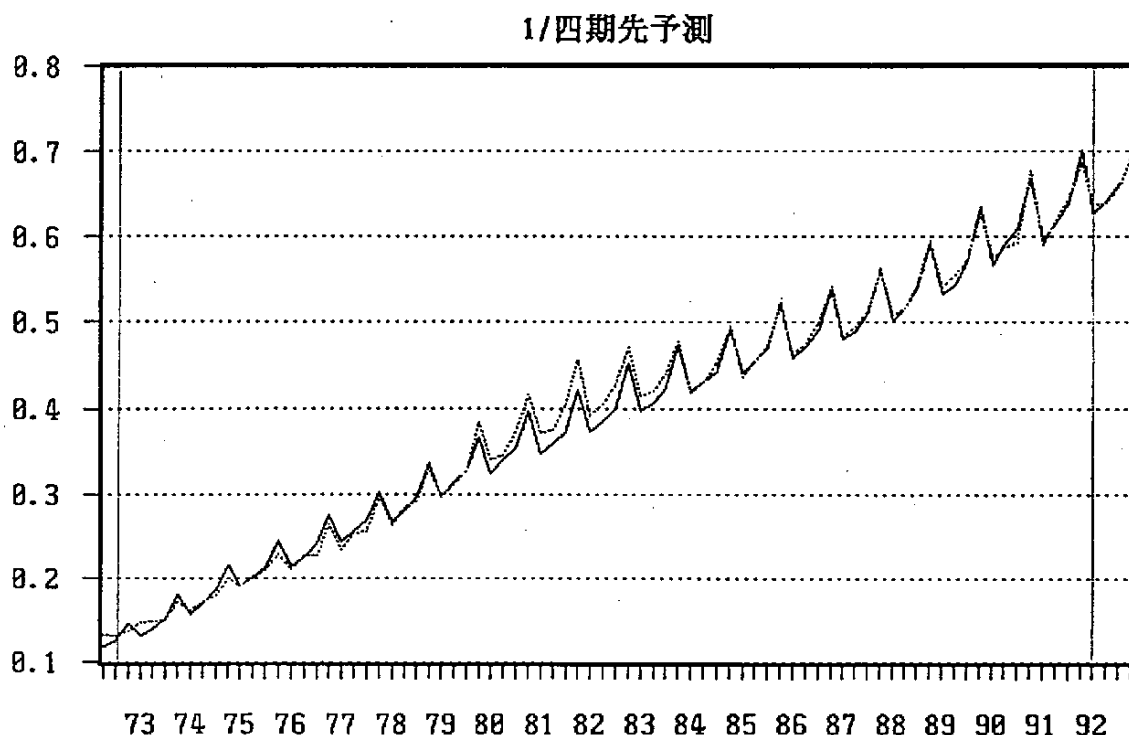
最初に、1972年CQ2から1992年CQ1までの20年間（80期）のデータを2000回学習させた後、学習領域での適合度を見るため実績値を入力データとして一期先の予測を80期にわたり行った結果を3-3図に示す。なお、予測値は見やすいように連続線としている。

図からわかるように指数化したデータを用い1四半期先の予測を行った場合、学習領域の期間においては非常に良好な結果を示した。

次に、1972CQ2から1982CQ1までの10年間（40期）のデータを用いて2000回学習を行った後、実績値を入力データとして1期先の予測を行った結果を3-4図に示す。また、1982CQ2から1992CQ1までの10年間（40期）のデータを用いて同じように予測を行った結果を3-5図に示す。なお、予測値は見やすいように連続線としている。

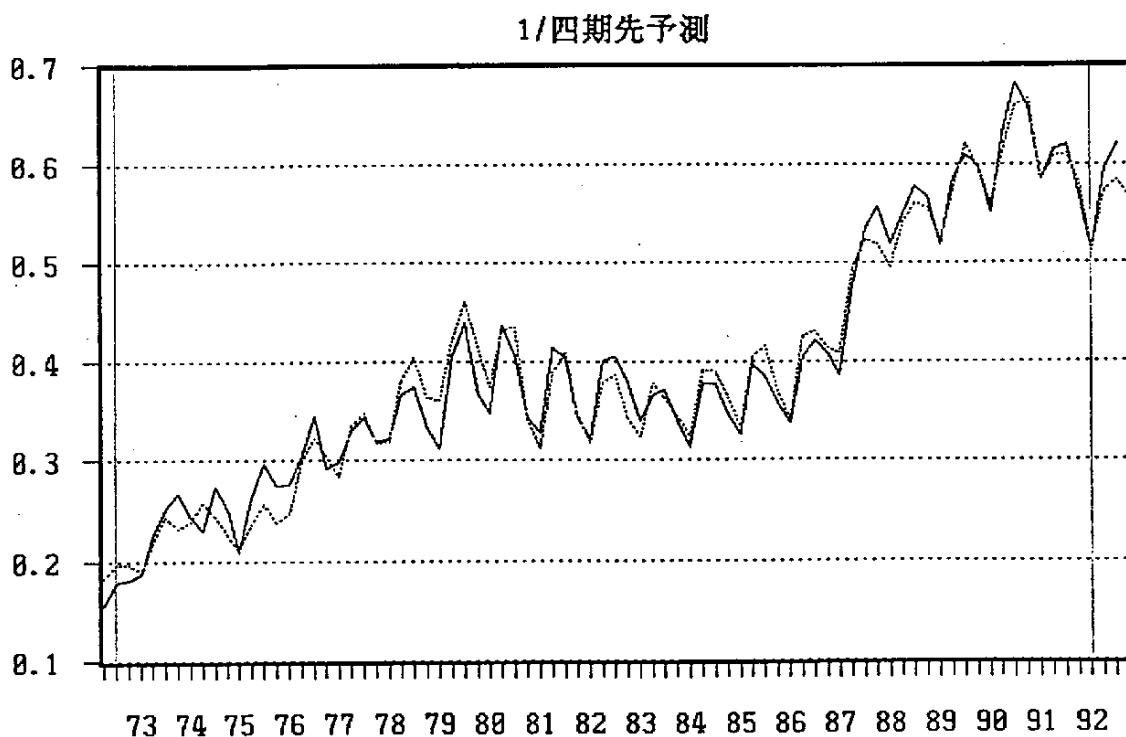
図からわかるように、学習済の期間においては非常に良好な予測結果を示すが、未学習期間では全く予測出来ない状態にある。これは、昨年の報告書でも述べたとおり右上がりの傾向のある時系列データを指数化した場合は、予測のための入力パターンが学習領域の中にないためであり、指数を用いた場合に十分予測された現象である。また、予測のパターンが未学習領域であれば、ニューラルネットでは未学習領域でのパターン処理は、学習済の領域に比べ精度が下がることと一致する。

よって、経済指標の様な時系列データで長期的な予測を行う場合、指数値を入出力データとして用いることはできない。



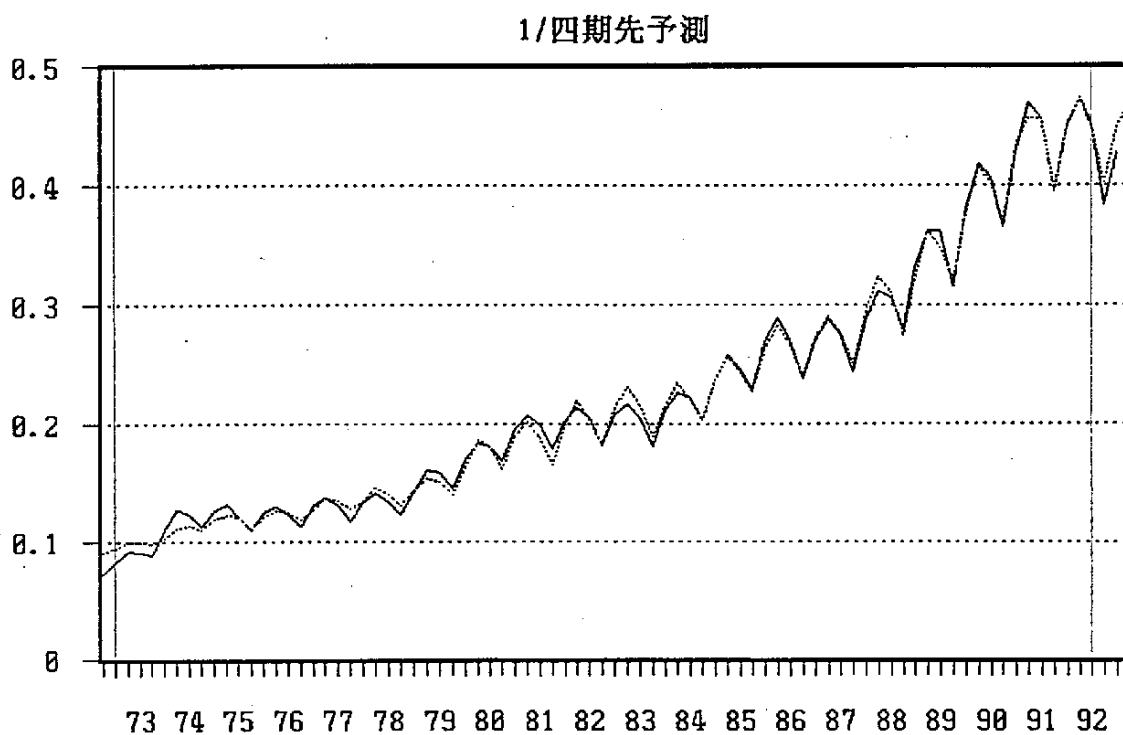
3-3図 指数による学習結果 I

a. 民間最終消費支出



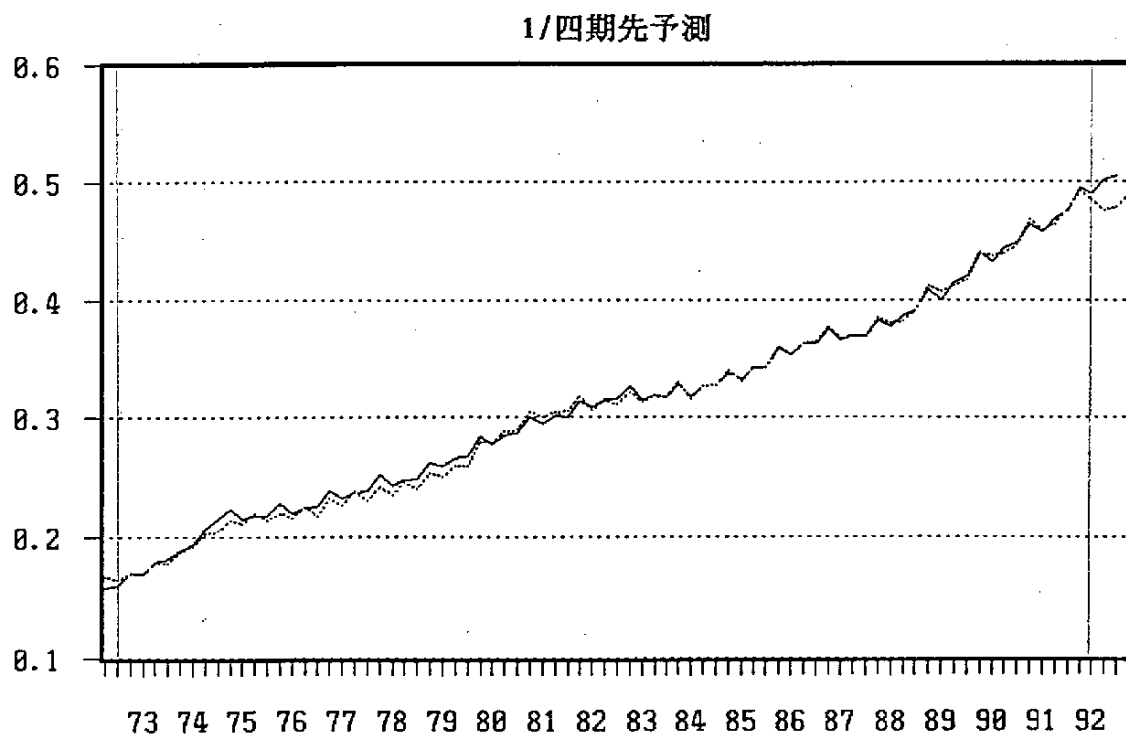
3-3図 指数による学習結果Ⅰ

b. 民間住宅投資

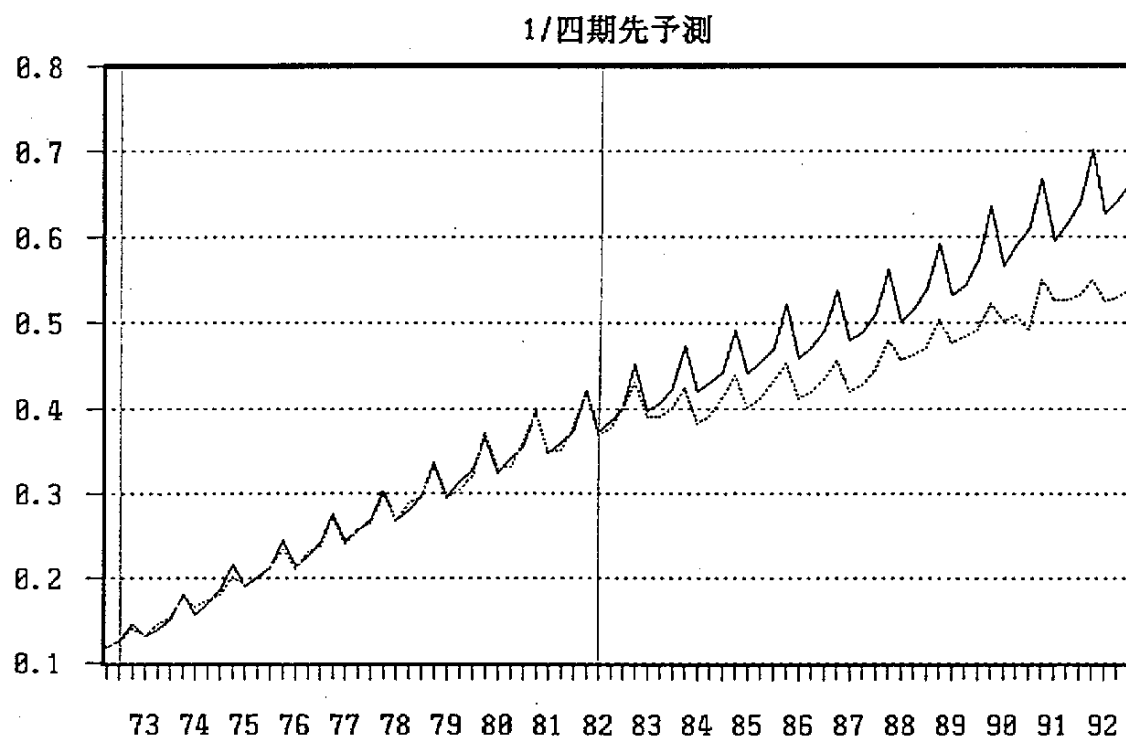


3-3図 指数による学習結果Ⅰ

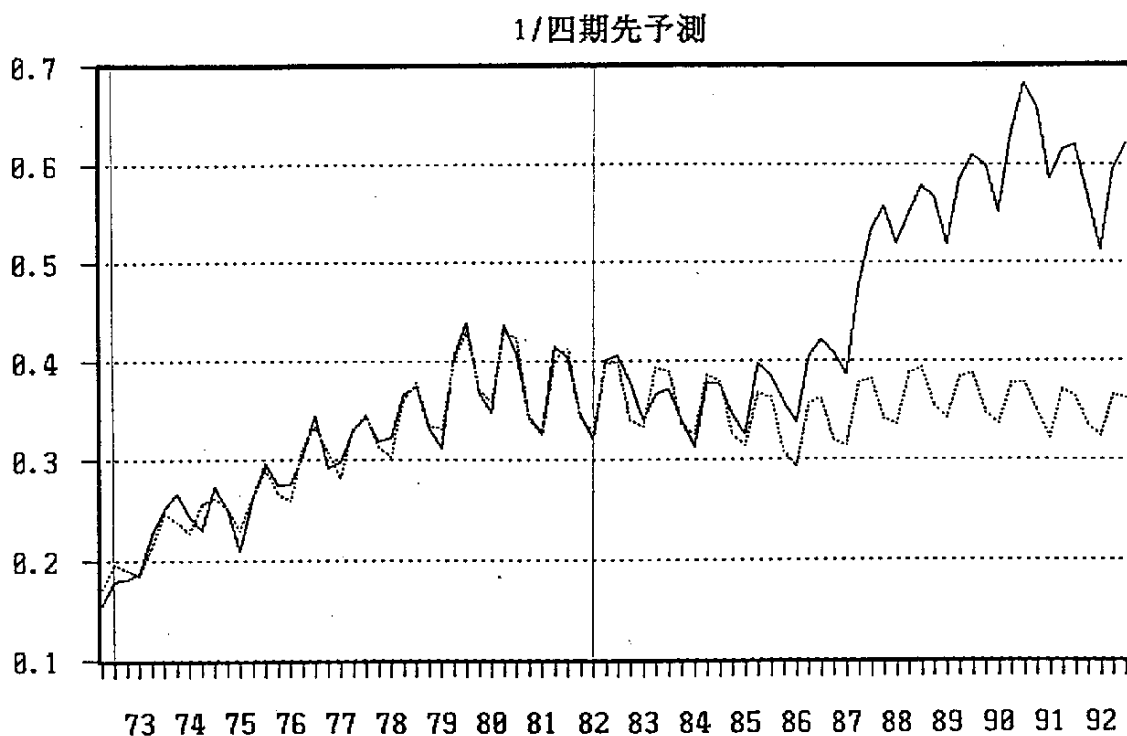
c. 民間企業設備投資



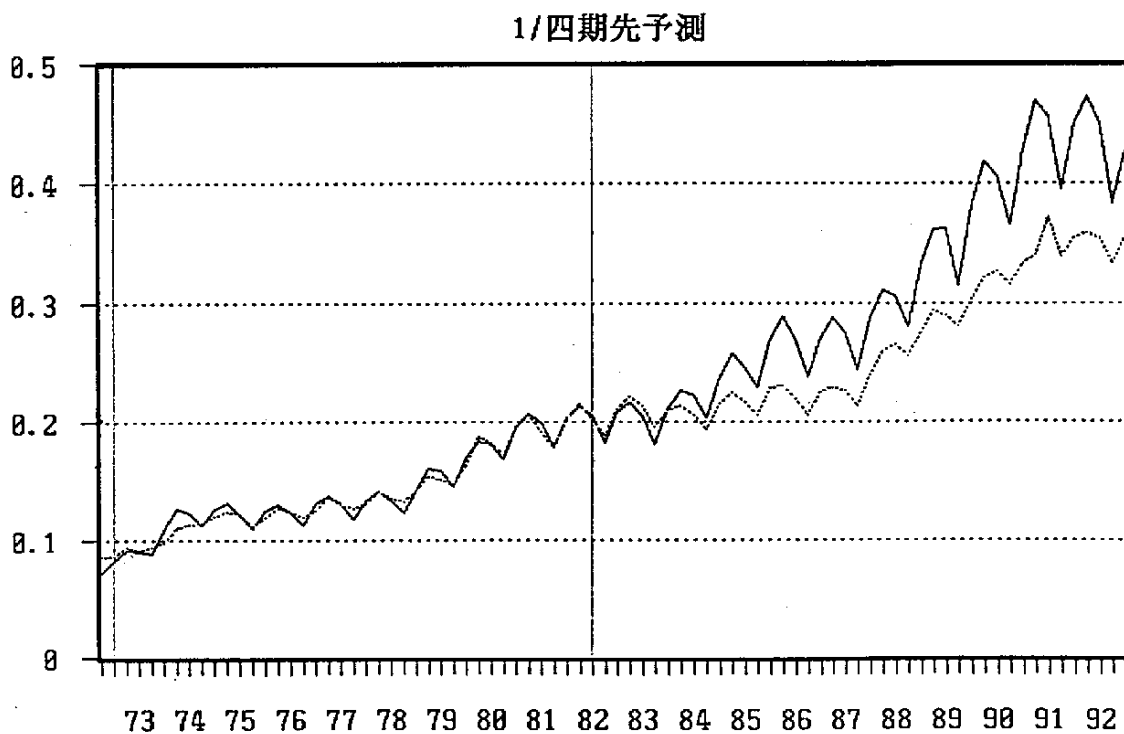
3-3図 指数による学習結果Ⅰ
d. 在庫



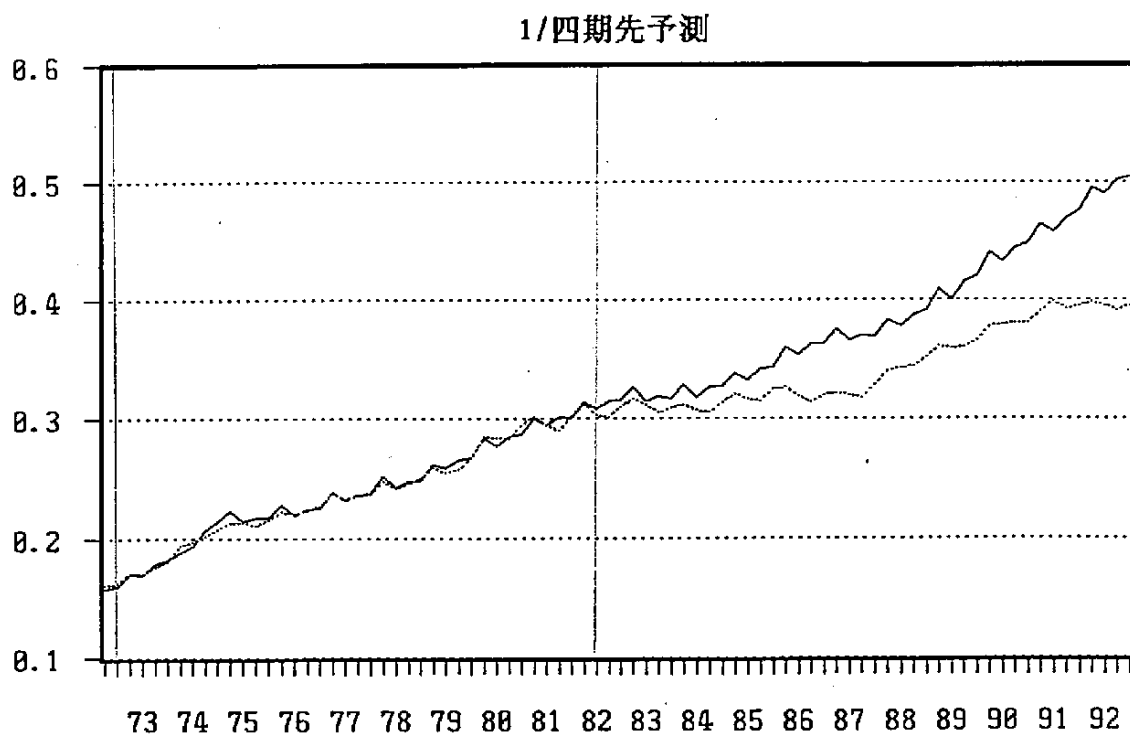
3-4図 指数による学習結果Ⅱ
a. 民間最終消費支出



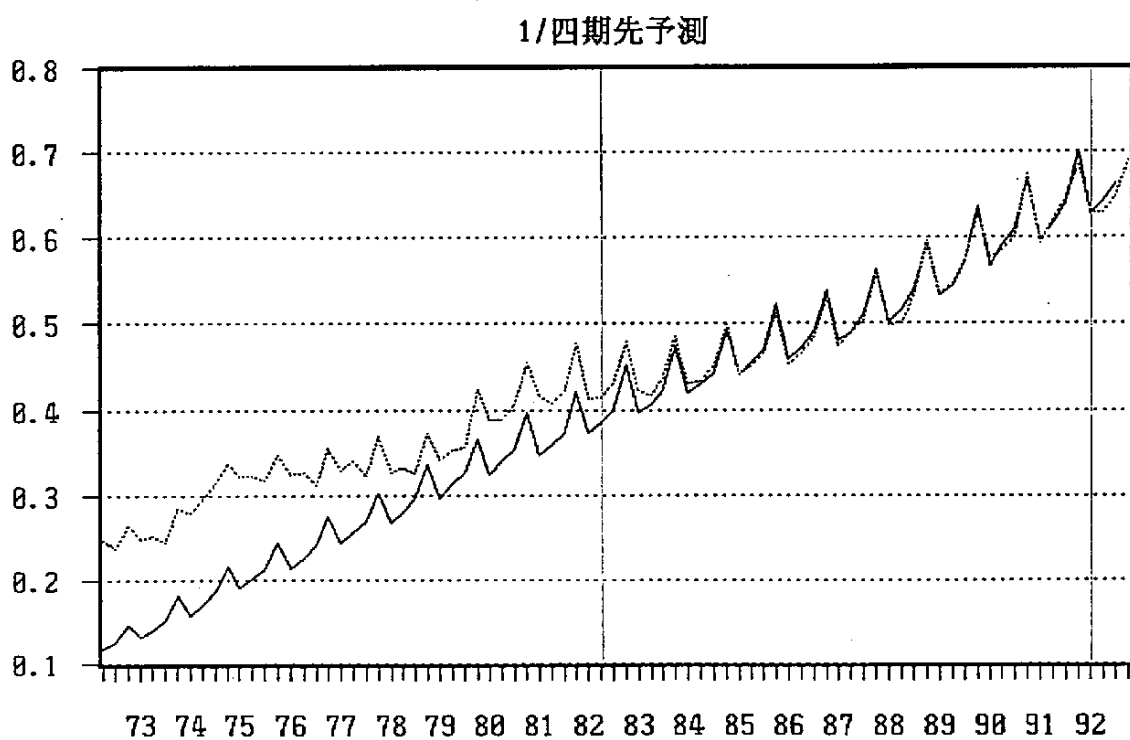
3-4図 指数による学習結果Ⅱ
b. 民間住宅投資



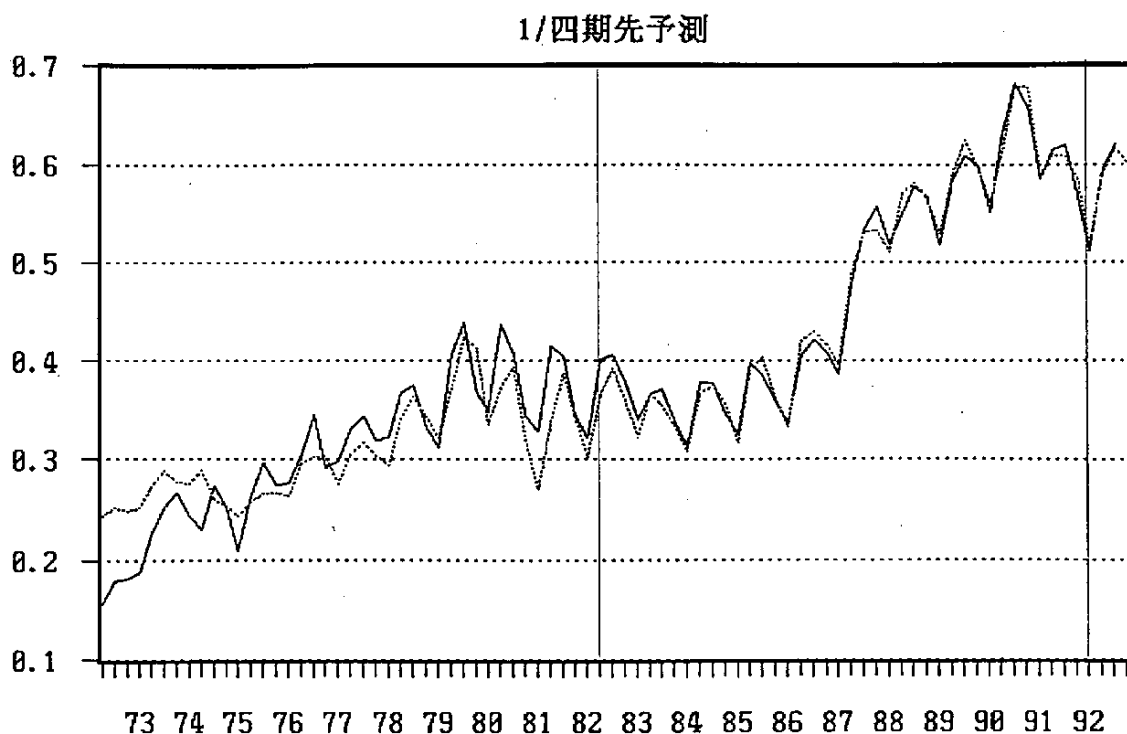
3-4図 指数による学習結果Ⅱ
c. 民間企業設備投資



3-4図 指数による学習結果Ⅱ
d. 在庫

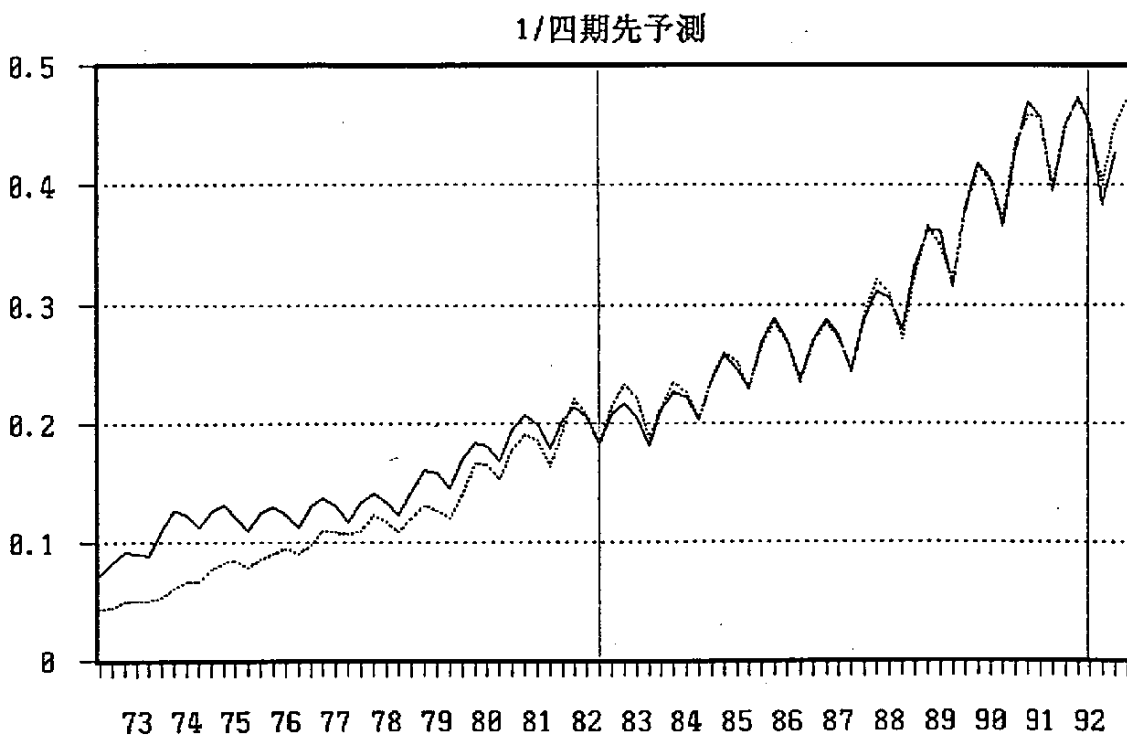


3-5図 指数による学習結果Ⅲ
a. 民間最終消費支出



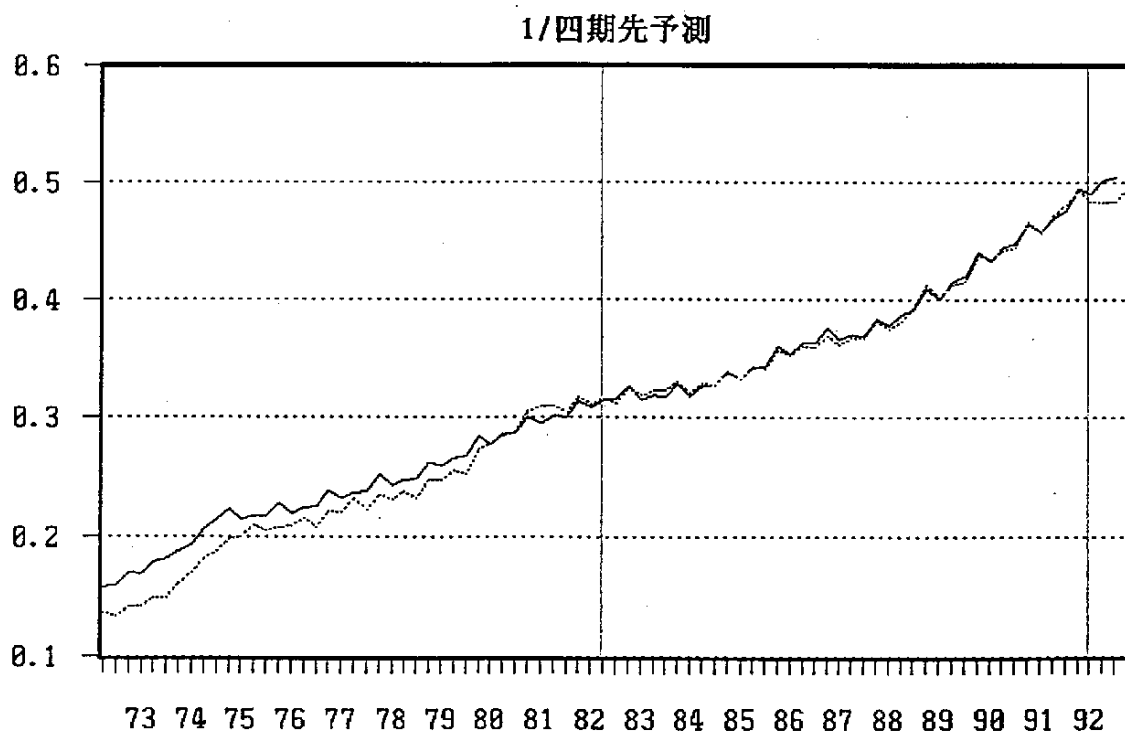
3 - 5 図 指数による学習結果Ⅲ

b. 民間住宅投資



3 - 5 図 指数による学習結果Ⅲ

c. 民間企業設備投資



3-5図 指数による学習結果Ⅲ

d. 在庫

b. 比による予測

予測に用いたデータは、指数の場合と同じものを用いた。1972年1四半期の民間最終消費支出(C)についてみると前期比は $1 / (1 + e - (10,951 / 12,545 - 1)) = 0.468$ として指数化される。また、同様に前々期比は $1 / (1 + e - (10,951 / 10,745 - 1)) = 0.505$ となる。なお、公定歩合については0から約10%の間を変動するため、比を取らず指数化 $1 / (1 + e - (\text{公定歩合の原数値} / 10))$ で計算している。このようにして比を取った学習データを3-6表に示す。

最初に、1972年CQ2から1992年CQ1までの20年間(80期)のデータを2000回学習させた後、学習領域での適合度を見るため実績値を入力データとして一期先の予測を行い評価を行った結果を3-7図に示す。なお、予測値は見やすいように連続線としている。

図からわかるように比を取ったデータを用い予測を行った場合、学習領域の期間においては民間消費支出を除いては良好な結果を示している。

次に、1972CQ2から1982CQ1までの10年間(40期)のデータを用いて2000回学習を行った後、実績値を入力データとして評価を行った結果を3-8図に示す。また、1982CQ2から1992CQ1までの10年間(40期)のデータを用いて同じように評価を行った結果を3-9図に示す。なお、予測値は見やすいように連続線としている。

図からわかるように、比を取ったデータを用いた場合は、学習済の期間においては20年間学

習を行った場合に比べ良好な結果を示す。また、未学習期間でも指数化したデータに比べかなり良い結果を示している。これは、昨年の報告書でも述べたとおり比であれば予測を行う入力データのパターンが学習領域の中にあるパターンとある程度近いため、良好な結果が得られるものと考えられる。

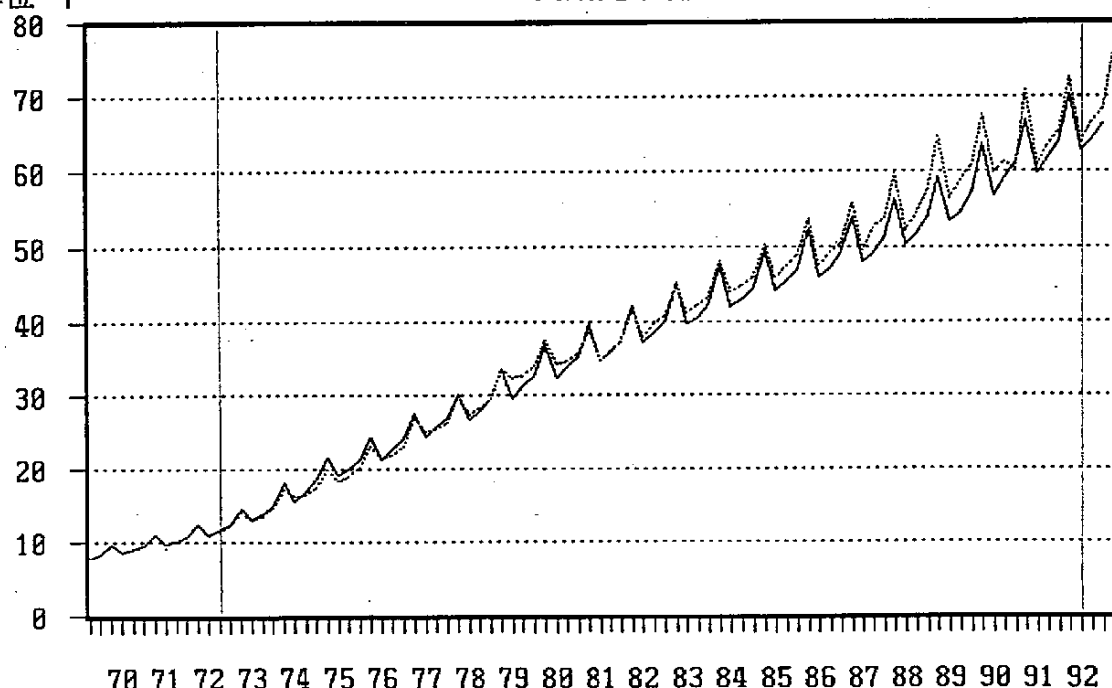
上記から、学習期間については長ければ長いほど良いものではないことがわかった。

3-6表 比をとった学習データ

	C	IH	IE	J	M	E	INT	CG	IG
1970CQ1	0.469	0.494	0.5	0.502	0.513	0.479	0.651	0.431	0.503
1970CQ2	0.514	0.539	0.492	0.512	0.517	0.532	0.651	0.543	0.401
1970CQ3	0.515	0.518	0.535	0.51	0.511	0.52	0.651	0.446	0.651
1970CQ4	0.54	0.485	0.507	0.525	0.505	0.519	0.647	0.714	0.554
1971CQ1	0.468	0.485	0.486	0.5	0.5	0.482	0.641	0.418	0.499
1971CQ2	0.514	0.529	0.477	0.504	0.503	0.536	0.636	0.556	0.415
1971CQ3	0.512	0.518	0.524	0.502	0.488	0.519	0.63	0.437	0.627
1971CQ4	0.542	0.49	0.503	0.515	0.504	0.496	0.628	0.708	0.561
1972CQ1	0.468	0.497	0.494	0.498	0.499	0.467	0.617	0.42	0.52
1972CQ2	0.518	0.552	0.484	0.504	0.502	0.513	0.616	0.565	0.403
1972CQ3	0.517	0.536	0.536	0.503	0.518	0.537	0.605	0.431	0.613
1972CQ4	0.542	0.504	0.527	0.515	0.522	0.524	0.605	0.709	0.557
1973CQ1	0.474	0.508	0.499	0.499	0.523	0.469	0.605	0.42	0.533
1973CQ2	0.516	0.551	0.493	0.514	0.537	0.514	0.626	0.571	0.409
1973CQ3	0.522	0.53	0.56	0.504	0.52	0.531	0.654	0.433	0.59
1973CQ4	0.549	0.514	0.538	0.509	0.549	0.54	0.673	0.752	0.553
1974CQ1	0.465	0.478	0.494	0.506	0.563	0.503	0.711	0.402	0.467
1974CQ2	0.523	0.487	0.478	0.518	0.53	0.559	0.711	0.616	0.45
1974CQ3	0.523	0.547	0.529	0.509	0.509	0.544	0.711	0.425	0.603
1974CQ4	0.54	0.479	0.512	0.51	0.509	0.526	0.711	0.762	0.563
1975CQ1	0.47	0.459	0.48	0.491	0.476	0.452	0.711	0.411	0.473
1975CQ2	0.513	0.564	0.476	0.503	0.496	0.507	0.7	0.568	0.424
1975CQ3	0.515	0.532	0.534	0.501	0.504	0.507	0.684	0.423	0.596
1975CQ4	0.537	0.482	0.511	0.511	0.517	0.525	0.663	0.729	0.55
1976CQ1	0.469	0.501	0.486	0.49	0.498	0.489	0.657	0.412	0.5

単位・千

1/四期先予測

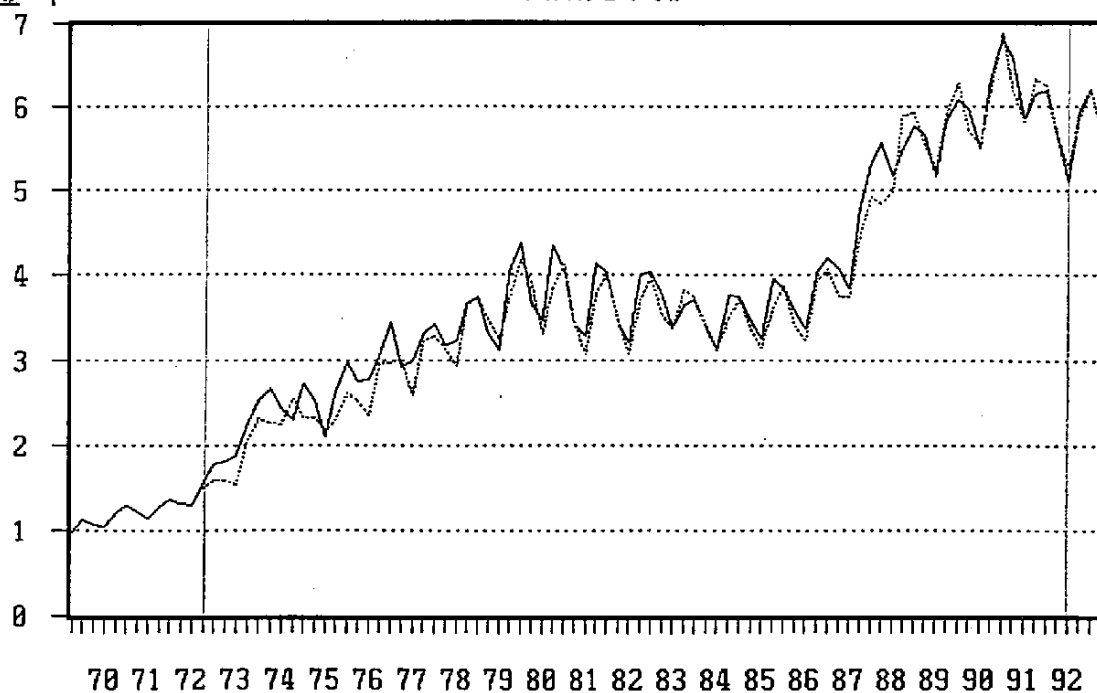


3-7図 比による学習結果 I

a. 民間最終消費支出

単位・千

1/四期先予測

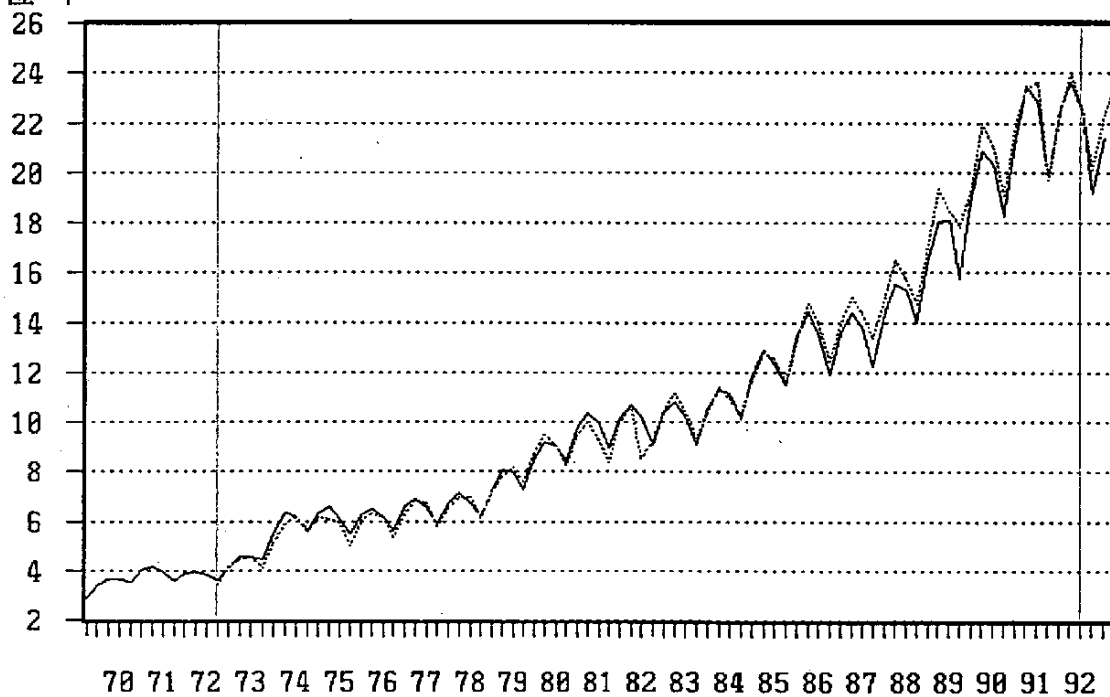


3-7図 比による学習結果 I

b. 民間住宅投資

単位・千

1/四期先予測

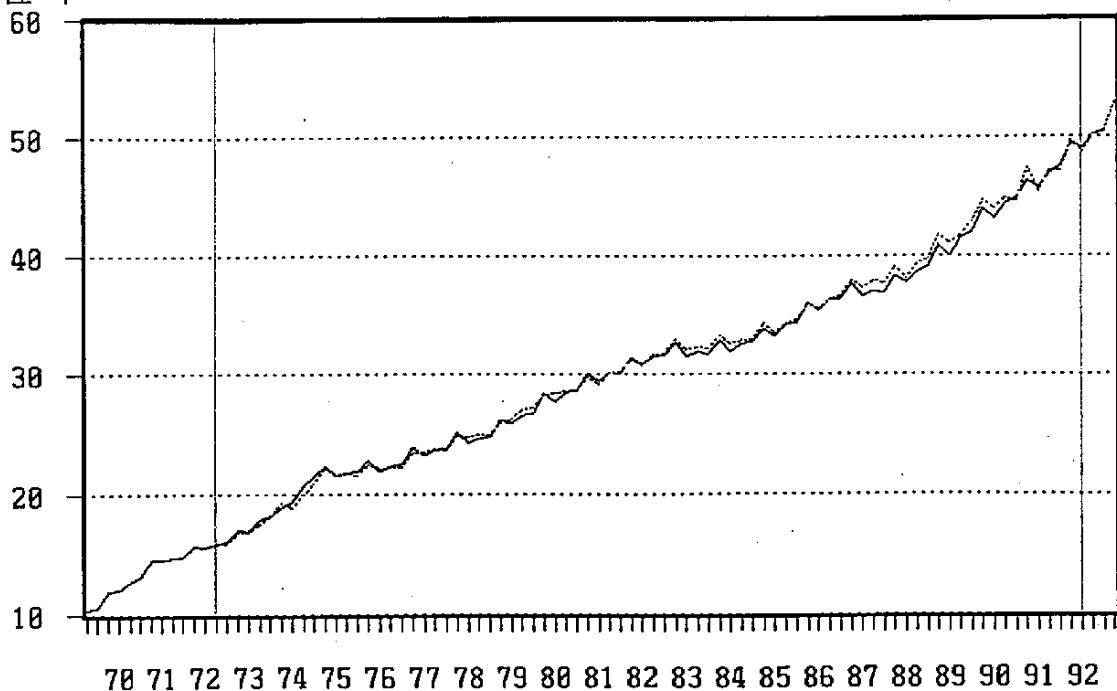


3-7図 比による学習結果 I

c. 民間企業設備投資

単位・千

1/四期先予測

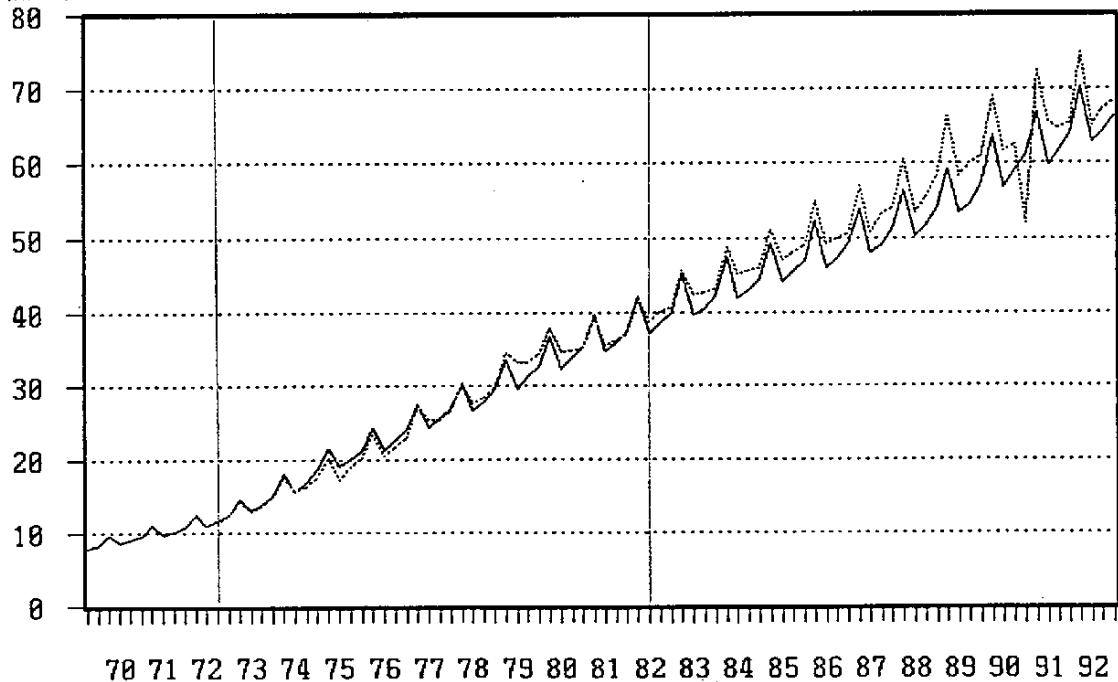


3-7図 比による学習結果Ⅰ

d. 在庫

単位・千

1/四期先予測

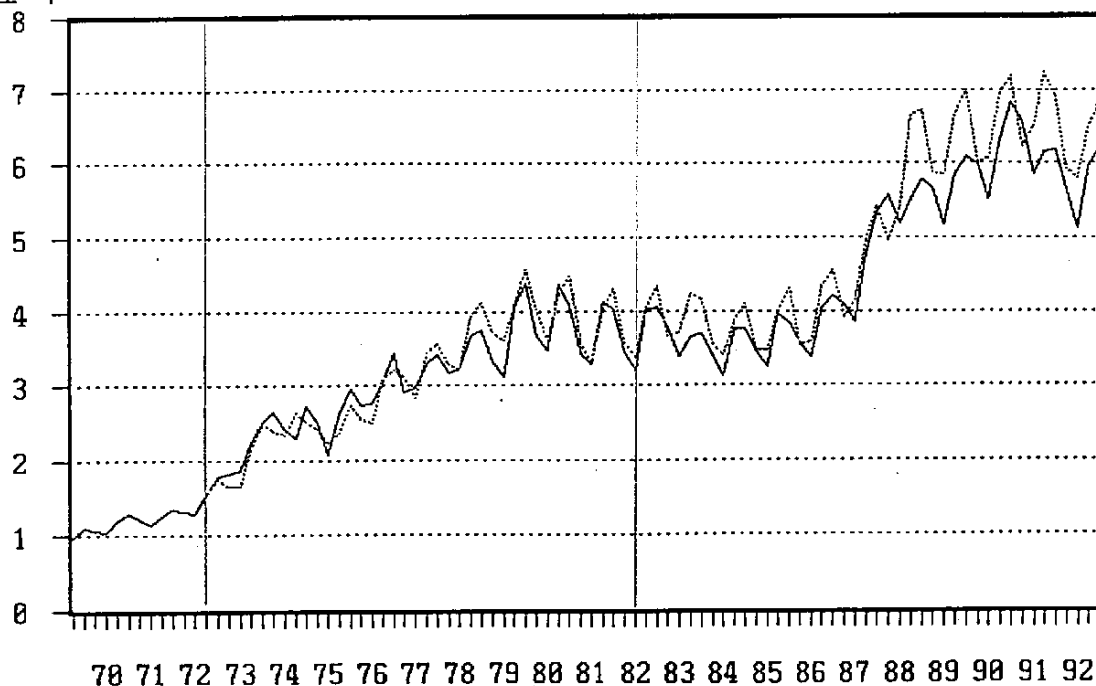


3-8図 比による学習結果Ⅱ

a. 民間最終消費支出

単位・千

1/四期先予測

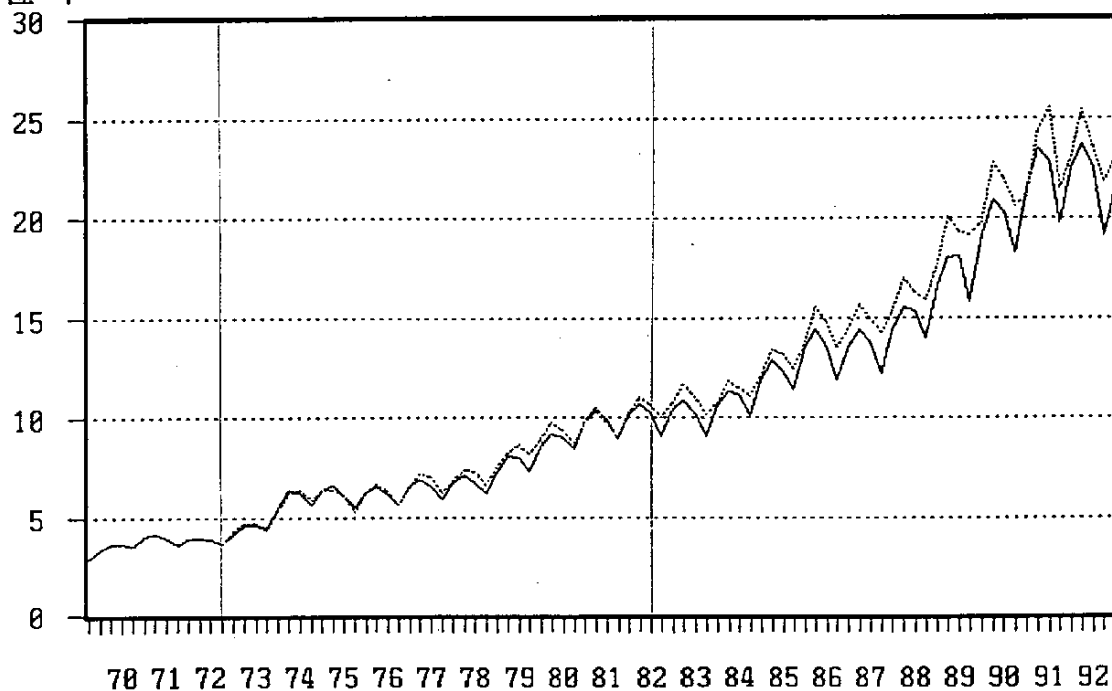


3-8図 比による学習結果Ⅱ

b. 民間住宅投資

単位・千

1/四期先予測

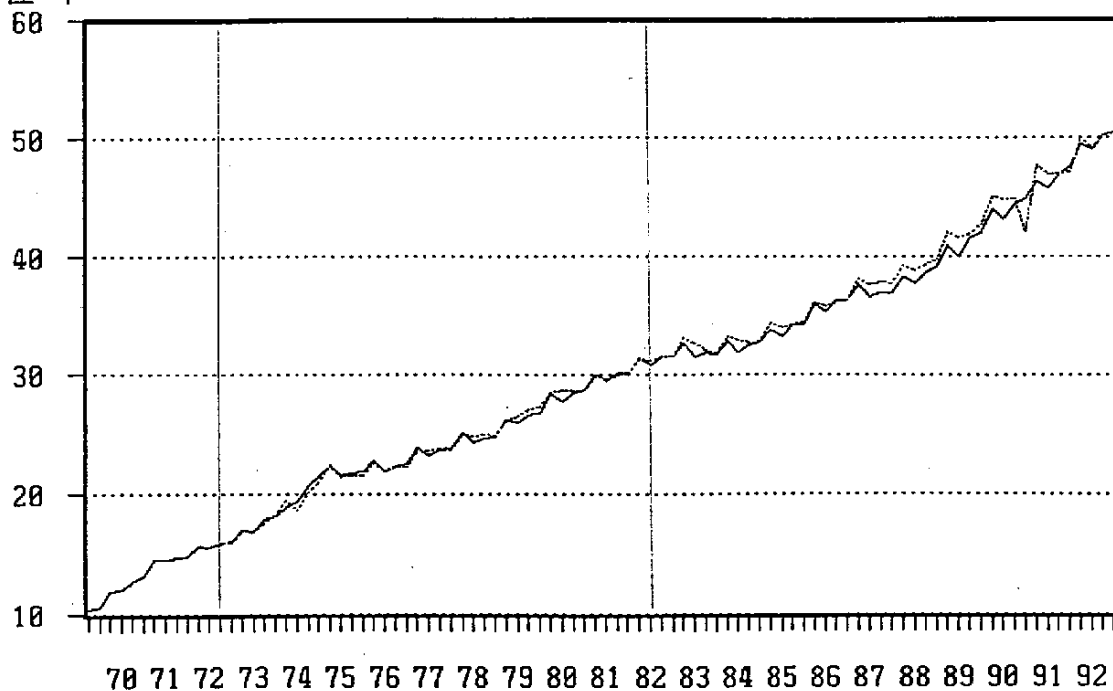


3-8図 比による学習結果Ⅱ

c. 民間企業設備投資

単位・千

1/四期先予測

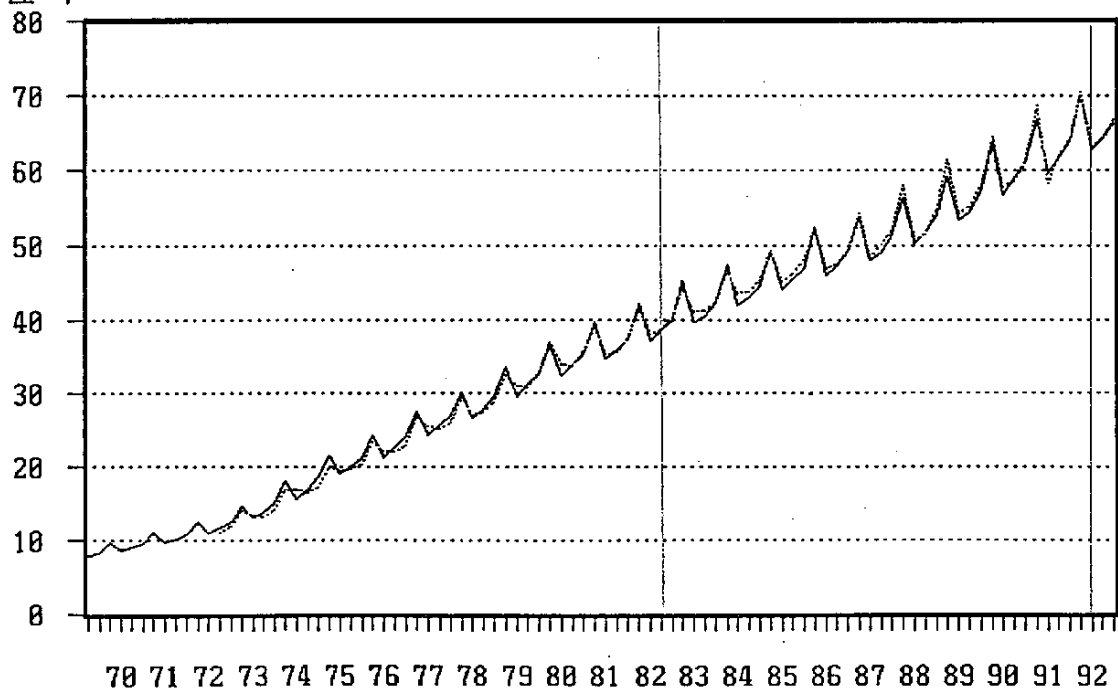


3-8図 比による学習結果Ⅱ

d. 在庫

単位・千

1/四期先予測

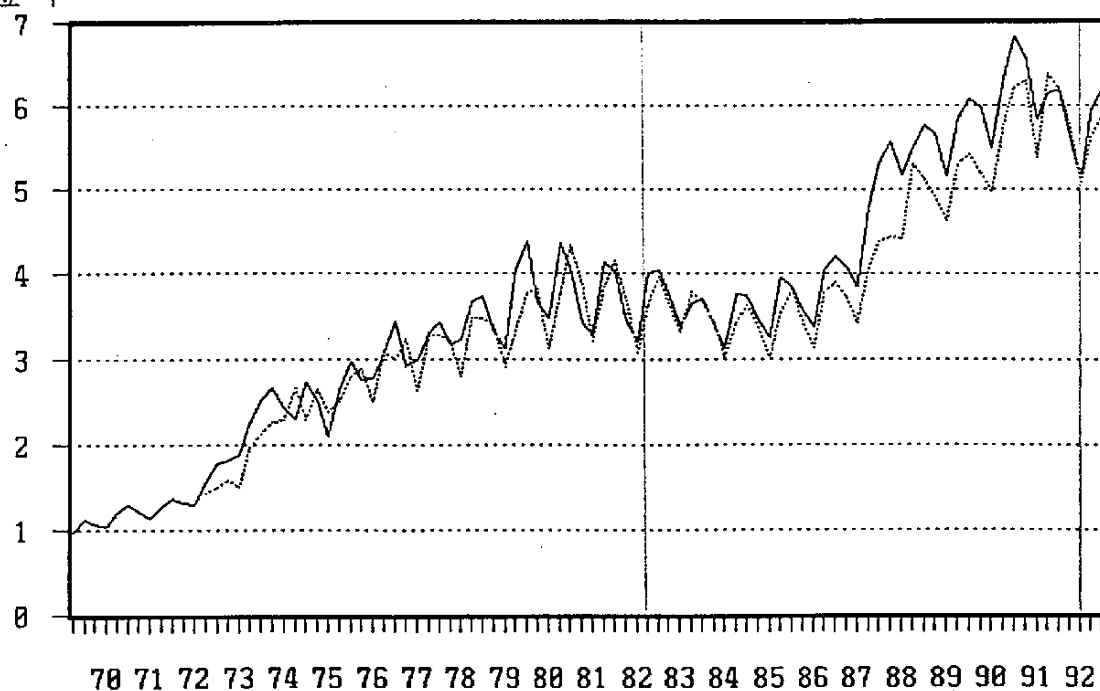


3-9図 比による学習結果Ⅲ

a. 民間最終消費支出

単位・千

1/四期先予測

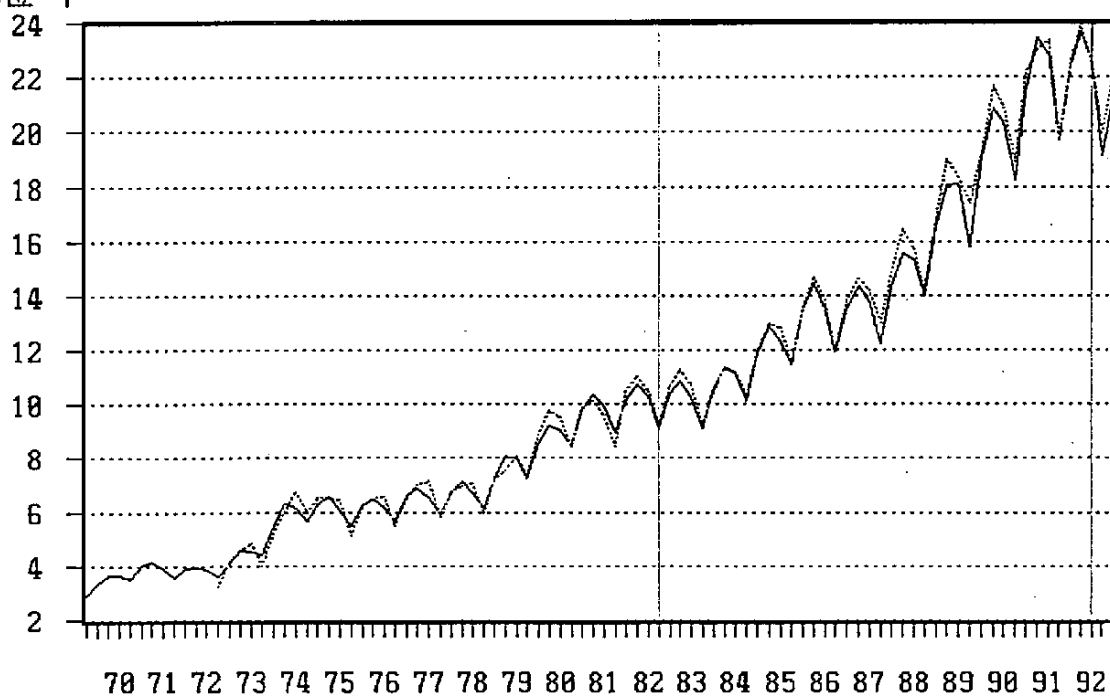


3-9図 比による学習結果Ⅲ

b. 民間住宅投資

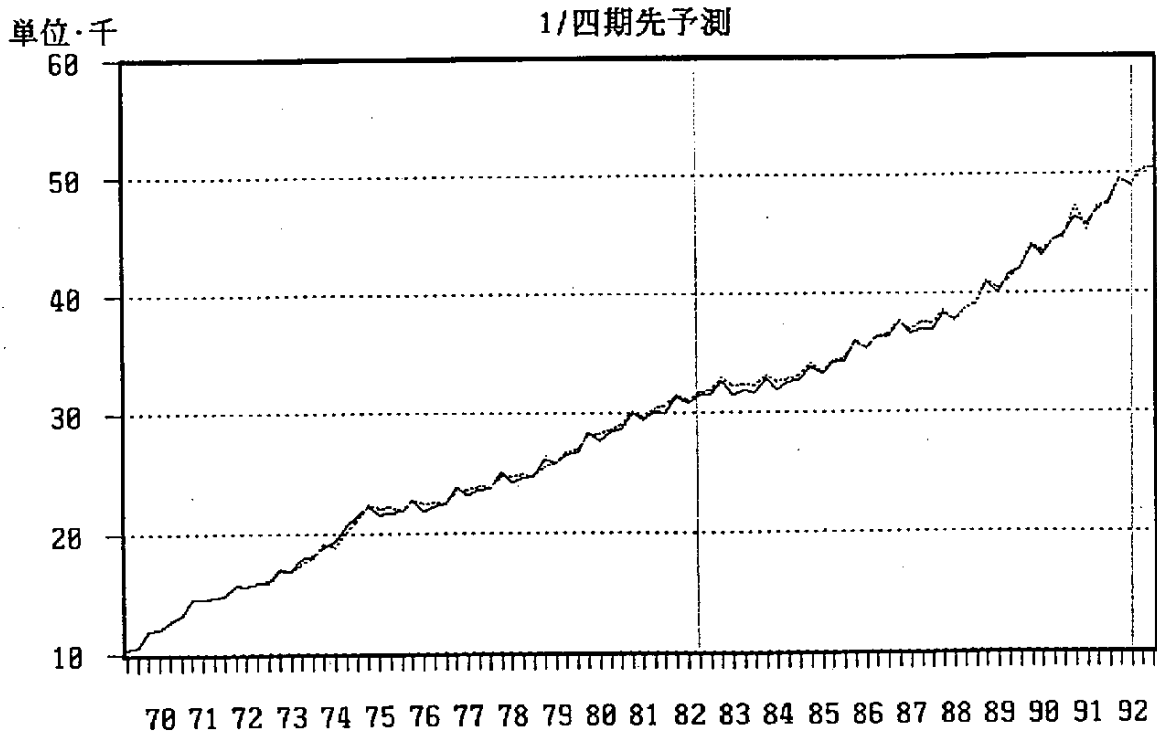
単位・千

1/四期先予測



3-9図 比による学習結果Ⅲ

c. 民間企業設備投資



3-9図 比による学習結果Ⅲ

d. 在庫

c. 結論

以上の結果から、ニューラルネットによる経済予測に用いるデータは、前期比等を用いたものがよく、指数を用いたものはその使用方法に工夫を行う必要があることがわかった。また、学習期間については長ければ良いというものではないことがわかった。

③ 学習時期による比較

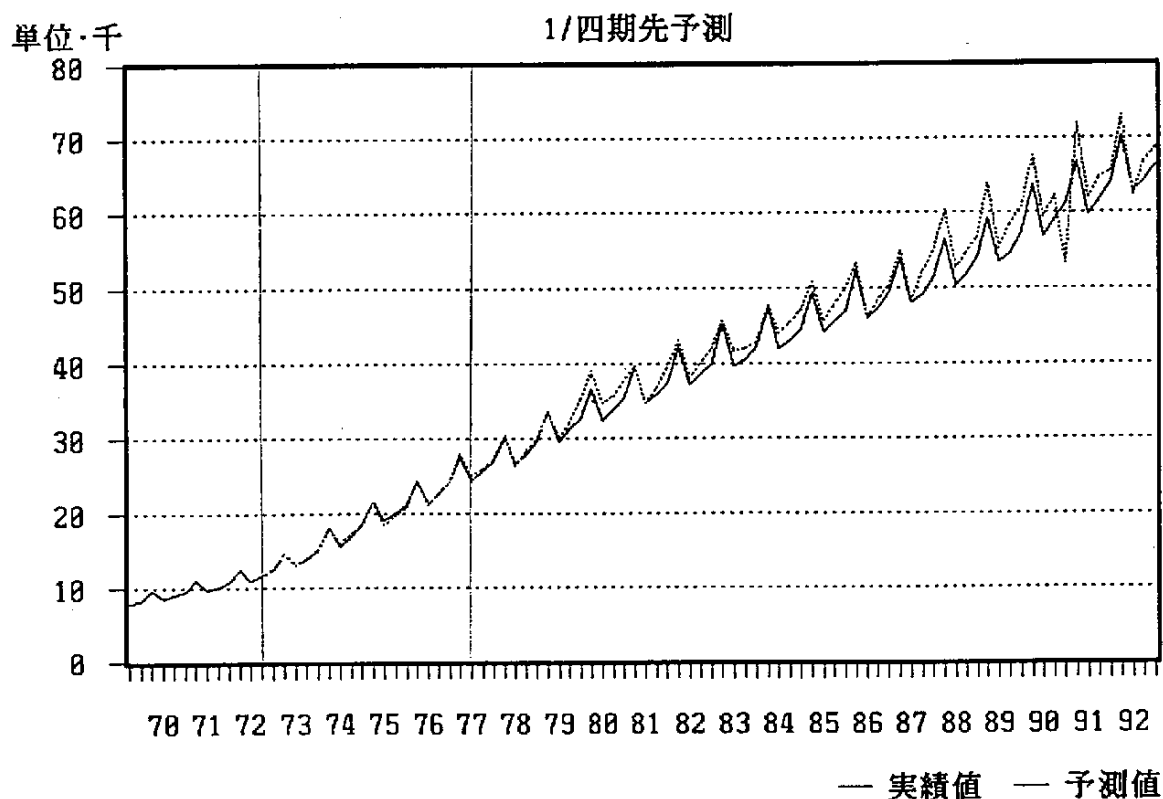
今回は、最適な学習時期を見つけることを目的に比較を行った。用いたデータ、ニューラルネットの仕様、及び学習回数は前節と同じである。

次に、1972年CQ 2から1992年CQ 1までの20年間で5年毎に分割し学習データを作成した。1972CQ 2から1977CQ 1までの5年間（20期）学習させた結果を3-10図に、1977CQ 2から1982CQ 1までの5年間（20期）学習させた結果を3-11図に、1982CQ 2から1987CQ 1までの5年間（20期）学習させた結果を3-12図に、1987CQ 2から1992CQ 1までの5年間（20期）の学習結果を3-13図に示す。なお、予測値は見やすいよう連続線としている。

以上4とおりの結果から、学習時期については、最新の5年間で学習させた結果が一番よかった。

また、3-14図に民間消費支出の四半期ごとの伸び率をシグモイド変換した値をグラフ化したものを示す。この図から伸び率のパターンは常に一定ではないことがわかる。

以上のことから、学習データは最新のパターンを学習させることが重要であり、また、学習期間があまり短いと短期的な変動に強く影響を受けるため、学習期間は最新のデータで5年程度学習させるのがよいことがわかった。

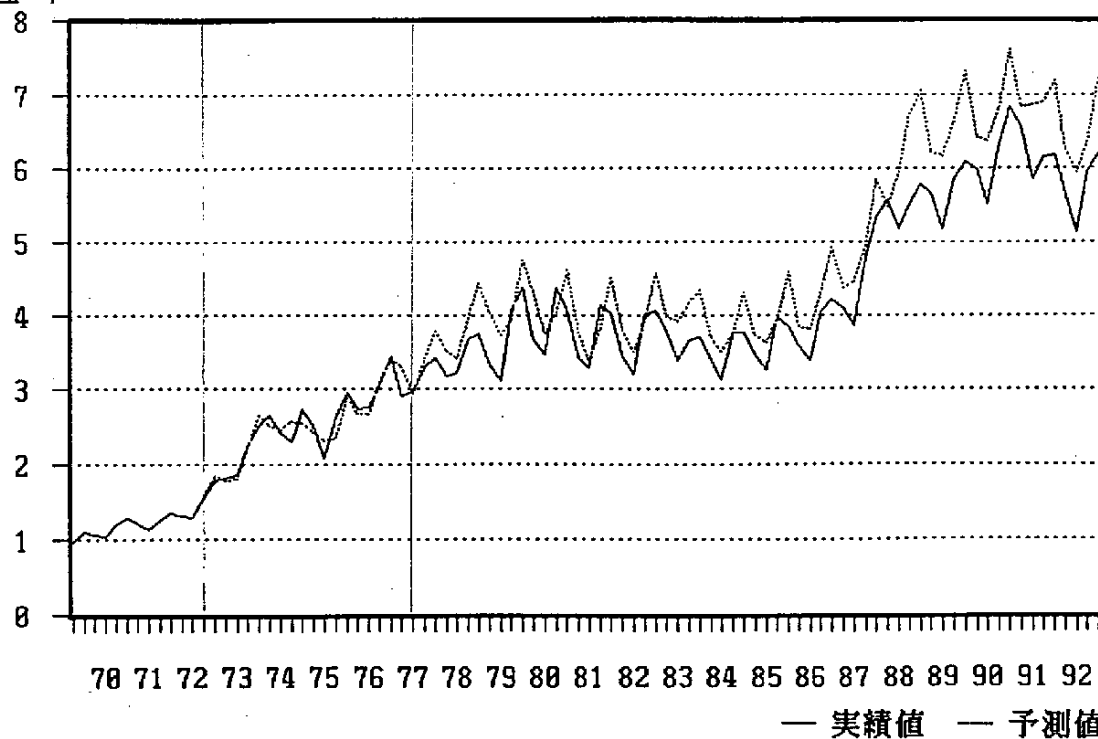


3-10図 学習時期Ⅰによる結果

a. 民間最終消費支出

単位・千

1/四期先予測

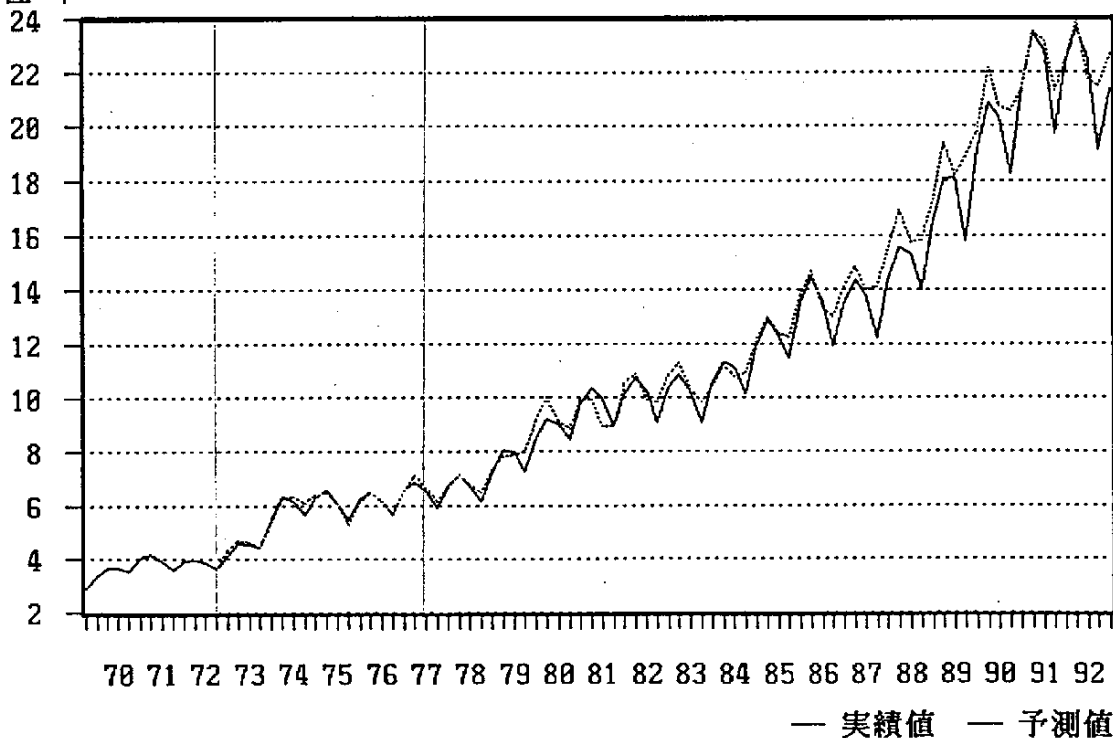


3-10図 学習時期Ⅰによる結果

b. 民間住宅投資

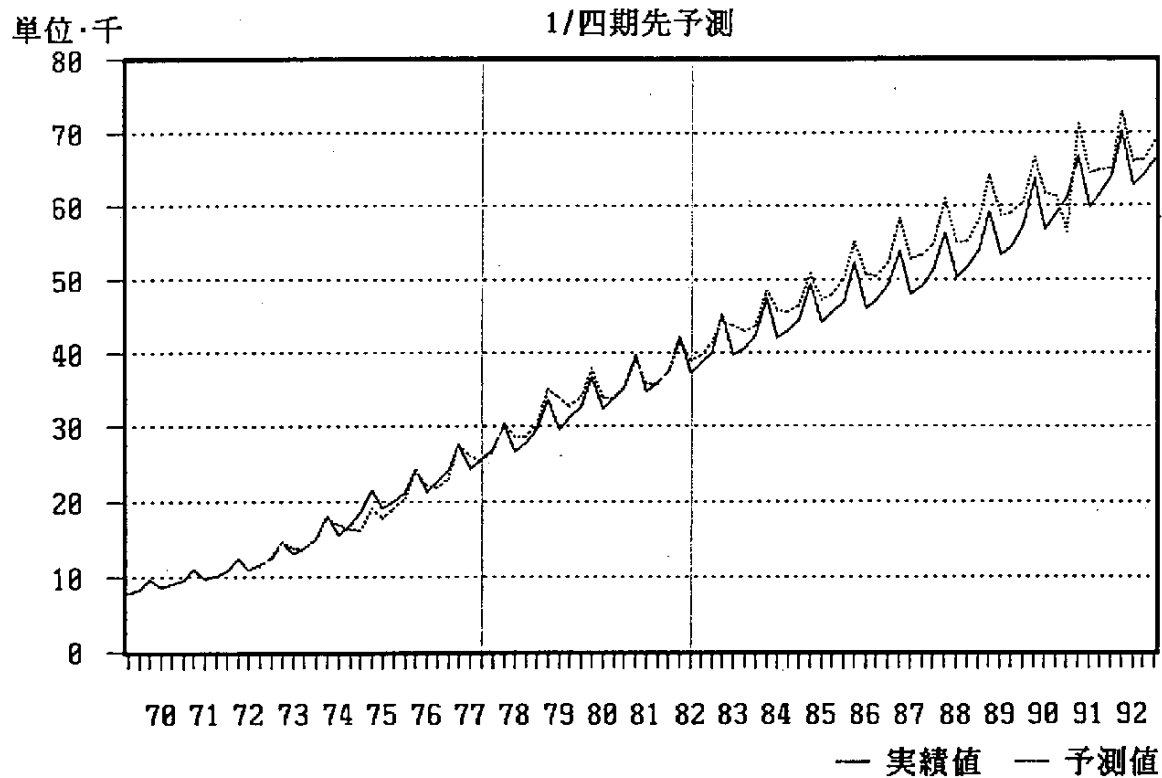
単位・千

1/四期先予測



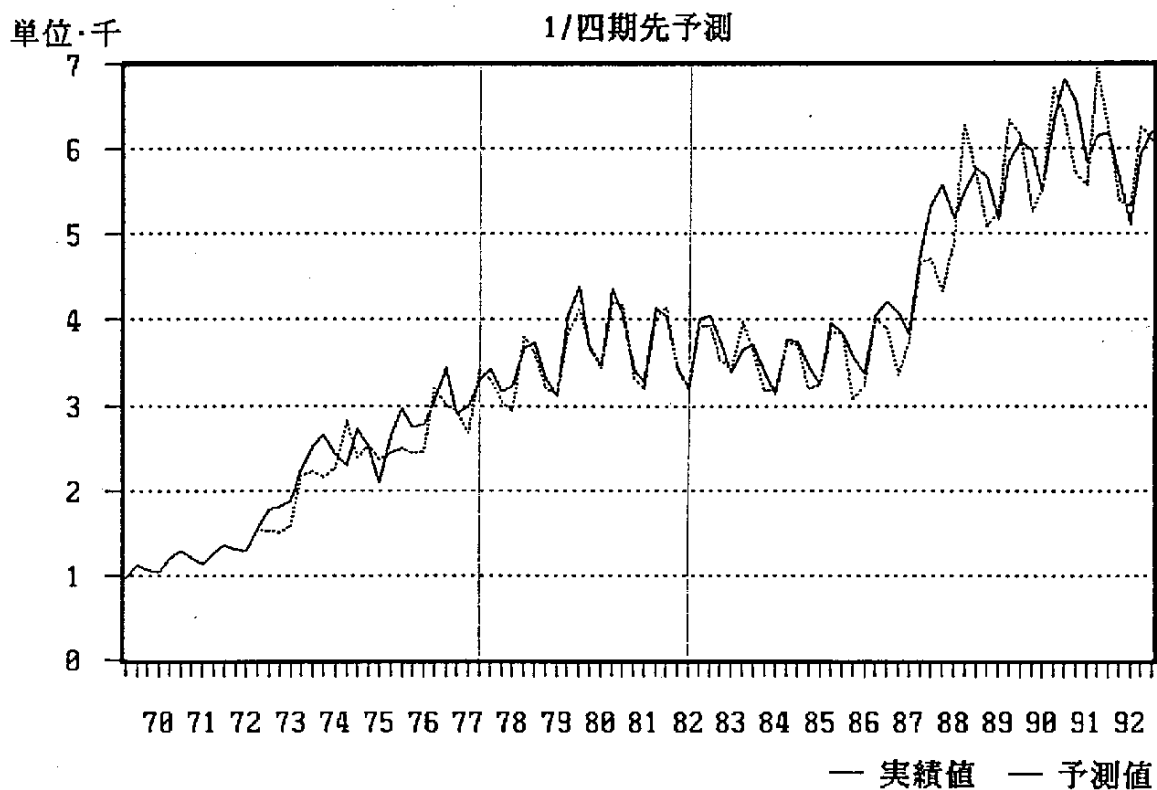
3-10図 学習時期Ⅰによる結果

c. 民間企業設備投資



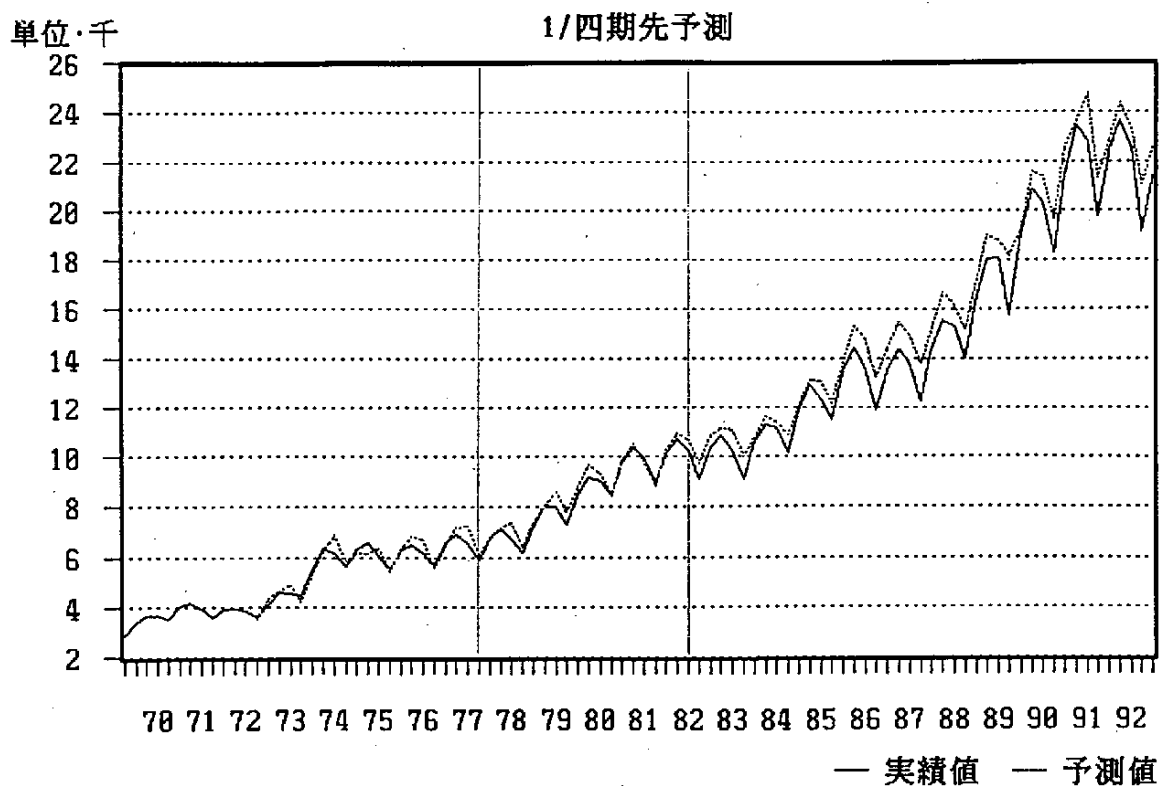
3-11図 学習時期Ⅱによる結果

a. 民間最終消費支出



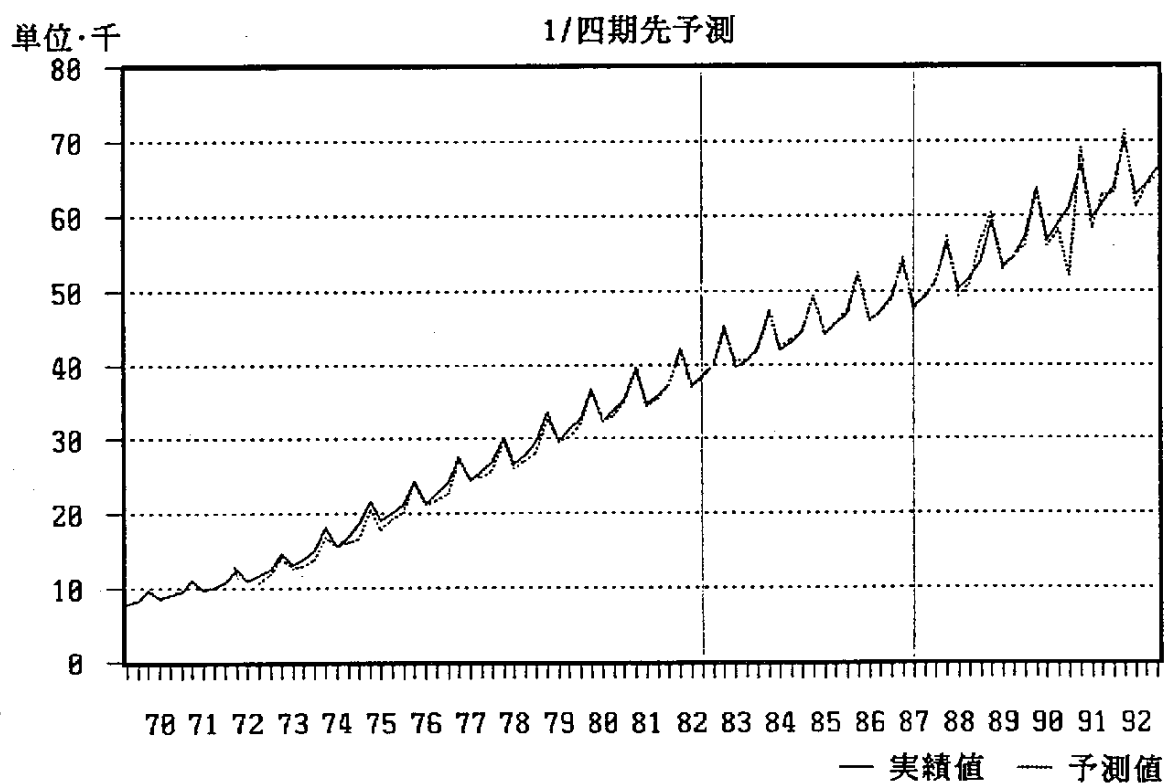
3-11図 学習時期Ⅱによる結果

b. 民間住宅投資



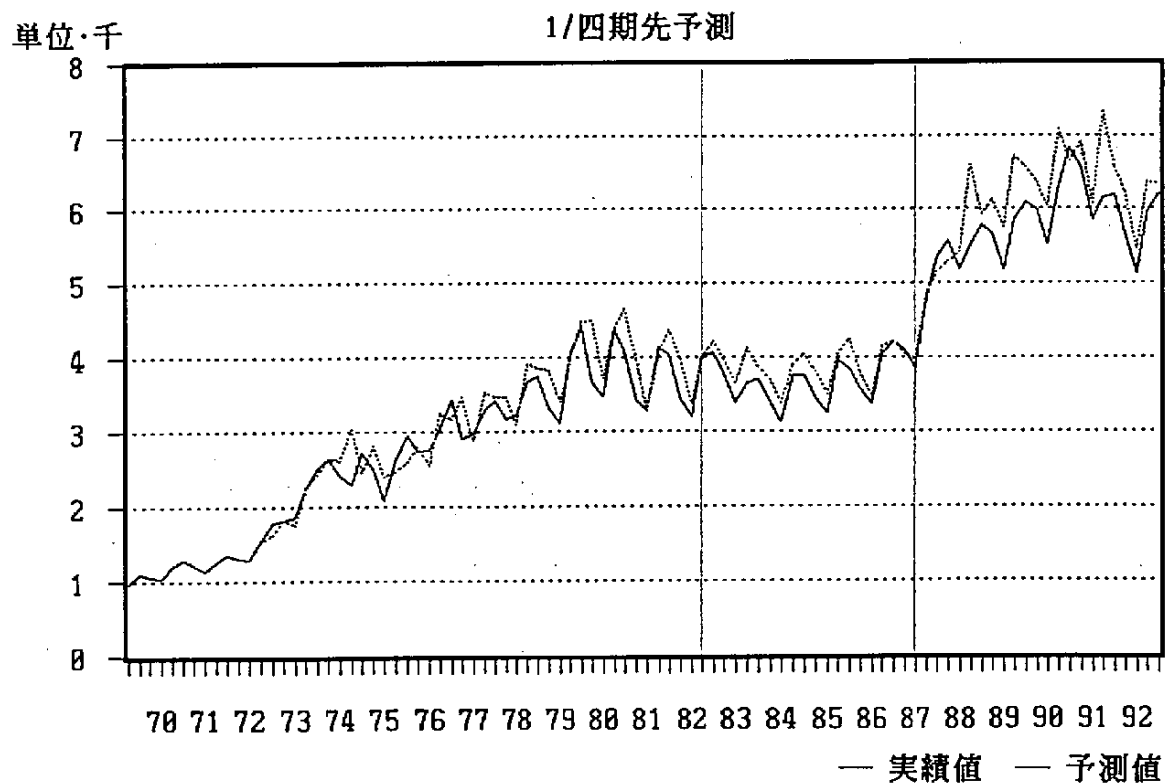
3-11図 学習時期Ⅱによる結果

c. 民間企業設備投資



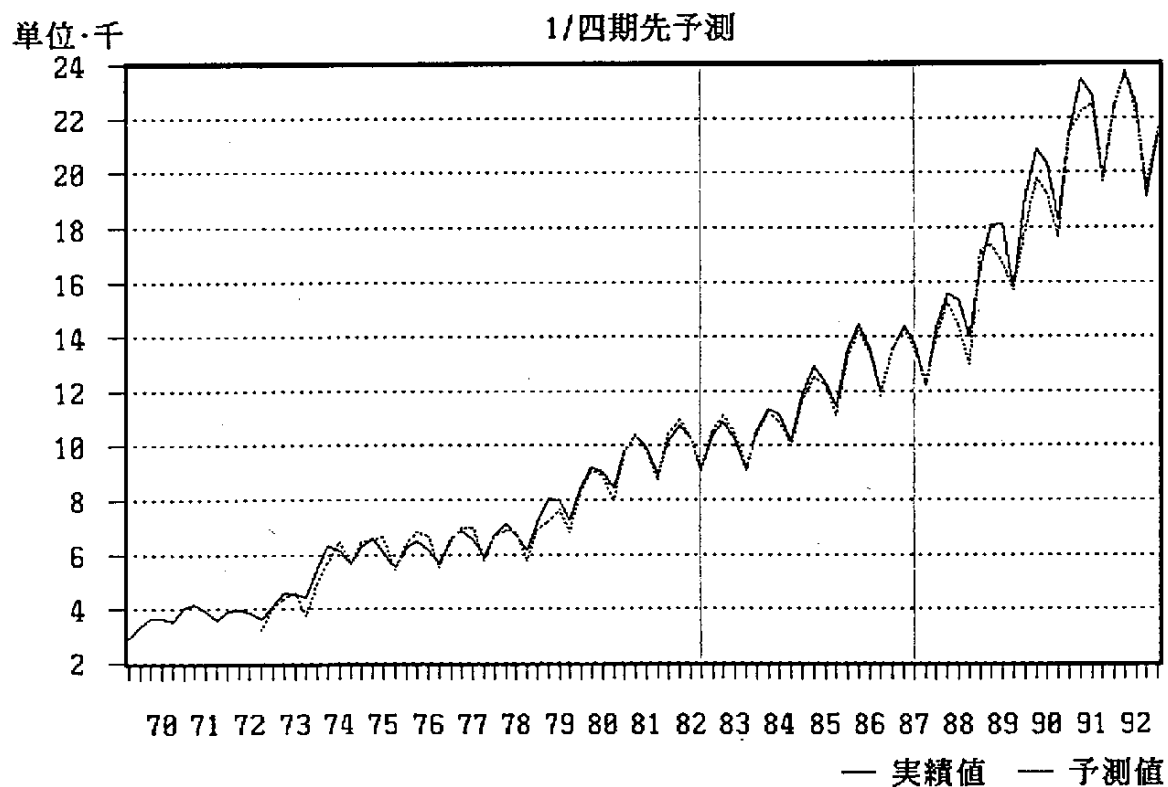
3-12図 学習時期Ⅲによる結果

a. 民間最終消費支出



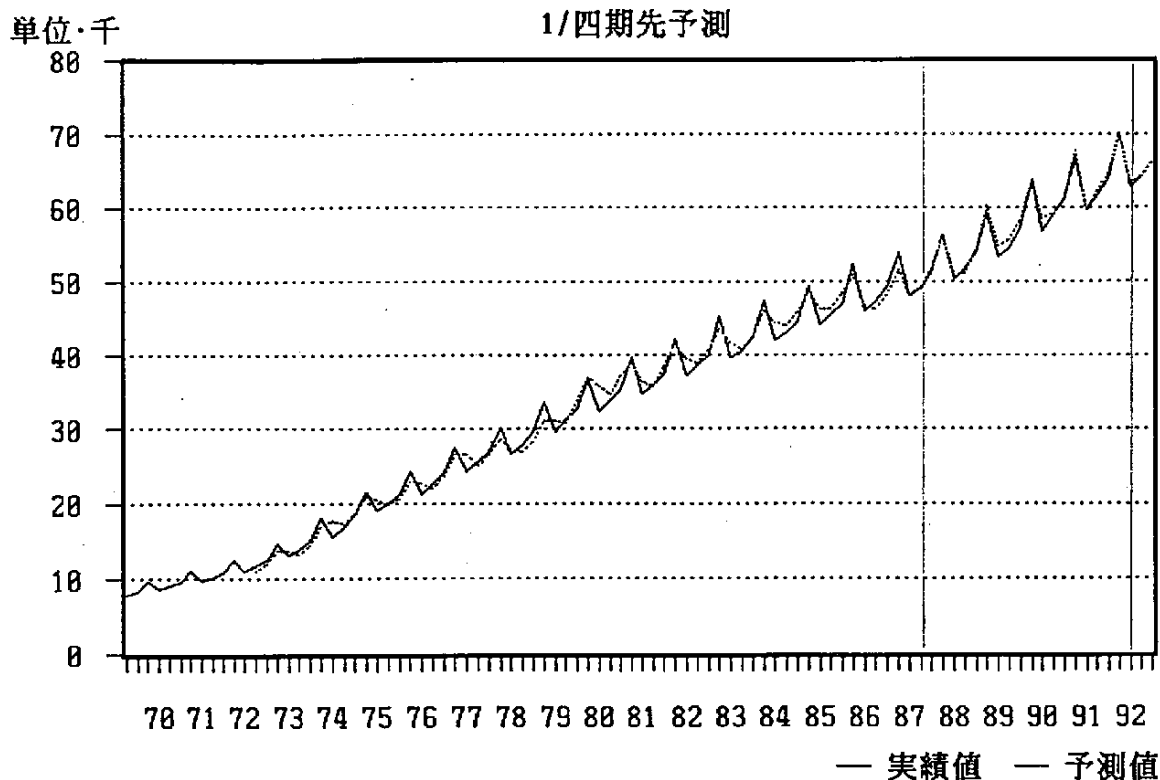
3-12図 学習時期Ⅲによる結果

b. 民間住宅投資



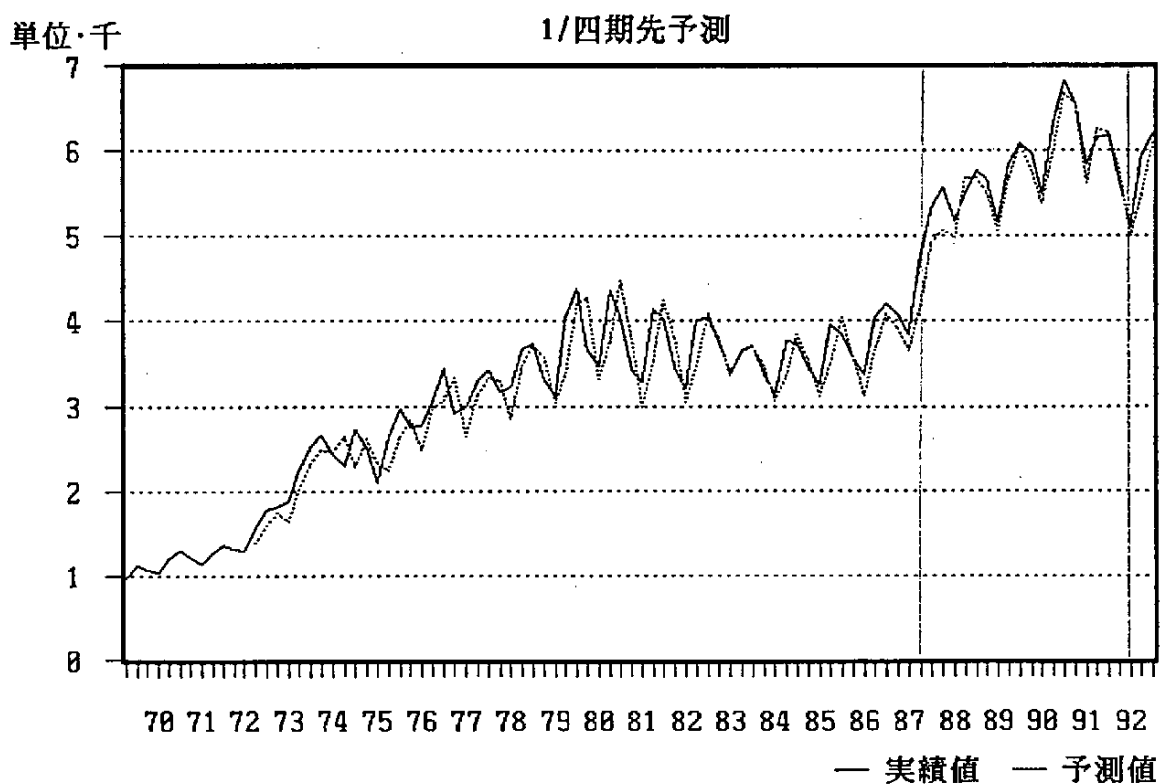
3-12図 学習時期Ⅲによる結果

c. 民間企業設備投資



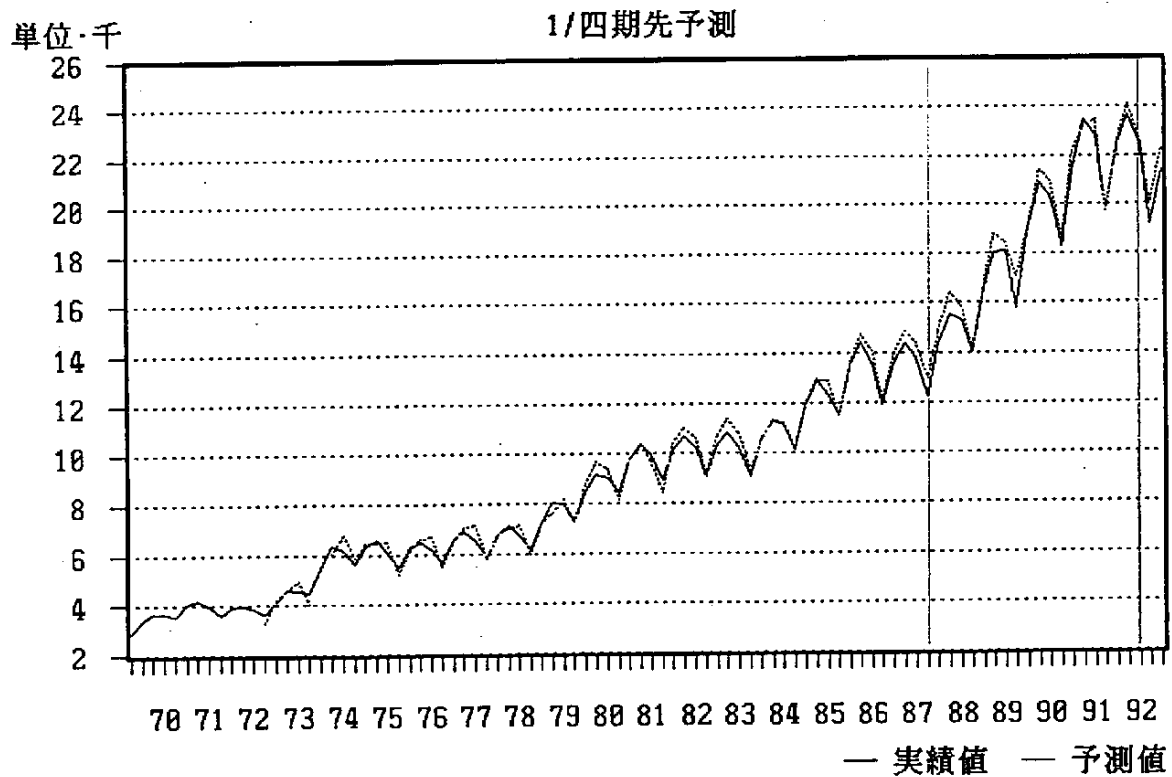
3-13図 学習時期Ⅳによる結果

a. 民間最終消費支出



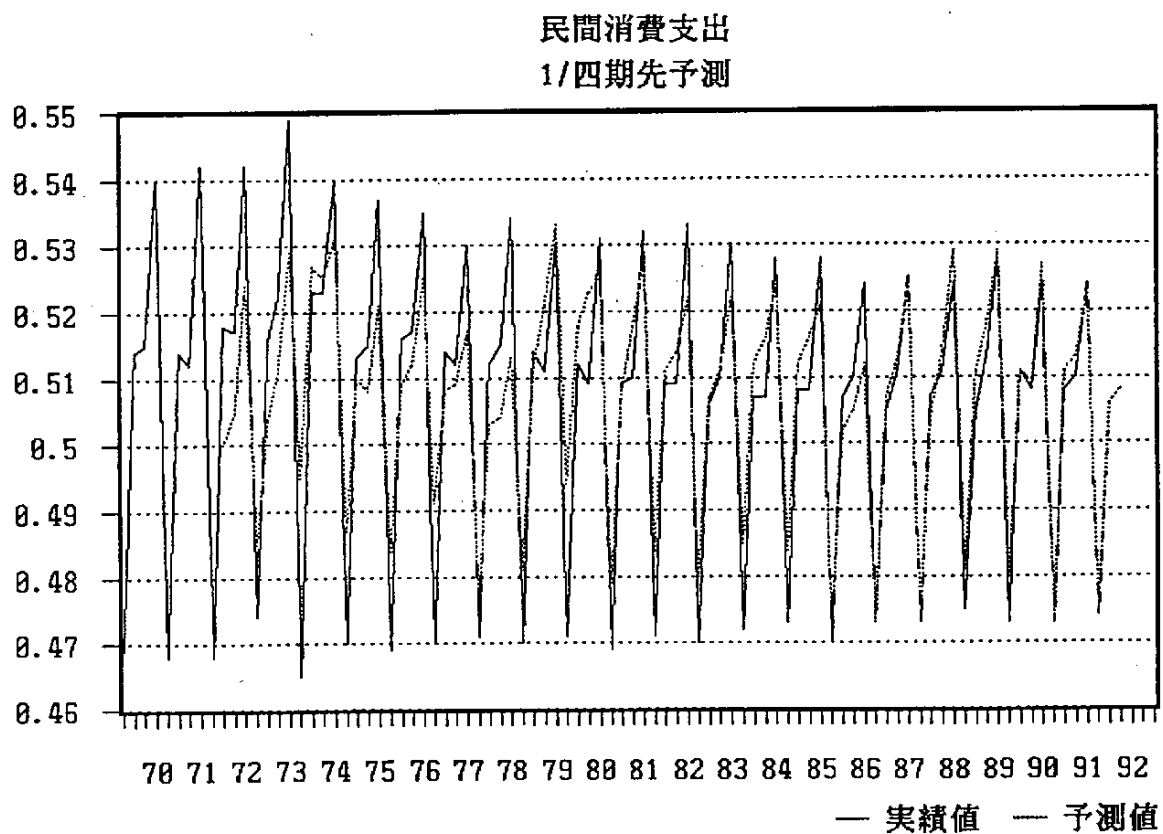
3-13図 学習時期Ⅳによる結果

b. 民間住宅投資



3-13図 学習時期Ⅳによる結果

c. 民間企業設備投資

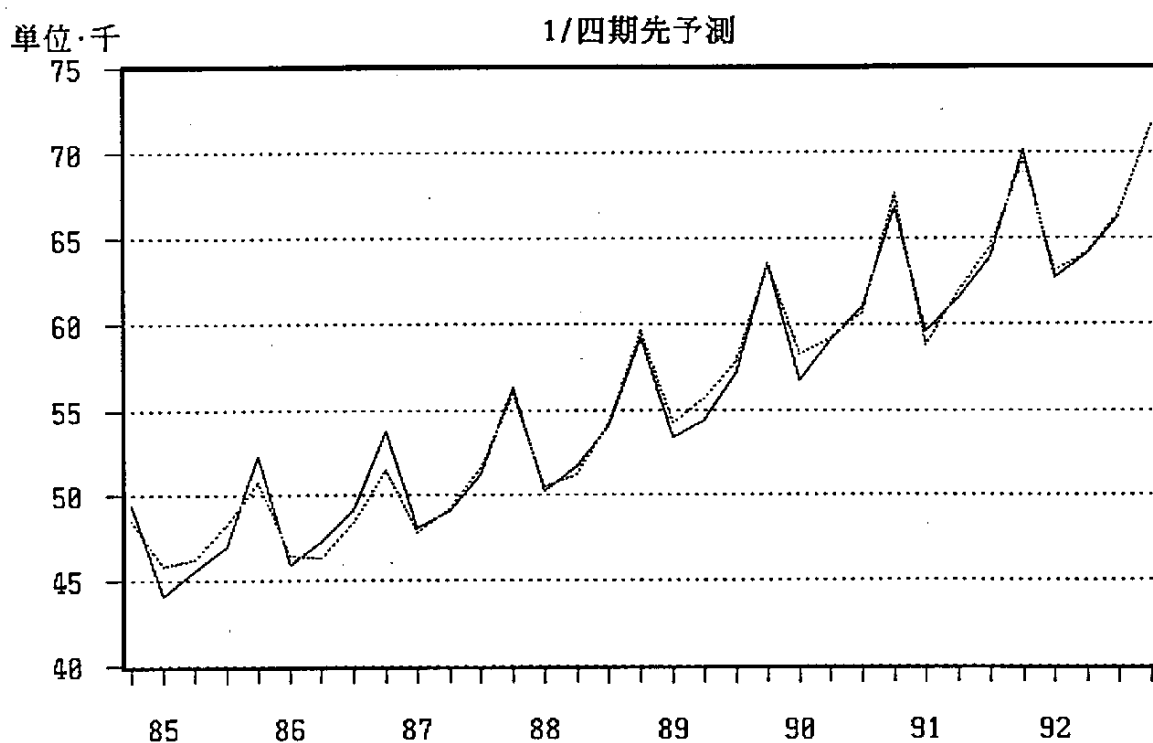


3-14図 四半期伸び率のシグモイド変換値

④ 中間層のニューロン数の変化に伴う予測への影響

中間層のニューロン数の決定は、経験により求めざるを得ない。今回1987CQ 2 から1992CQ 1 までの5年間(20期)の学習データを用い、入力層に9指標×3期=27、出力層に4指標×1期=4のニューラルネットで中間層のニューロン数を10、11、12、15、20、30に変化させて学習を行った。その結果を3-15図～3-20図に示す。

以上から、中間層についてはその差による大きな違いは出なかったが、中間層10、11、12の結果がよいことがわかった。

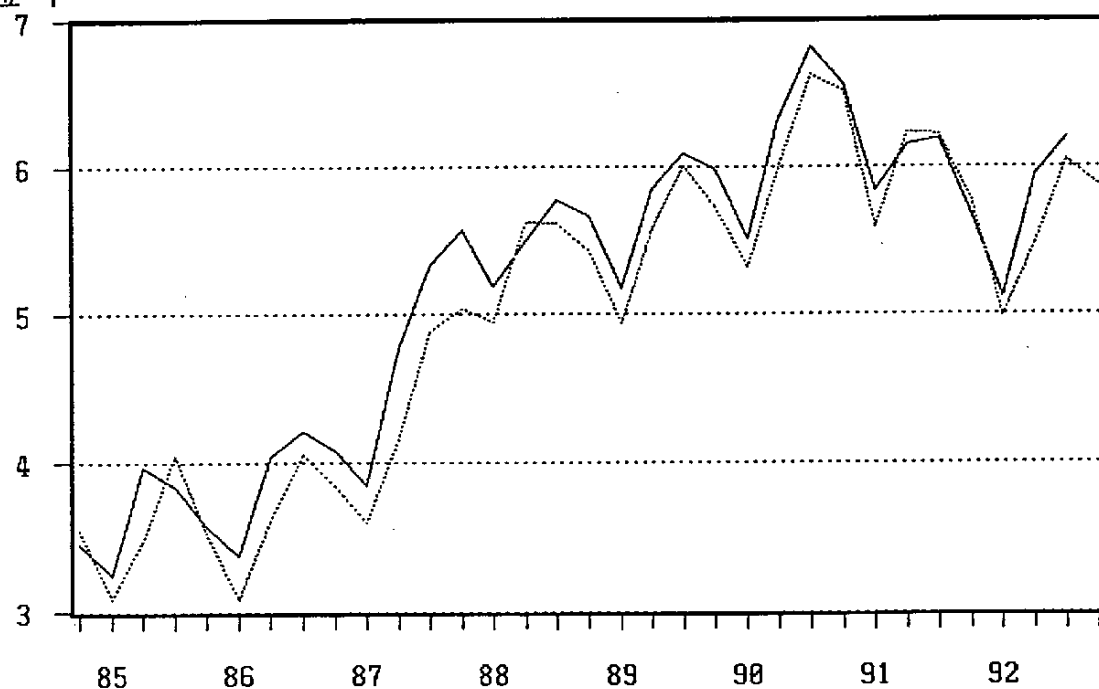


3-15図 ニューロン数10による結果

a. 民間最終消費支出

単位・千

1/四期先予測

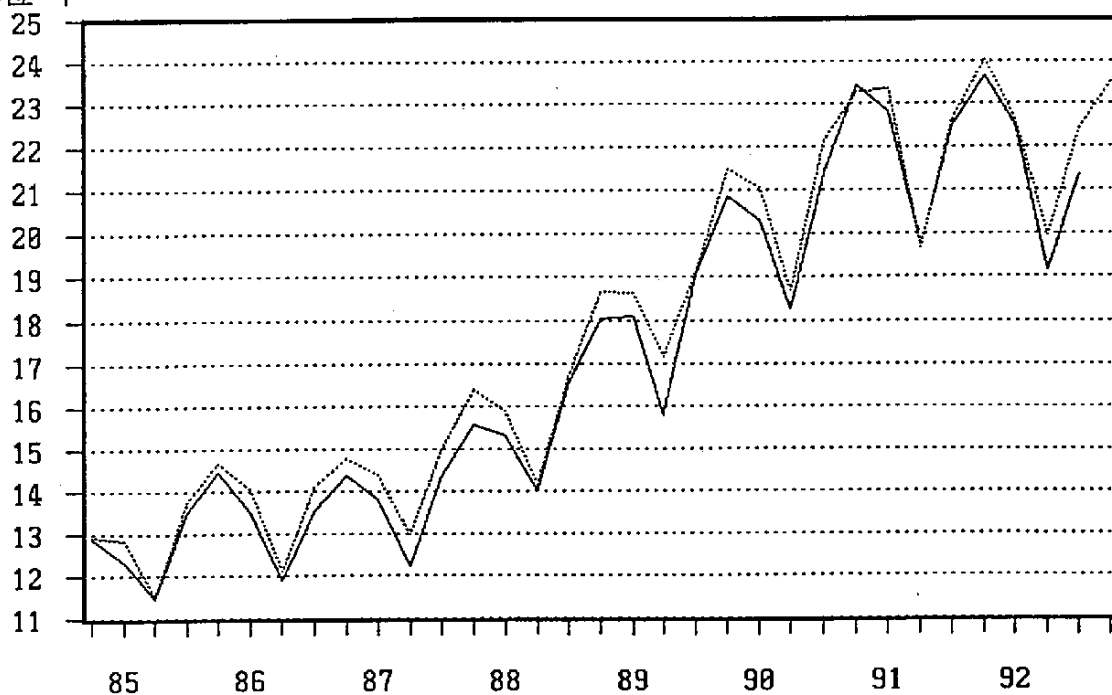


3-15図 ニューロン数10による結果

b. 民間住宅投資

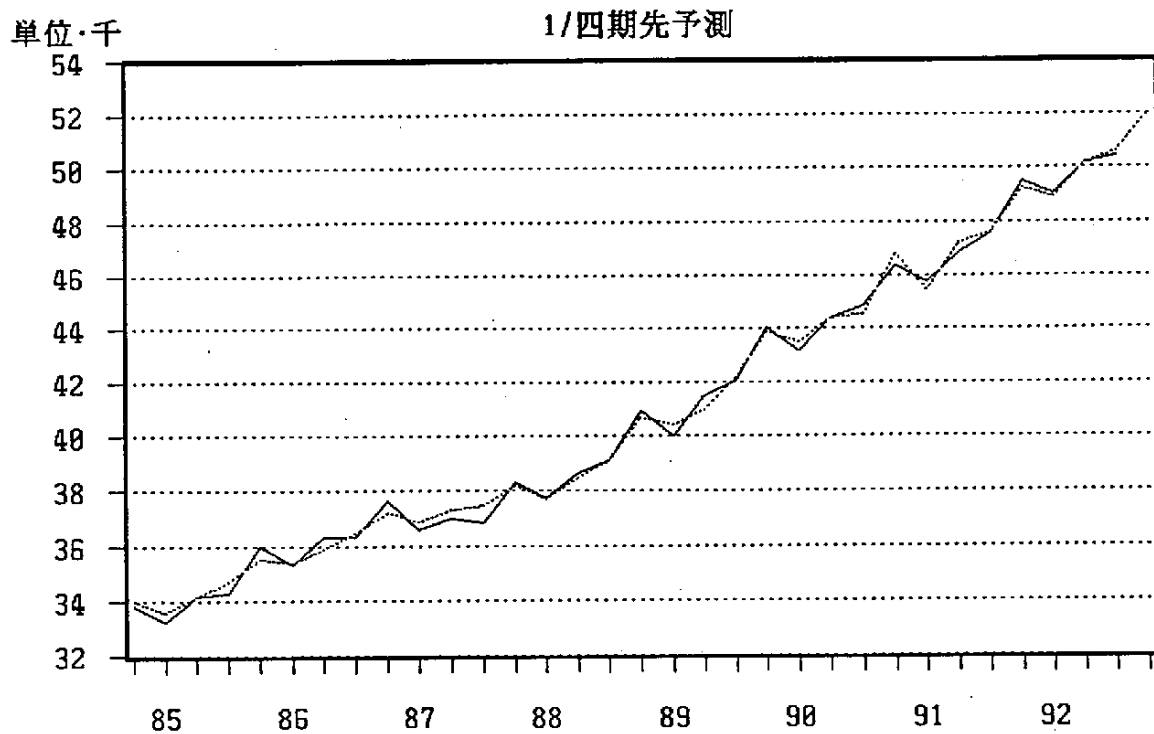
単位・千

1/四期先予測

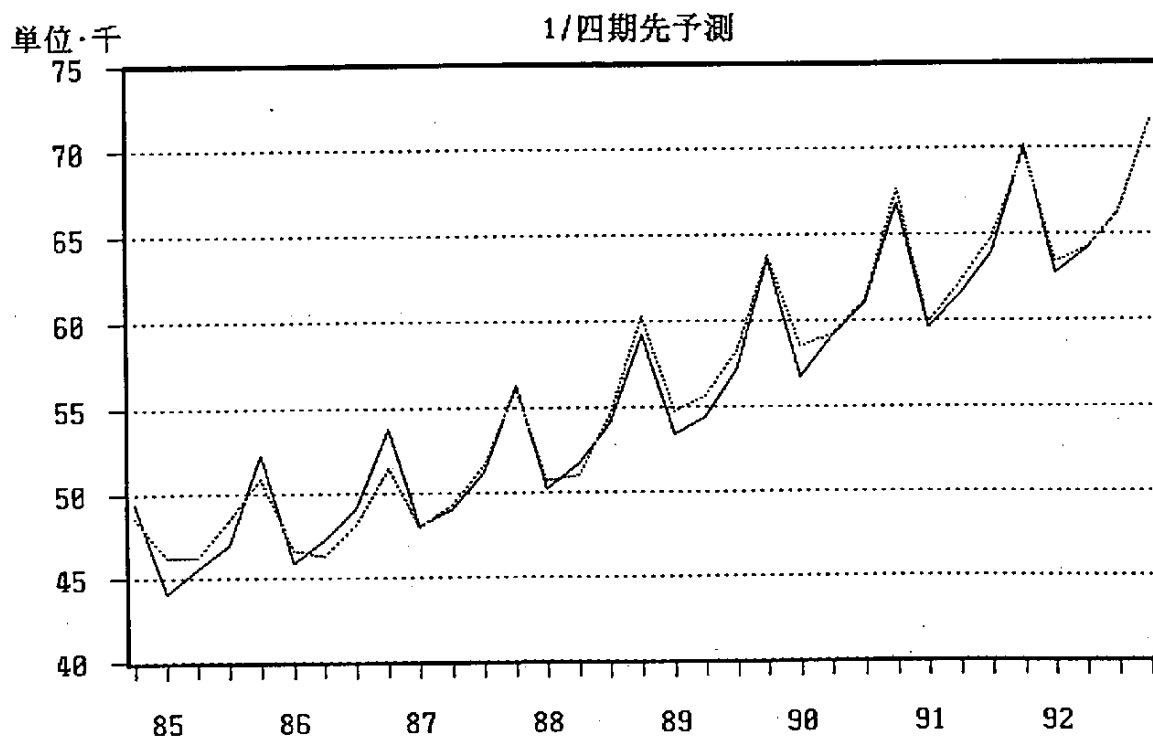


3-15図 ニューロン数10による結果

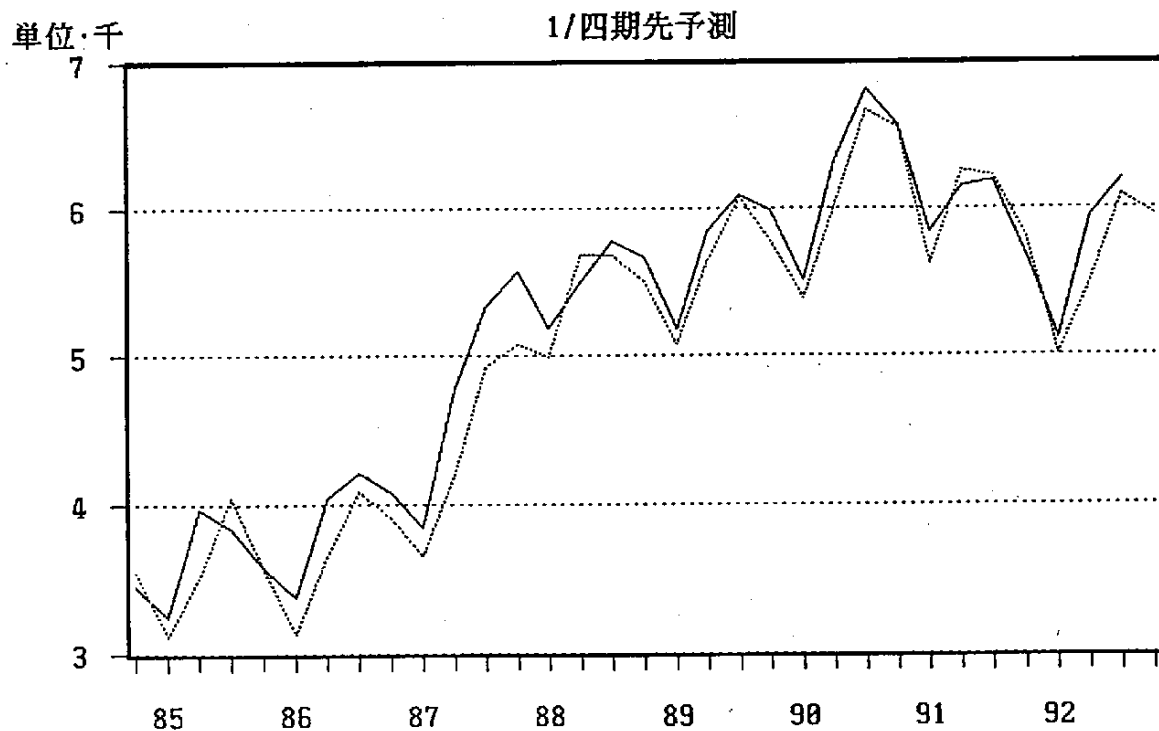
c. 民間企業設備投資



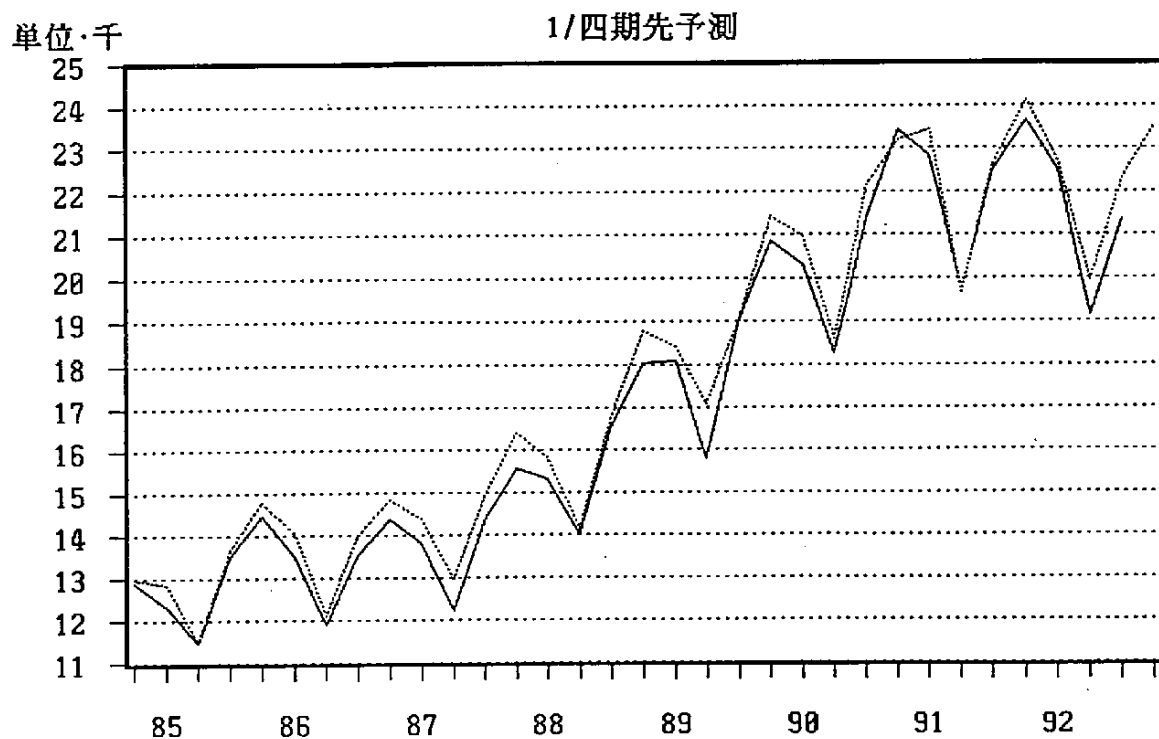
3-15図 ニューロン数10による結果
d. 在庫



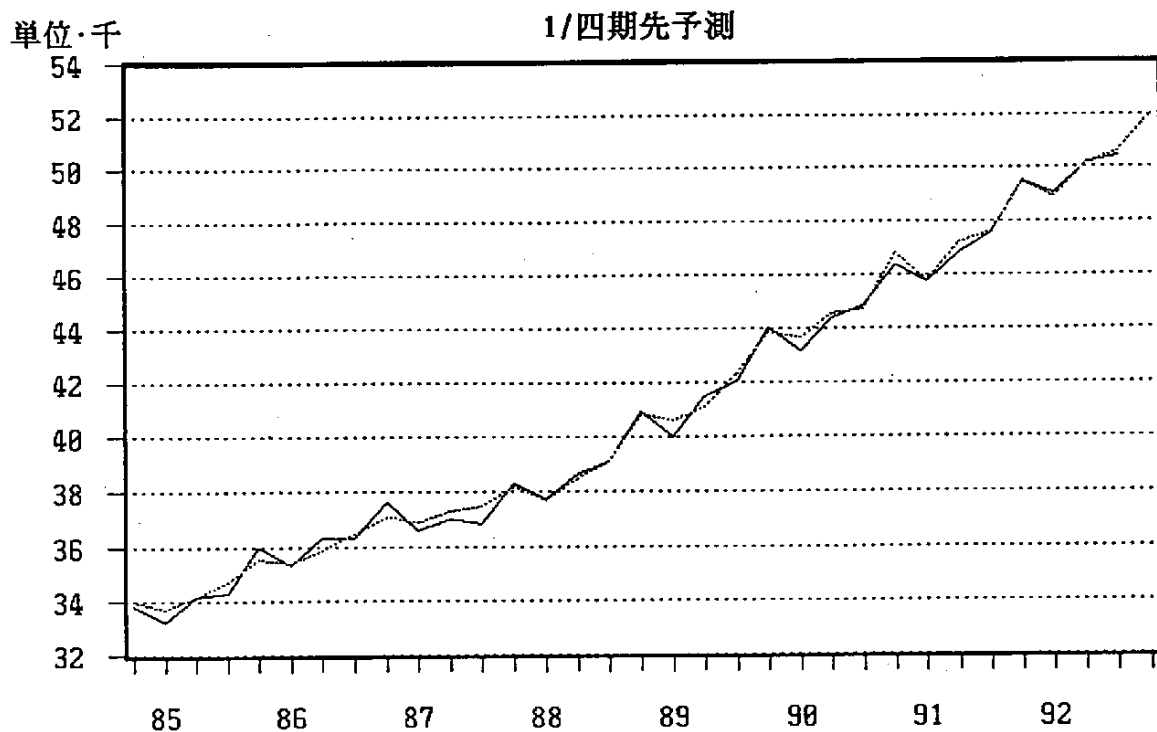
3-16図 ニューロン数11による結果
a. 民間最終消費支出



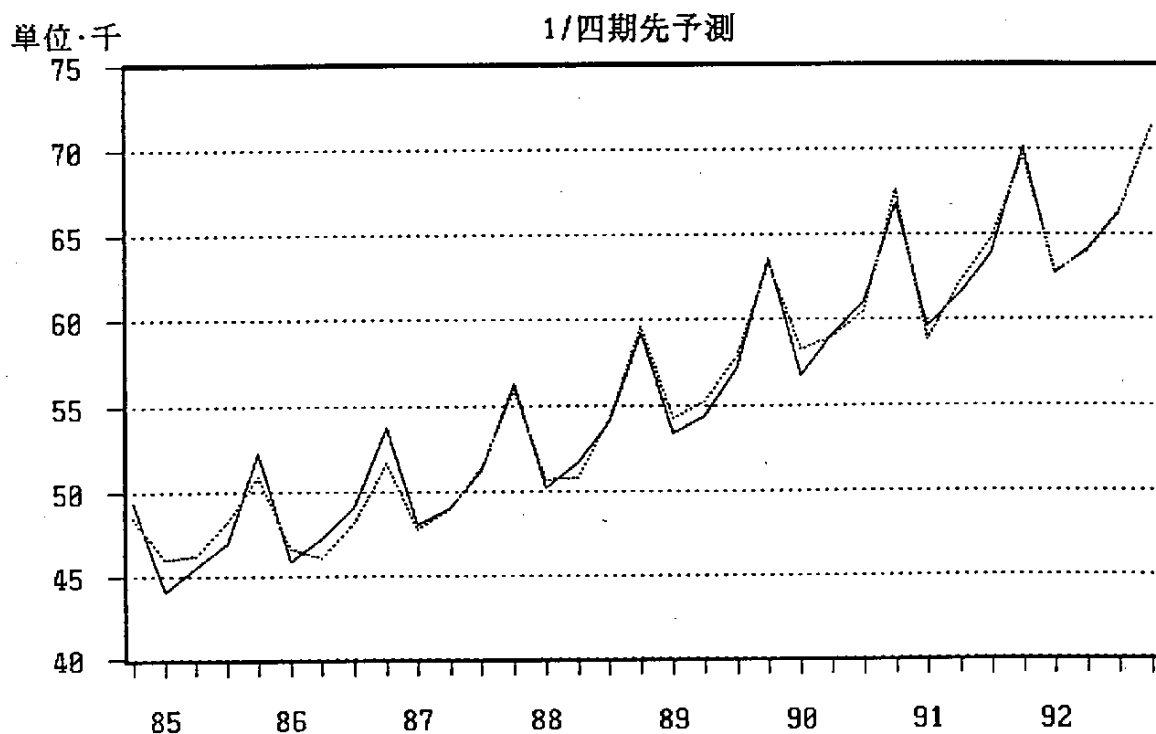
3-16図 ニューロン数11による結果
b. 民間住宅投資



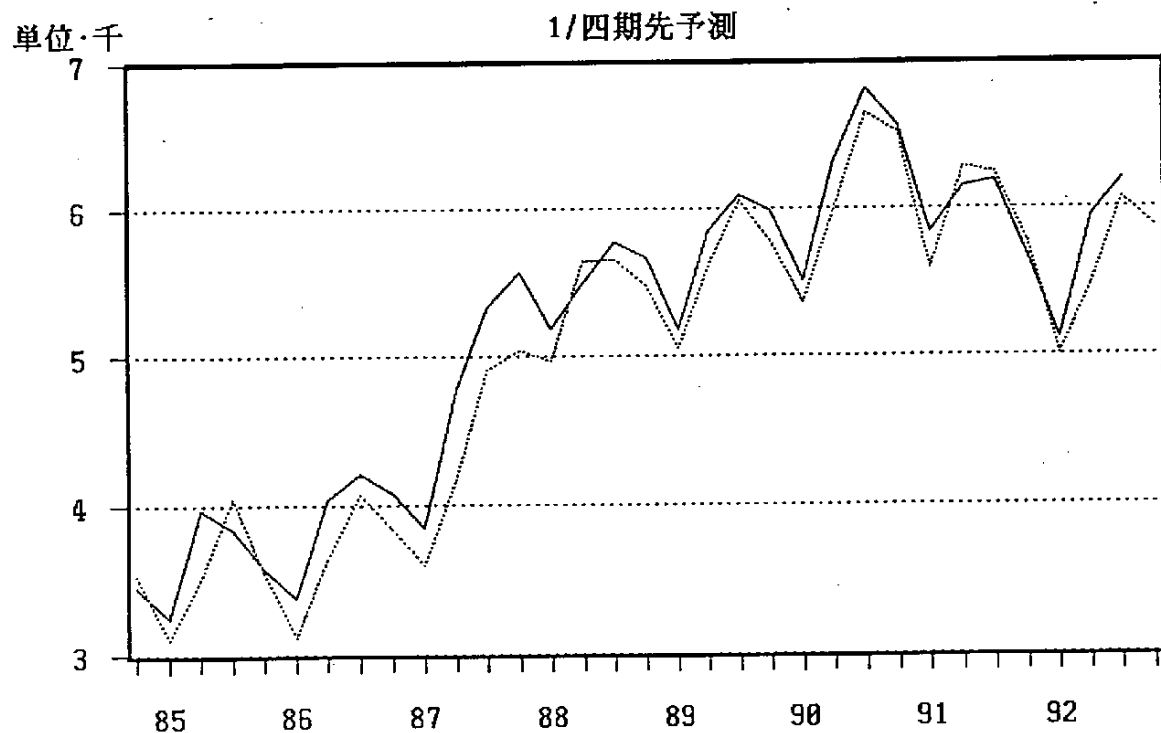
3-16図 ニューロン数11による結果
c. 民間企業設備投資



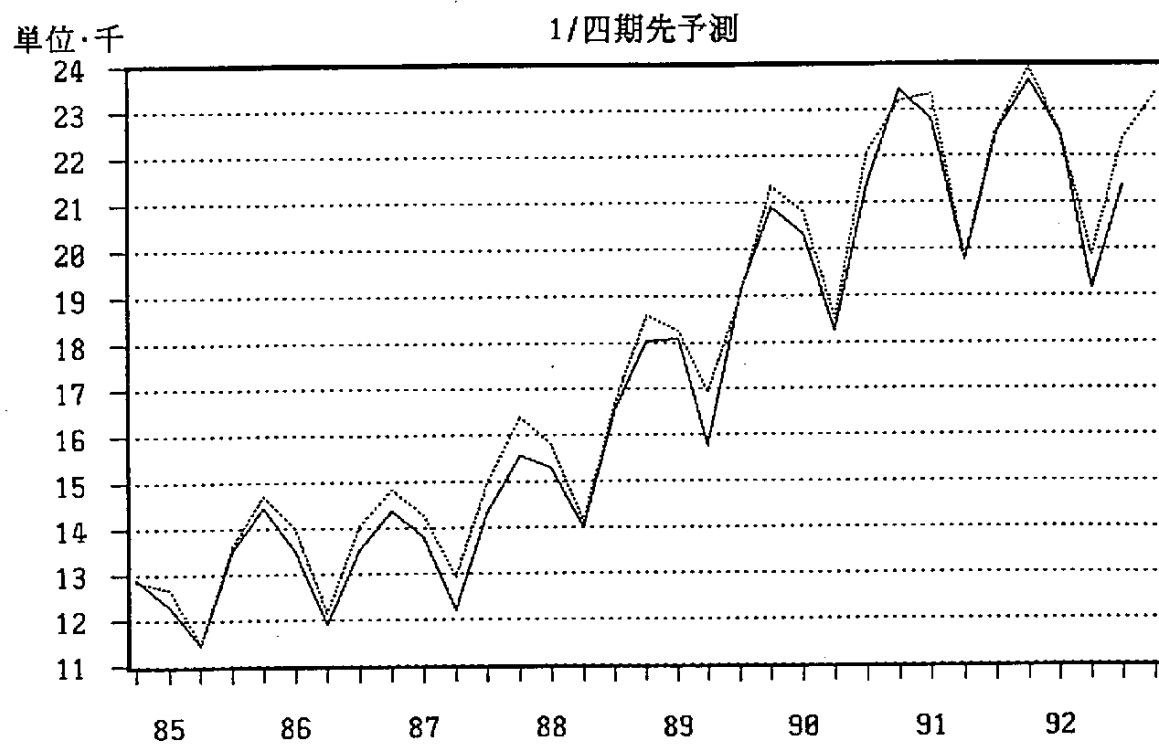
3-16図 ニューロン数11による結果
d. 在庫



3-17図 ニューロン数12による結果
a. 民間最終消費支出



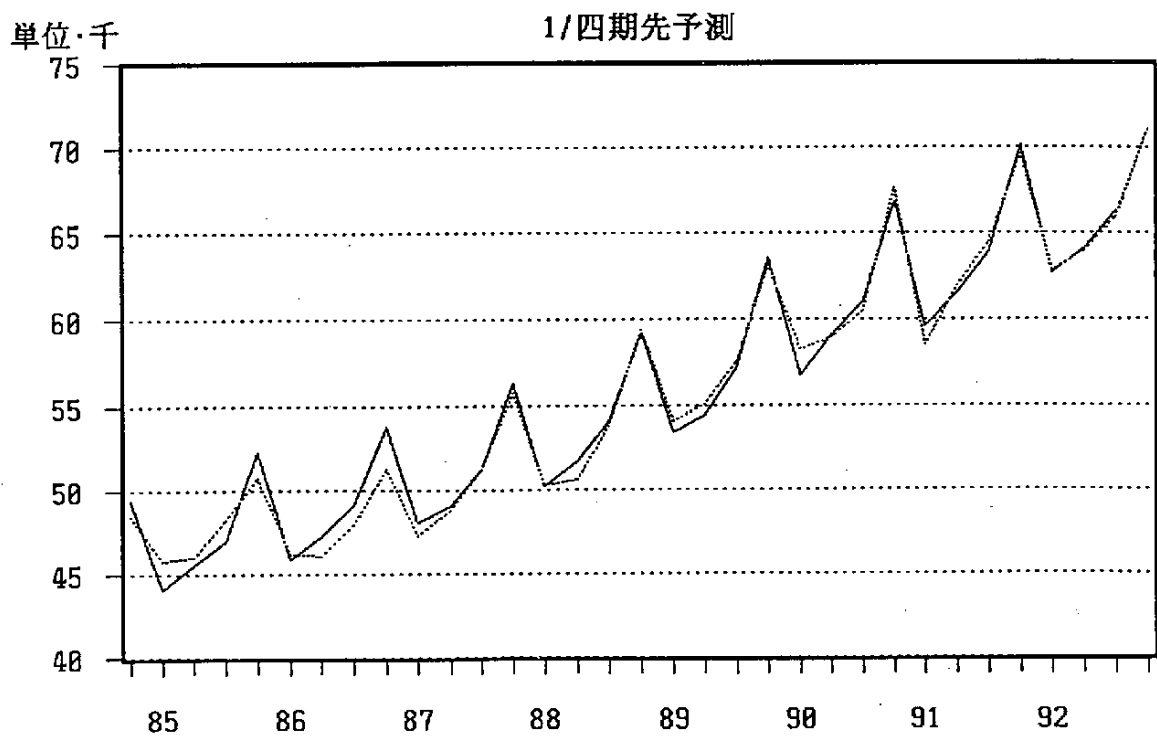
3-17図 ニューロン数12による結果
b. 民間住宅投資



3-17図 ニューロン数12による結果
c. 民間企業設備投資



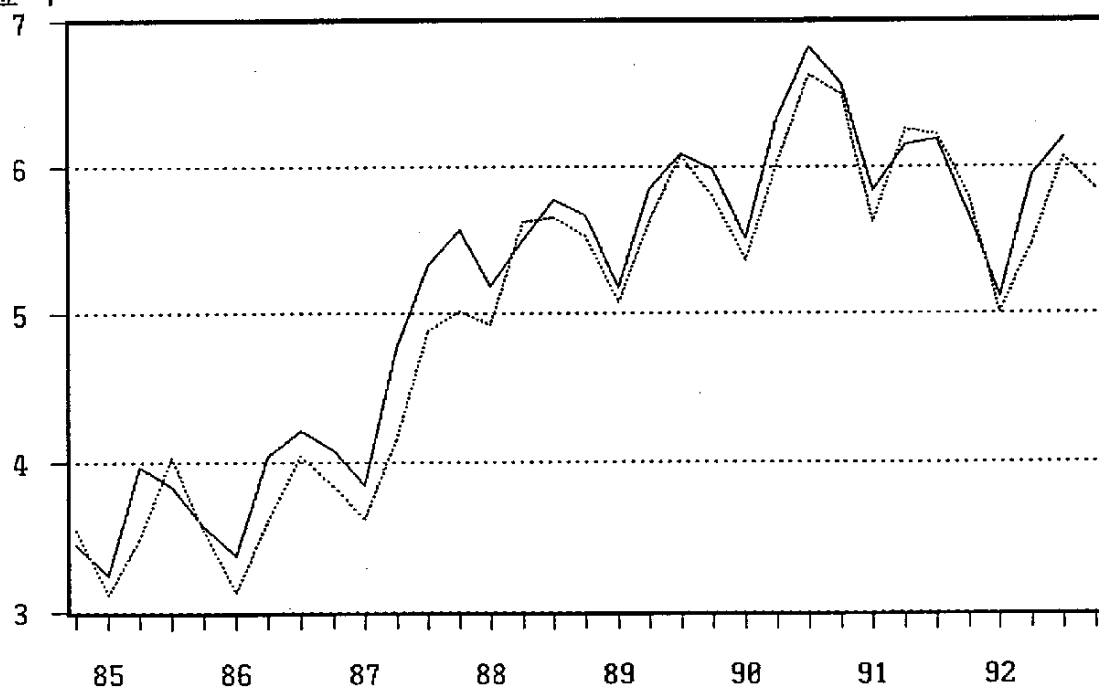
3-17図 ニューロン数12による結果
d. 在庫



3-18図 ニューロン数15による結果
a. 民間最終消費支出

単位・千

1/四期先予測

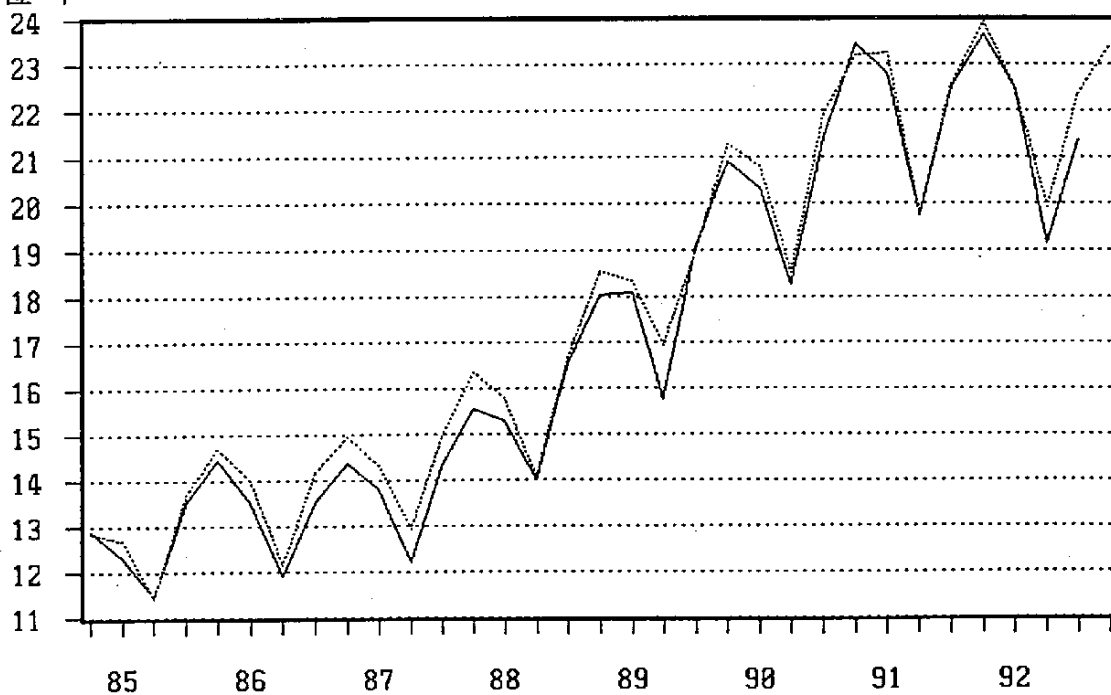


3-18図 ニューロン数15による結果

b. 民間住宅投資

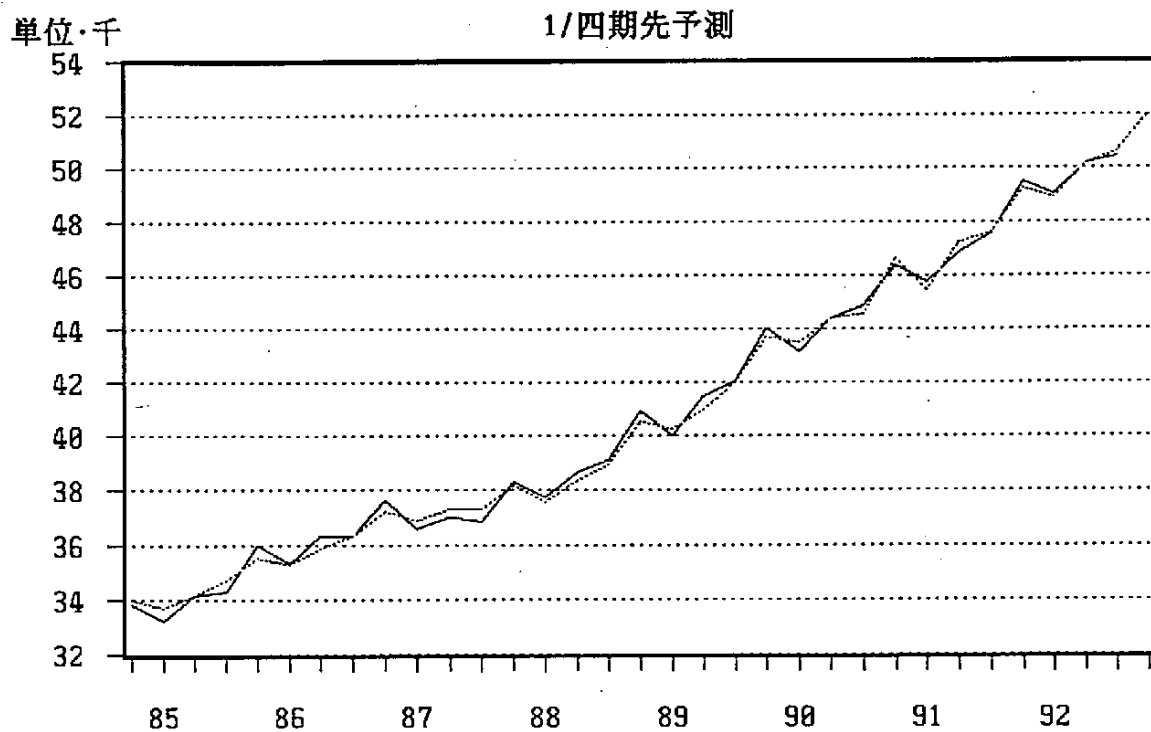
単位・千

1/四期先予測

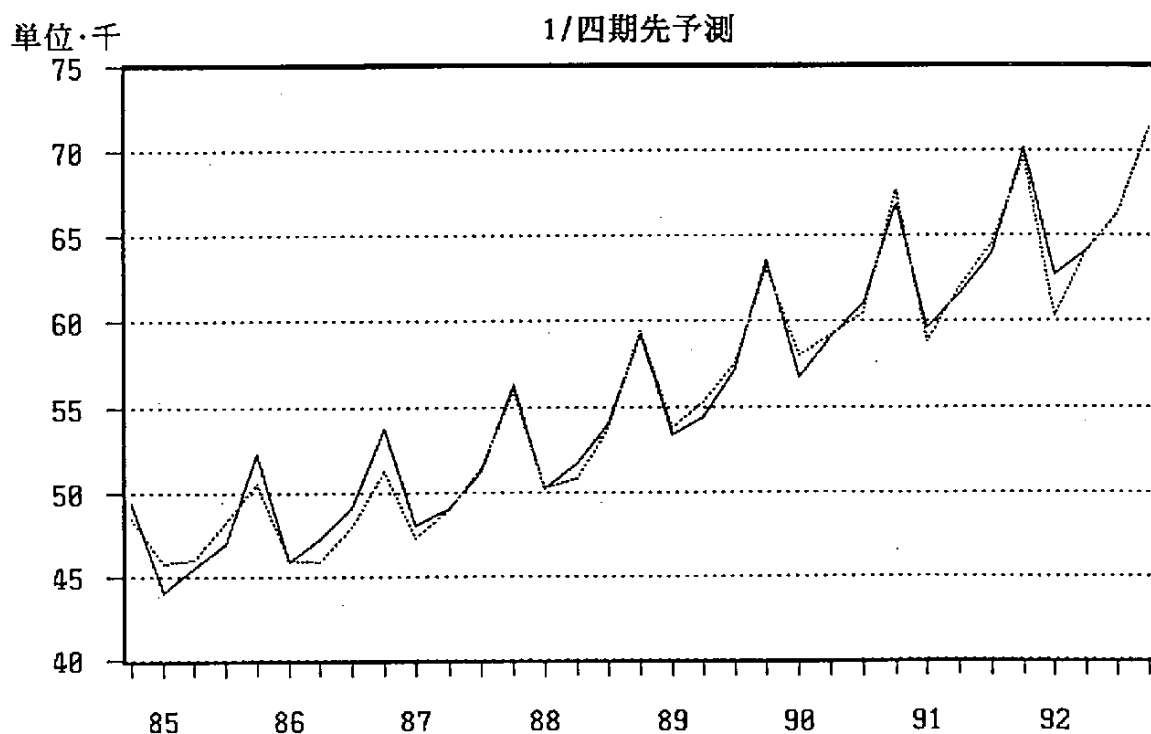


3-18図 ニューロン数15による結果

c. 民間企業設備投資



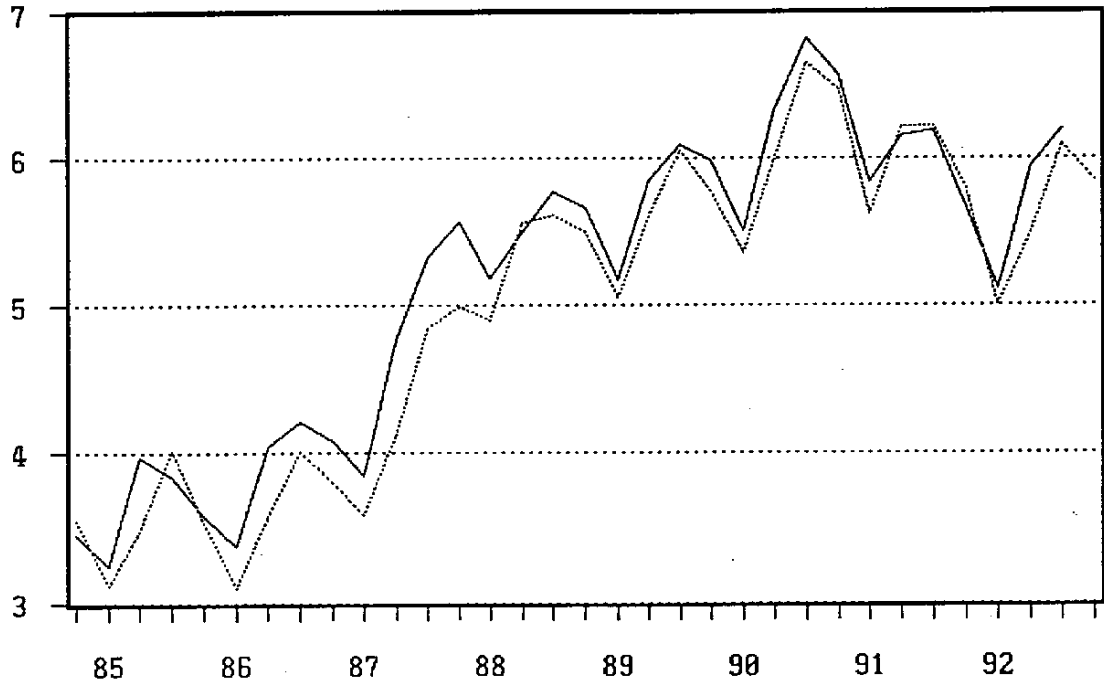
3-18図 ニューロン数15による結果
d. 在庫



3-19図 ニューロン数20による結果
a. 民間最終消費支出

単位・千

1/四期先予測

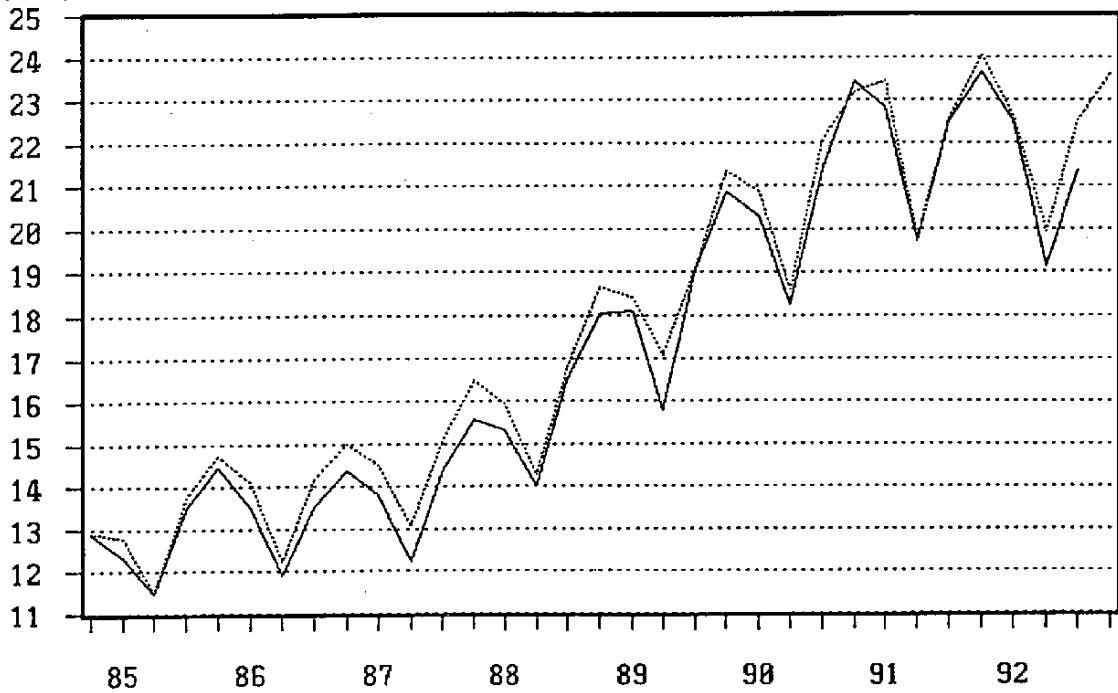


3-19図 ニューロン数20による結果

b. 民間住宅投資

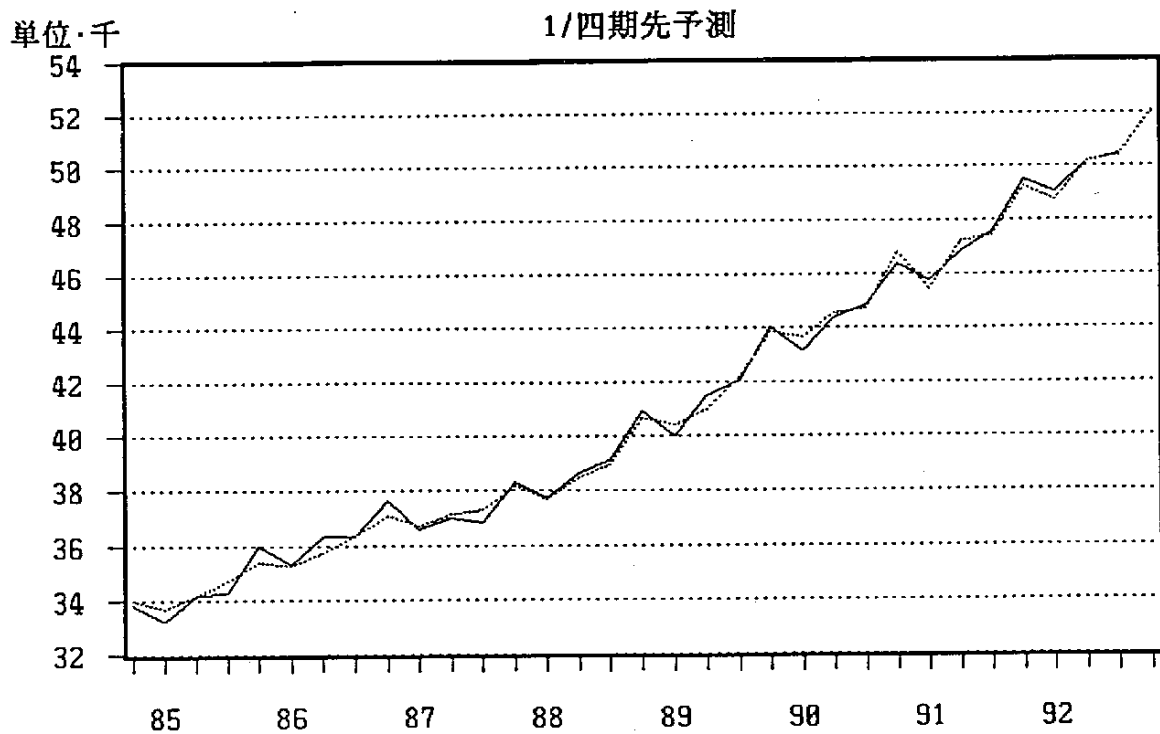
単位・千

1/四期先予測

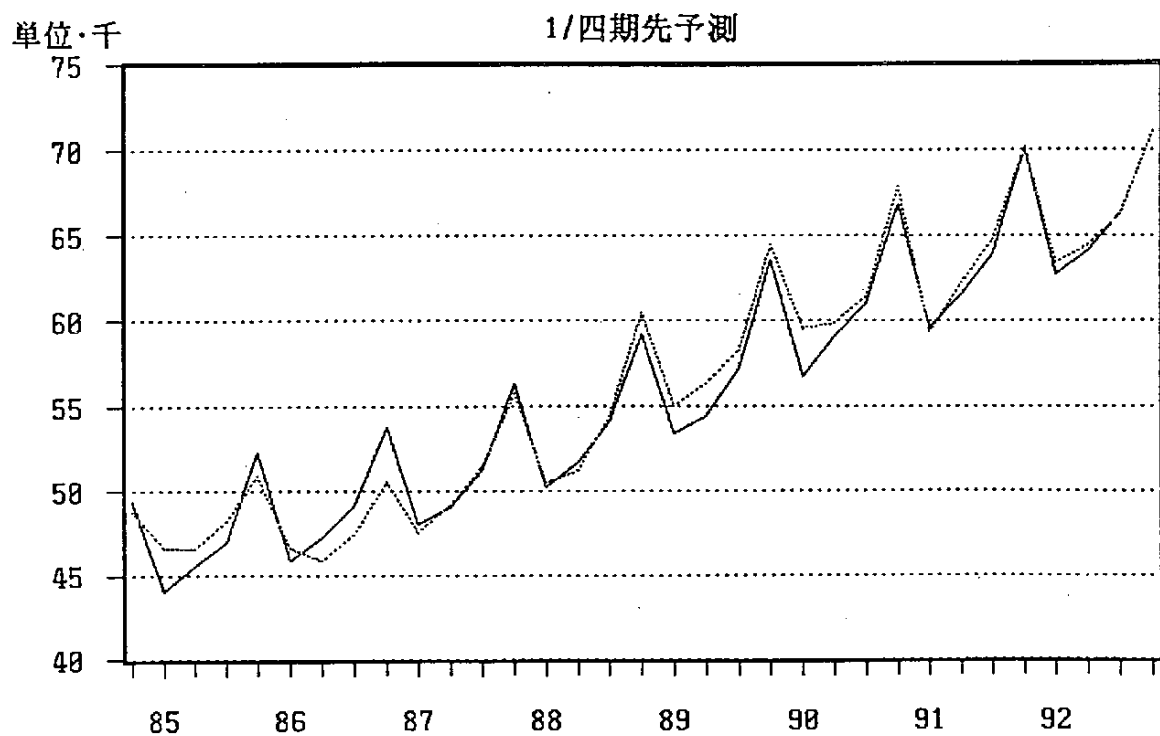


3-19図 ニューロン数20による結果

c. 民間企業設備投資



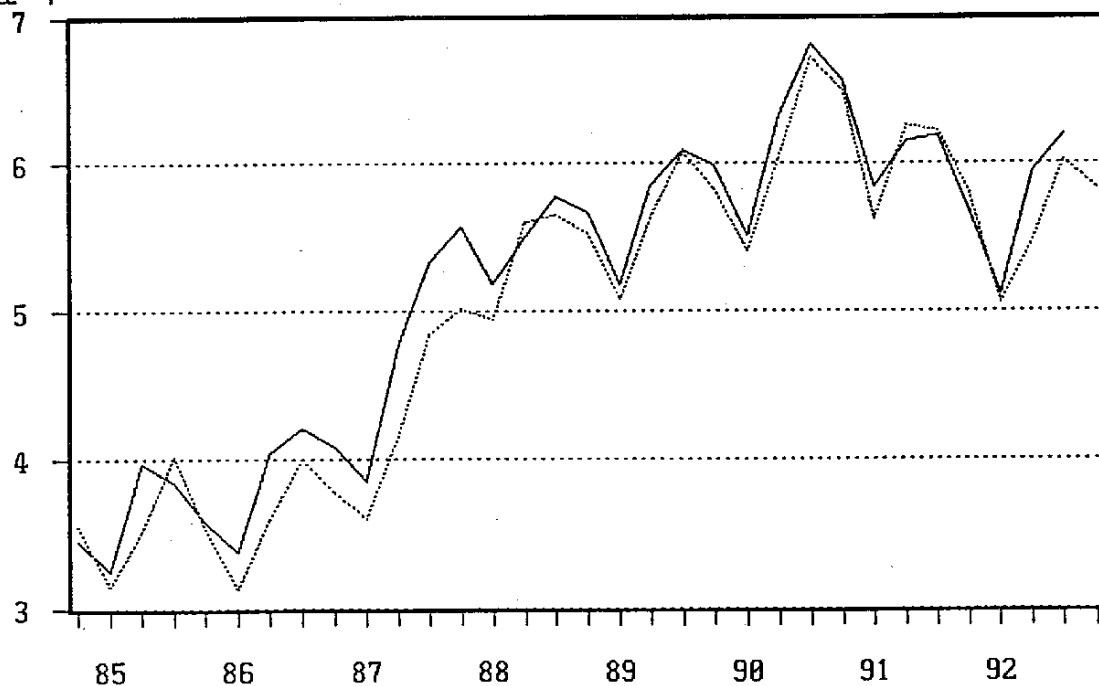
3-19図 ニューロン数20による結果
d. 在庫



3-20図 ニューロン数30による結果
a. 民間最終消費支出

単位・千

1/四期先予測

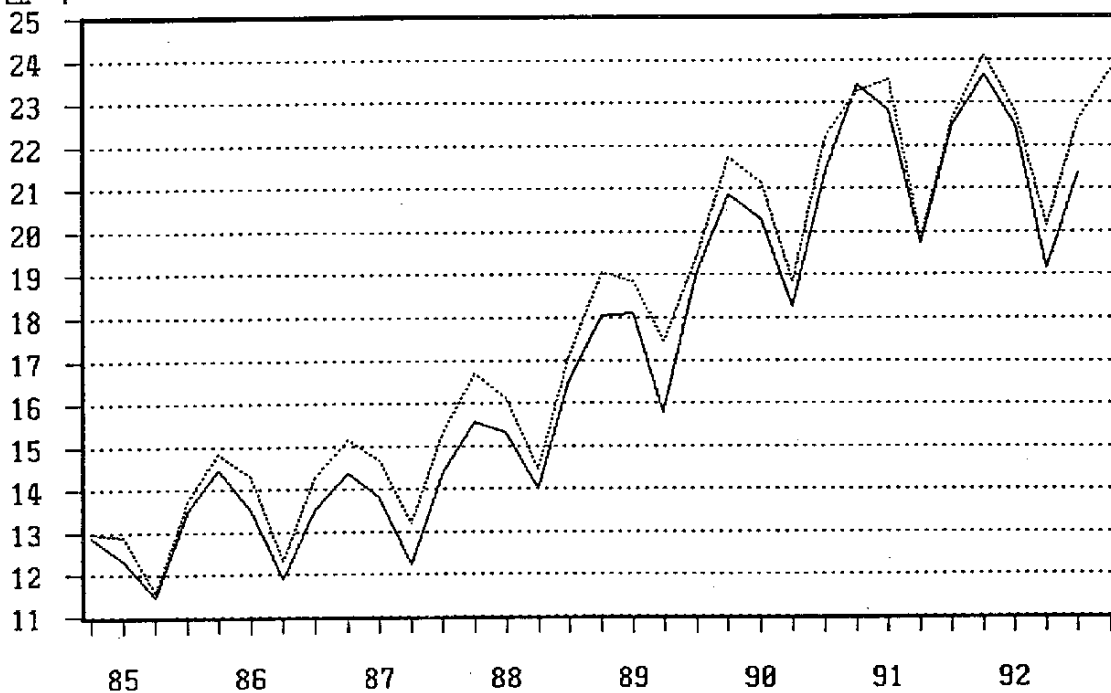


3-20図 ニューロン数30による結果

b. 民間住宅投資

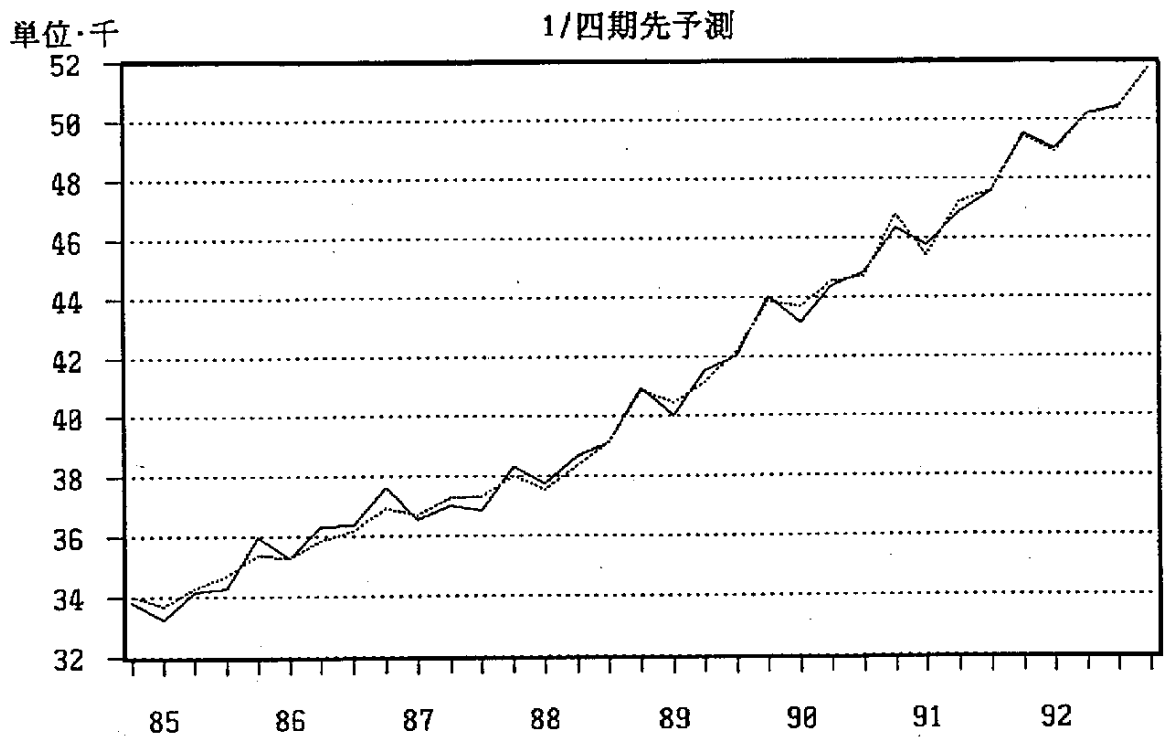
単位・千

1/四期先予測



3-20図 ニューロン数30による結果

c. 民間企業設備投資



3-20図 ニューロン数30による結果
d. 在庫

(4) 予測手法について

① ニューラルネットの経済理論への適合

ニューラルネットを用いた予測を行う前に、ニューラルネットによる予測結果が消費関数等、経済理論に準じた変数の変化を示すか確かめる必要がある。

昨年実施した年次データを用いた計量モデルによる予測の中では、民間住宅投資は一年前の公定歩合が上昇すると減少する関係が構造方程式で表現されている。

$$IH=6778.73+0.0432226 \text{ GNP}-397.353 \text{ INT}(-1)$$

また、民間企業設備投資も同様に一年前の公定歩合が上昇すると減少する関係が構造方程式で表現されている。

$$IE=3587.26+0.171688 \text{ GNP}-1032.1 \text{ INT}(-1)$$

今回、中間層の変化で用いた中間層ニューロン数11のニューラルネットを用いて、1991年3・四半期から1992年1・四半期までの3期間で1四半期前に公定歩合が2.5%に下げられた場合、及び同2四半期、3四半期前に公定歩合が2.5%に下げられた場合を想定して入力データを作成し、四半期先の予測を行った結果を下の表に示す。以上の結果から公定歩合引き下げの効果は2～3四半期後に民間住宅投資、民間企業設備投資を増加させる効果があることがわかる。このことは、ニューラルネットでも経済理論に準じた変数の動きをしていると考えられる。

公定歩合引き下げの効果

時点	変数名	実績値	1 四半期前	2 四半期前	3 四半期前
1991 3/4	C	0.476	0.473	0.473	0.473
	I H	0.471	0.471	0.472	0.473
	I E	0.490	0.492	0.495	0.498
	J	0.497	0.496	0.496	0.495
1991 4/4	C	0.506	0.504	0.504	0.504
	I H	0.519	0.519	0.520	0.521
	I E	0.472	0.474	0.477	0.479
	J	0.506	0.505	0.505	0.505
1992 1/4	C	0.508	0.507	0.507	0.507
	I H	0.506	0.506	0.507	0.508
	I E	0.541	0.542	0.545	0.547
	J	0.502	0.501	0.501	0.501

数値は、伸び率をシグモイド変換した値（数値が大きいほど伸びは大きい）

② 予測手法について

ニューラルネットで予測を行う場合の方法として、2とおりの方法が考えられる。1つの方法として、四半期先の予測を行い、その結果を用い次の四半期先を予測していく、これを繰り返していく方法である。

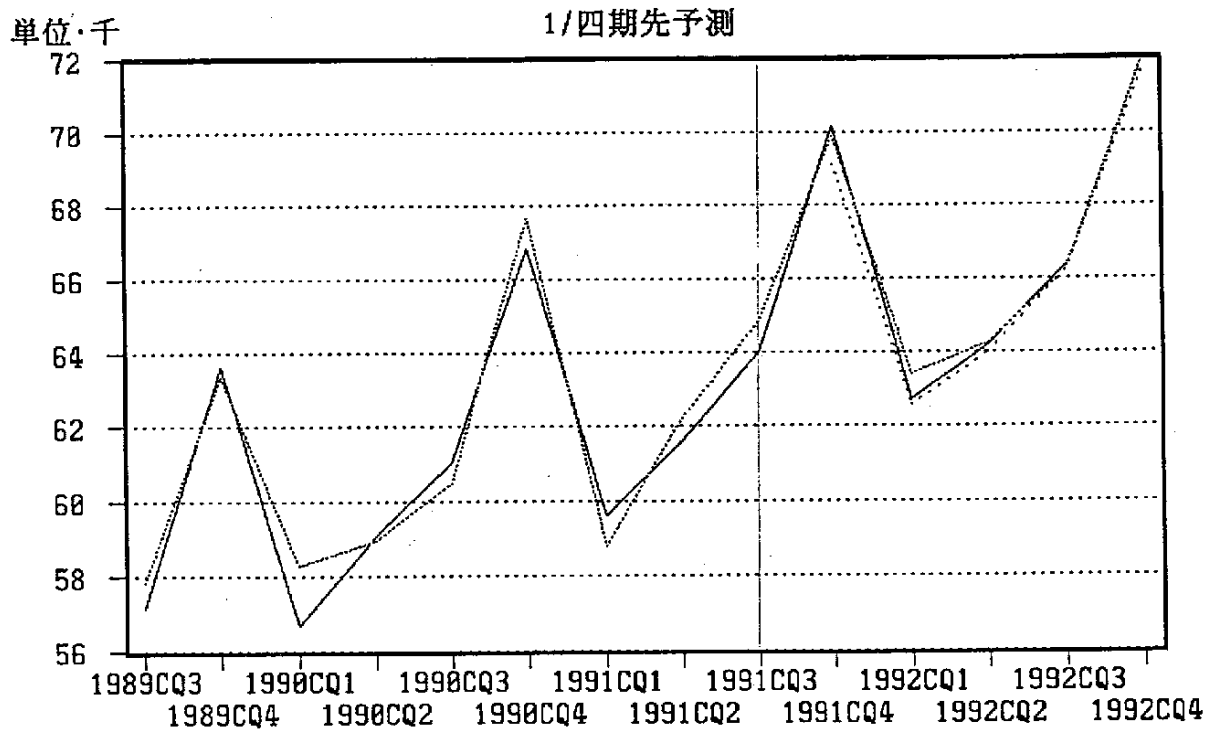
もう一つの方法は、ニューラルネットの出力パターンに1四半期先・2四半期先・3四半期先と予測したい先のパターンを学習データに組み込み、予測時に1四半期、2四半期、3四半期先まで予測する方法である。

a. 繰り返しによる予測

今回、中間層の変化で用いた中間層ニューロン数11のニューラルネットを用いて、1991年3四半期で公定歩合を2.5%に引き下げた場合、その後の予測を行った。まず、1991年4四半期の予測結果を元に1992年1四半期、2四半期の結果を元に3四半期と、順に1992年4四半期まで予測を行った。同時に公定歩合を引き下げなかった場合についても予測を行った。その結果を3-21図に示す。予測値を公定歩合を引き下げた場合と、引き下げなかった場合とを比較して民間消費支出については時間の経過にともなった大きな変化が見られないが、住宅投資はわずかに増加、民間企業設備投資は増加する結果が得られた。

一方、実績値と公定歩合を引き下げなかった場合の予測値を比較すると、実績値は、民間最終消費支出は予測値が常に低い値となっている。このことは、ニューラルネットの1四半期先の予測誤差は±5%以内であることから、予測結果をもとに次の予測を行う場合誤差は±10%、2期先の予測を行う場合±20%になりかなりの誤差が生じることになる。よって長期間にわたり予測を繰り返した場合、誤差が大きくなることと一致する。

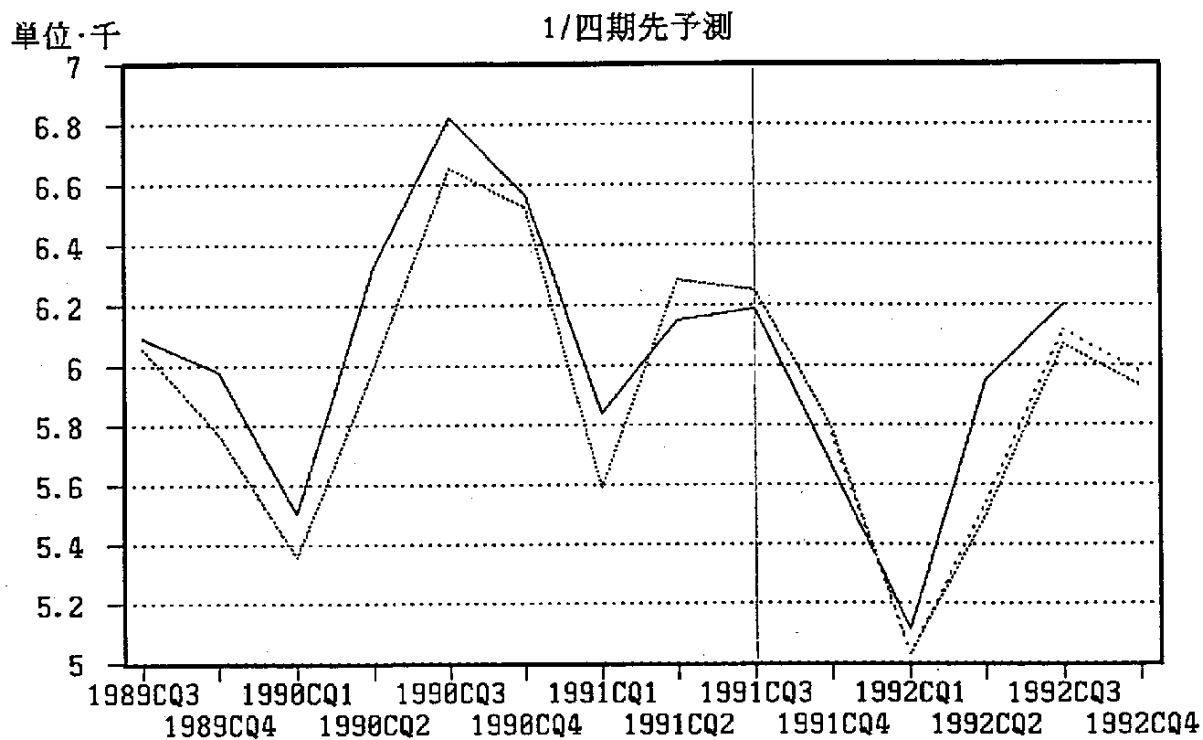
以上の結果から繰り返しによる予測は四半期データであれば1年、年次データであれば3年先の予測が限界であると考ええる。



3-21図 繰り返しによる予測

a. 民間最終消費支出

—— 公定歩合を下げなかった場合
 公定歩合を 2.5%下げた場合

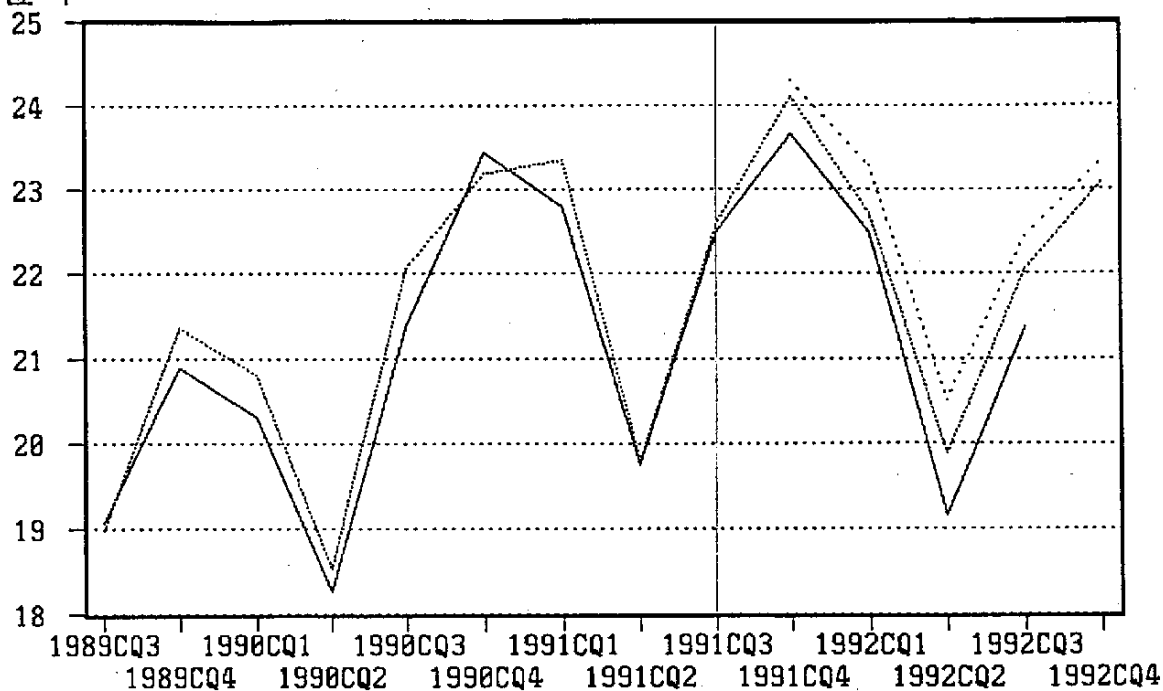


3-21図 繰り返しによる予測

b. 民間住宅投資

単位・千

1/四期先予測

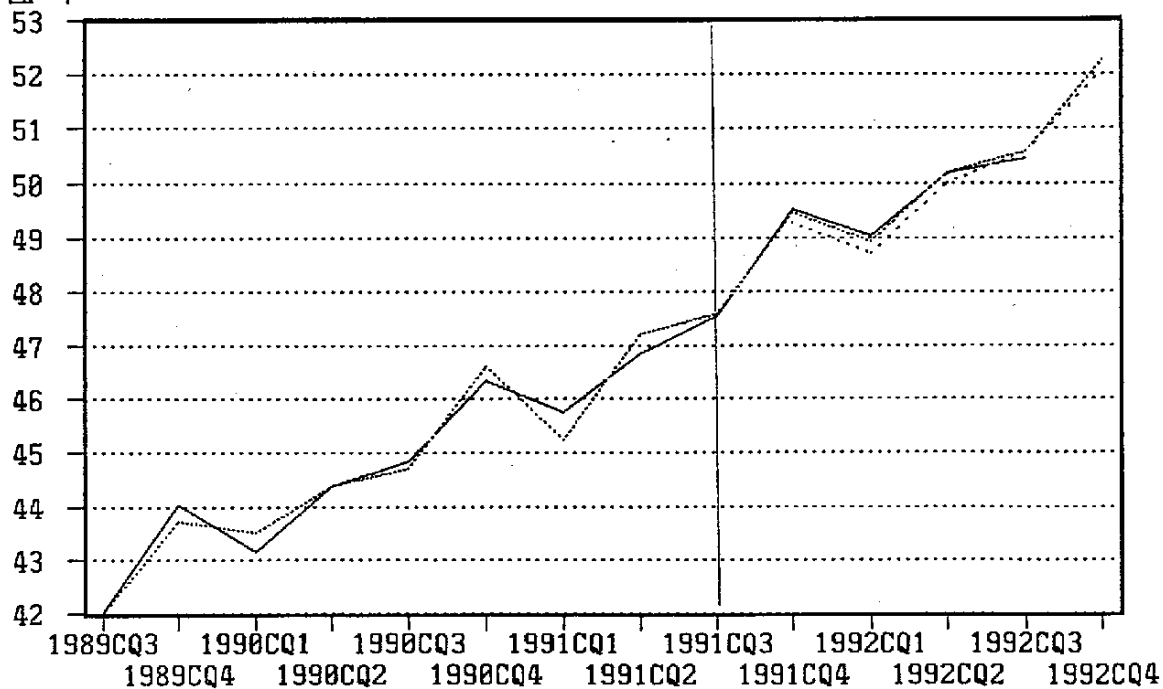


3-21図 繰り返しによる予測

c. 民間企業設備投資

単位・千

1/四期先予測



3-21図 繰り返しによる予測

d. 在庫

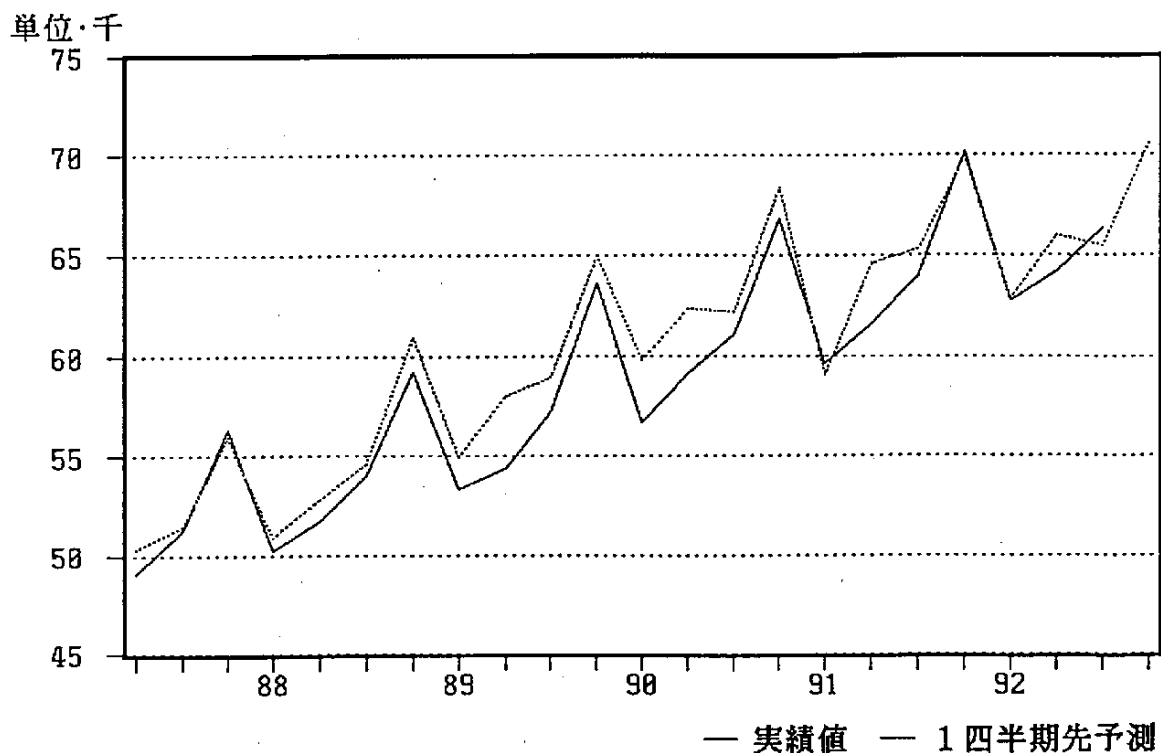
b. 同時に予測を行う場合

今回用いた方法は、ニューラルネットの出力パターンに1四半期先・2四半期先・3四半期先と予測したい先のパターンを学習データに組み込み、予測時に一度に1四半期、2四半期、3四半期先まで予測する方法である。この方法は繰り返し法に比べ簡単に予測ができるメリットがある。

予測に用いたニューラルネットは入力27（繰り返し法と同じ）、中間層18、出力層12（1四半期先データ4、2四半期先データ4、3四半期先データ4）を用い2000回学習を行った。結果を3-22図に示す。

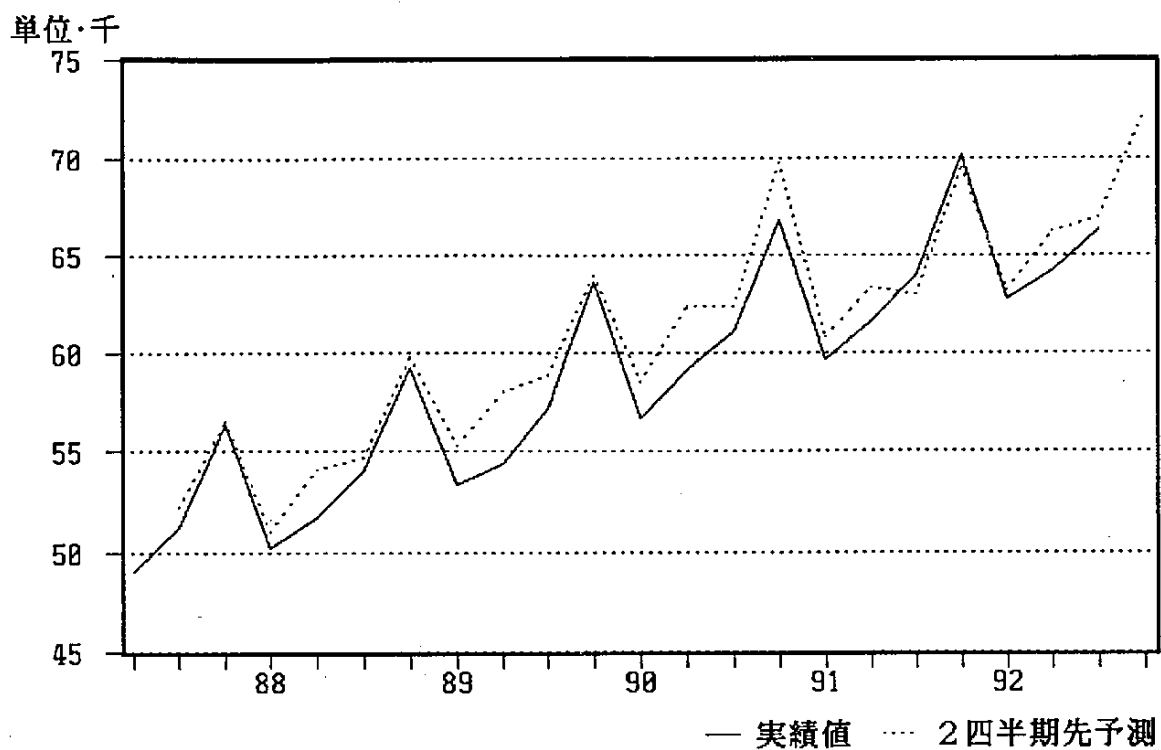
以上から、民間最終消費支出については3四半期先、民間住宅投資については1四半期、民間企業設備投資・在庫については1～3四半期でかなりよい結果が得られている。

しかし、一般的にどの四半期先の予測が良いかは上記の結果からは判断しにくく、実際に学習をさせてみてわかる試行錯誤的要素が大きいこと、また、3四半期先を予測する場合において、その間に公定歩合などの変化が生じてても全く予測結果に反映されないことなどを勘案すると、人為的、外的要因（外生変数）が頻繁に変化するような経済予測に用いるにはその適合が難しい。



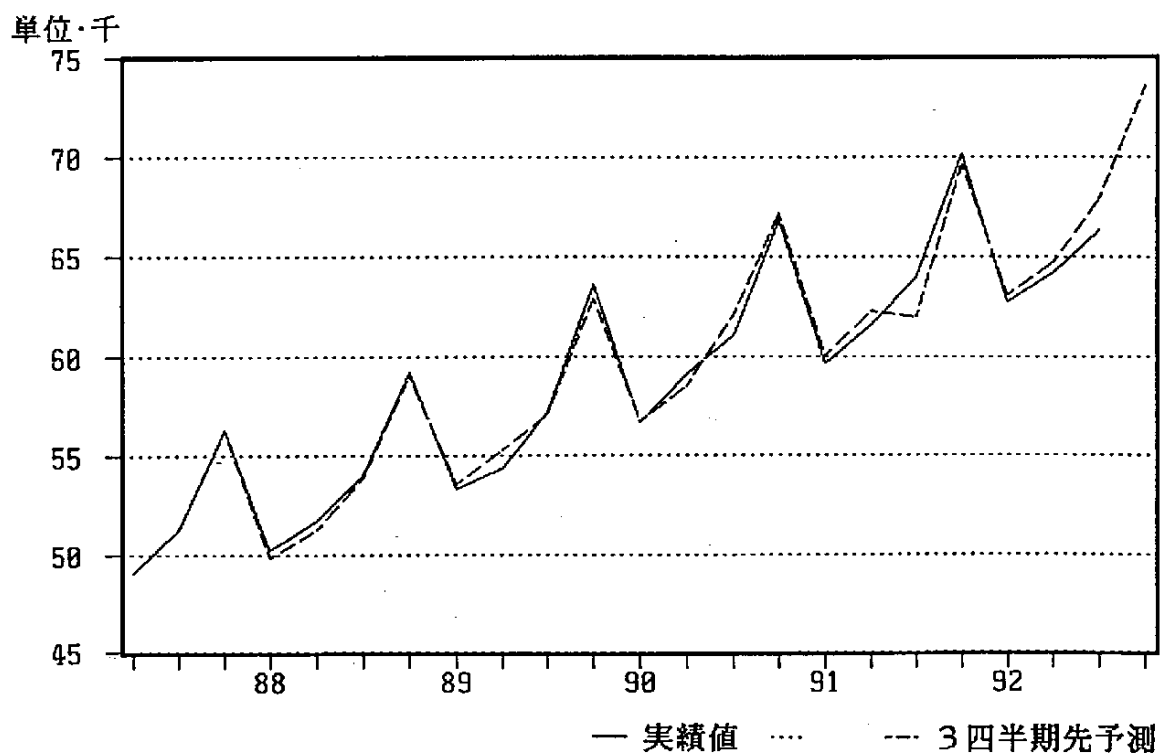
3-22図 同時に行う予測

a-① 民間最終消費支出



3-22図 同時に行う予測

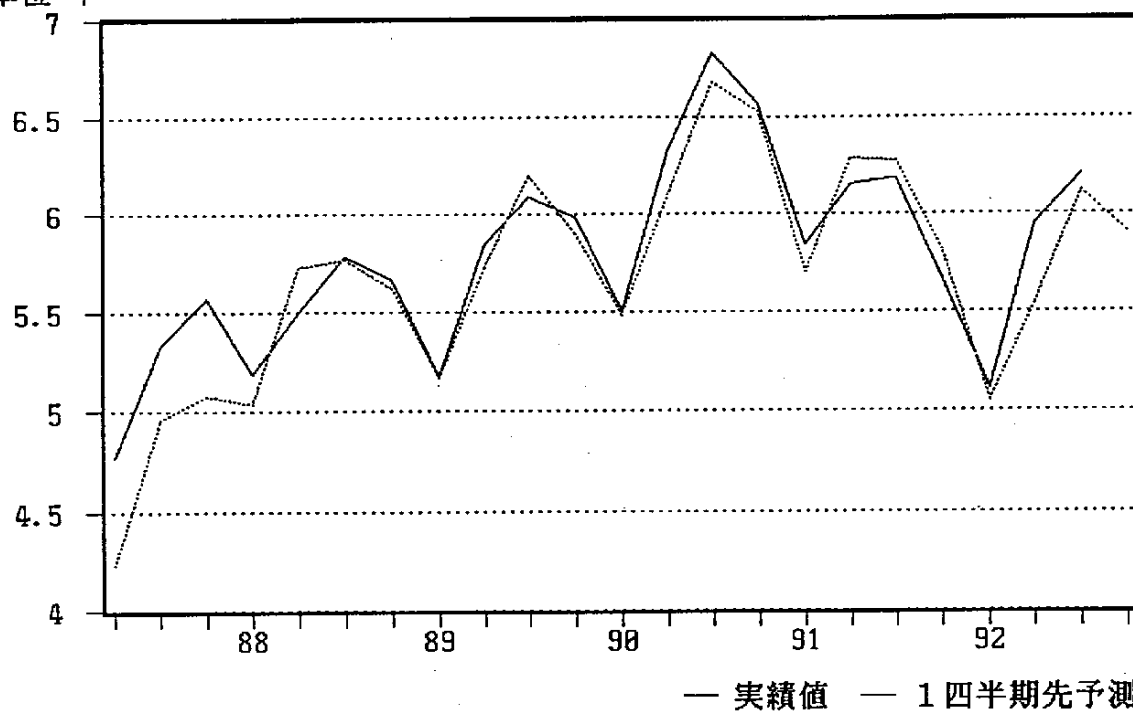
a-② 民間最終消費支出



3-22図 同時に行う予測

a-③ 民間最終消費支出

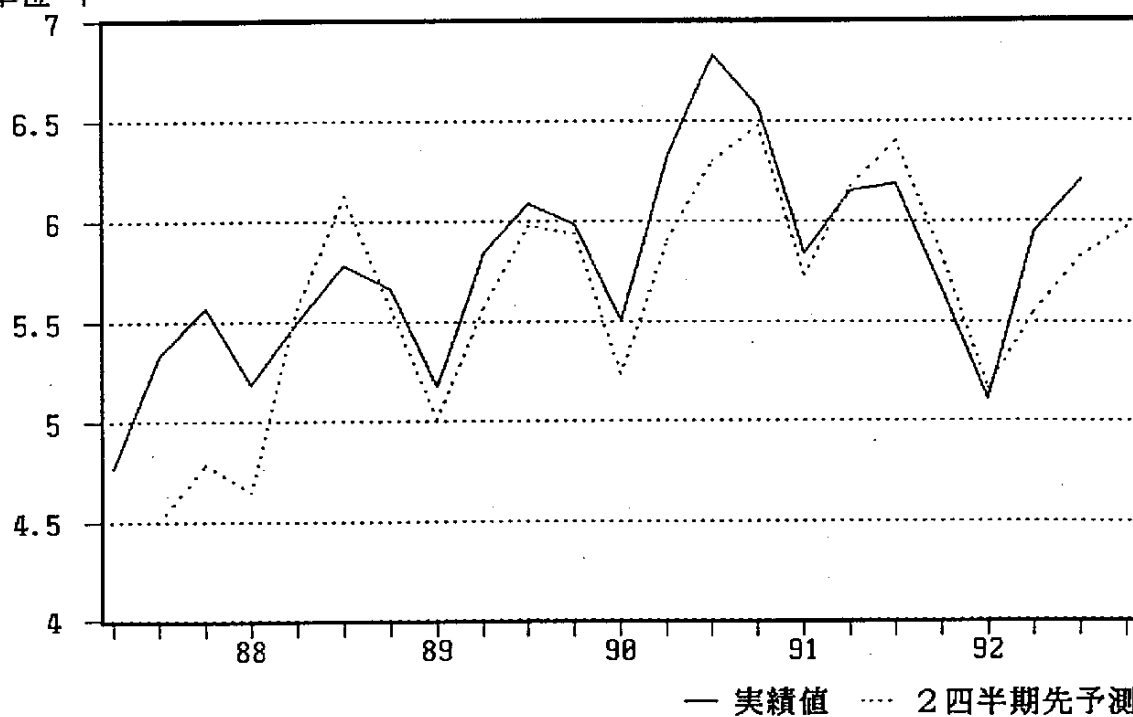
単位・千



3-22図 同時に行う予測

b-① 民間住宅投資

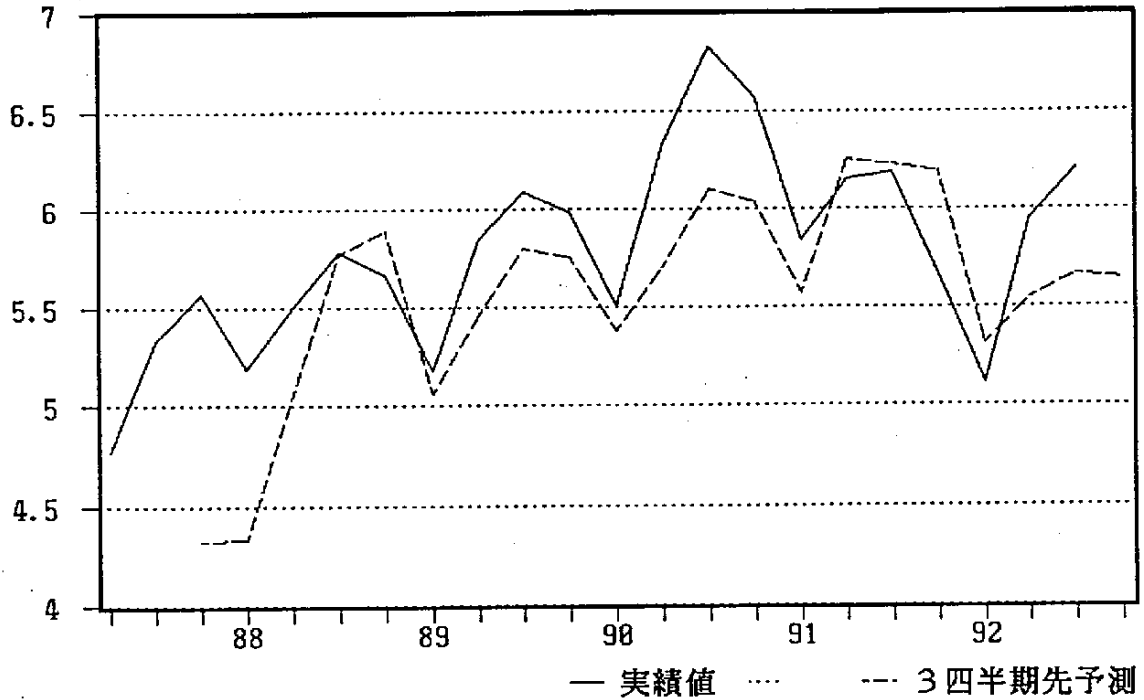
単位・千



3-22図 同時に行う予測

b-② 民間住宅投資

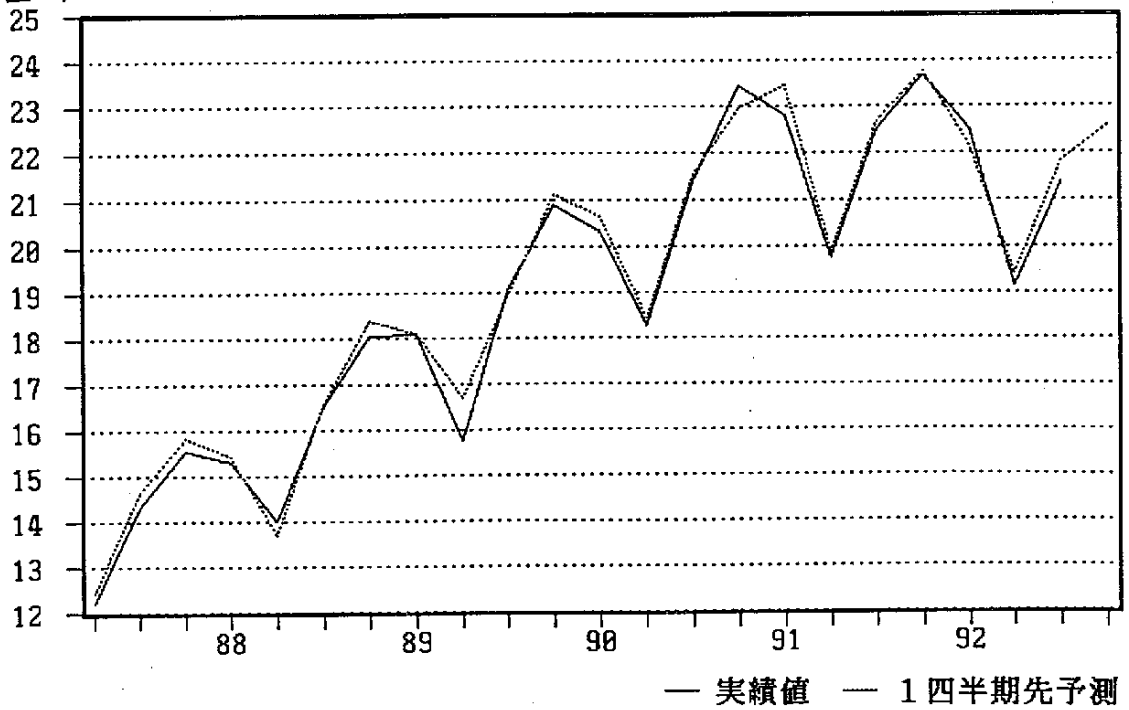
単位・千



3-22図 同時に行う予測

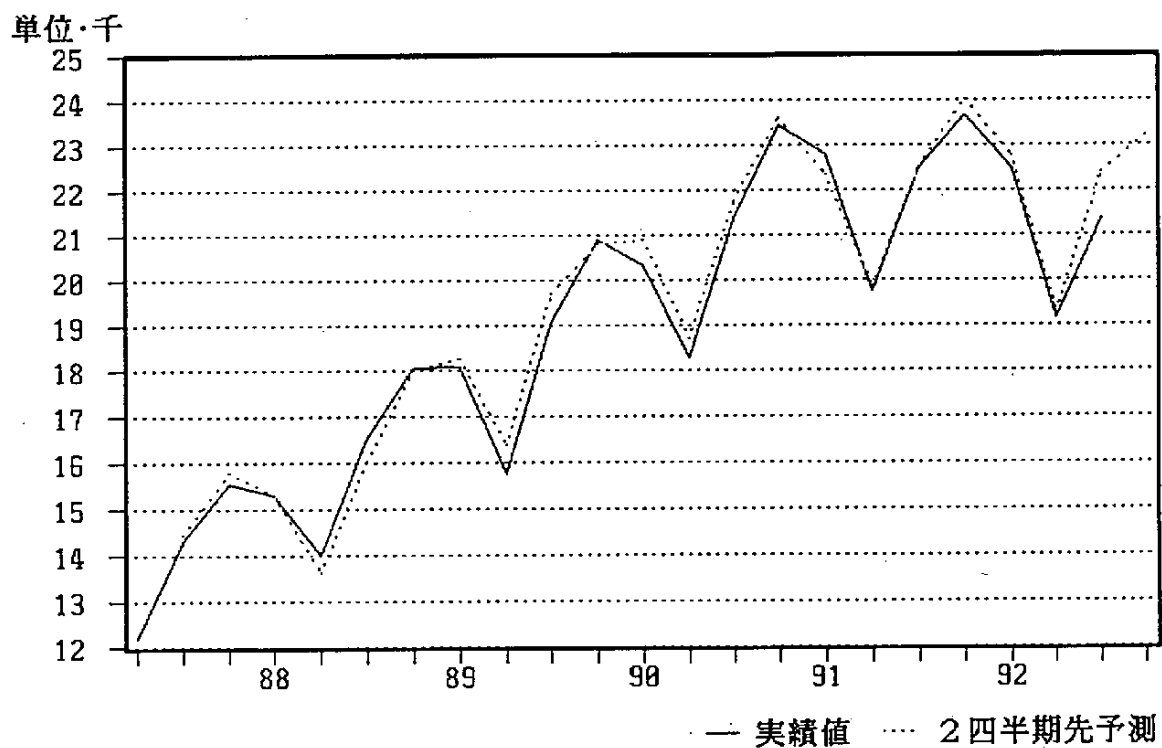
b-③ 民間住宅投資

単位・千



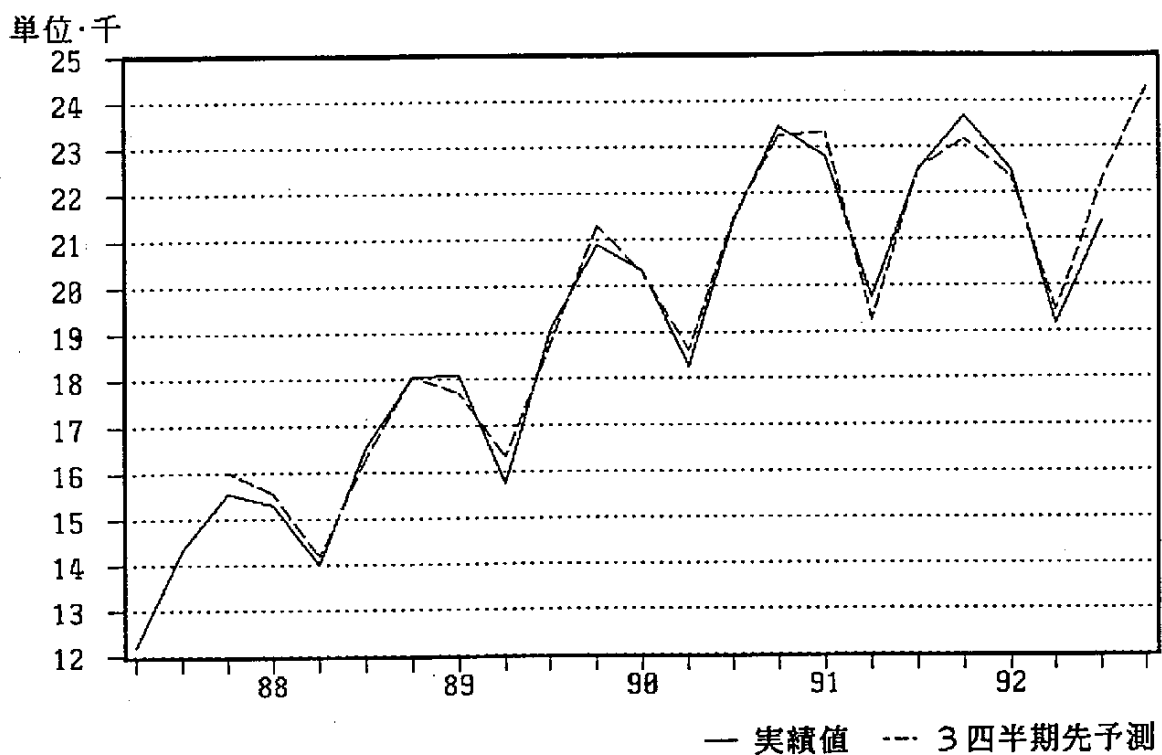
3-22図 同時に行う予測

c-① 民間企業設備投資



3-22図 同時に行う予測

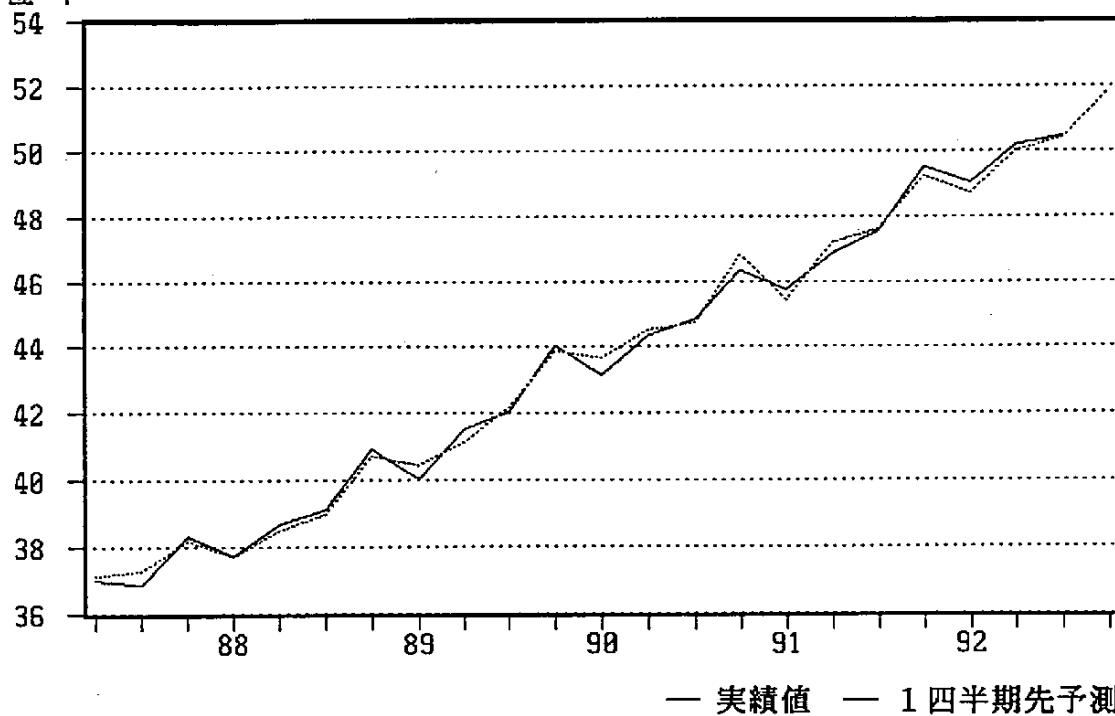
c-② 民間企業設備投資



3-22図 同時に行う予測

c-③ 民間企業設備投資

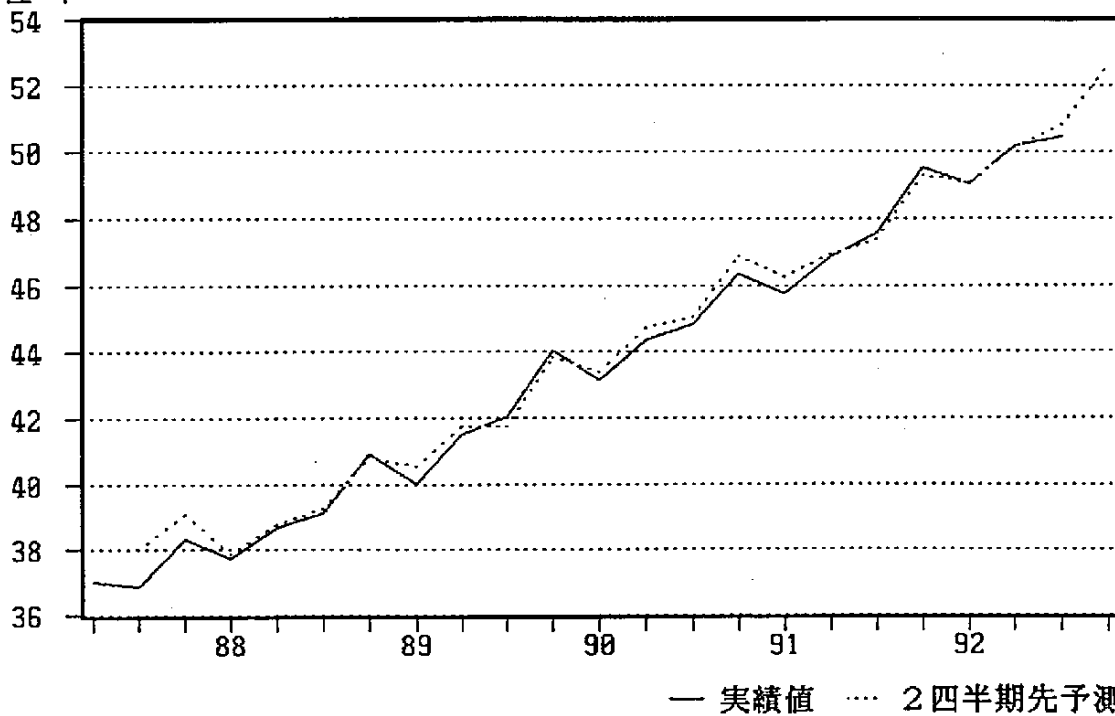
単位・千



3-22図 同時に行う予測

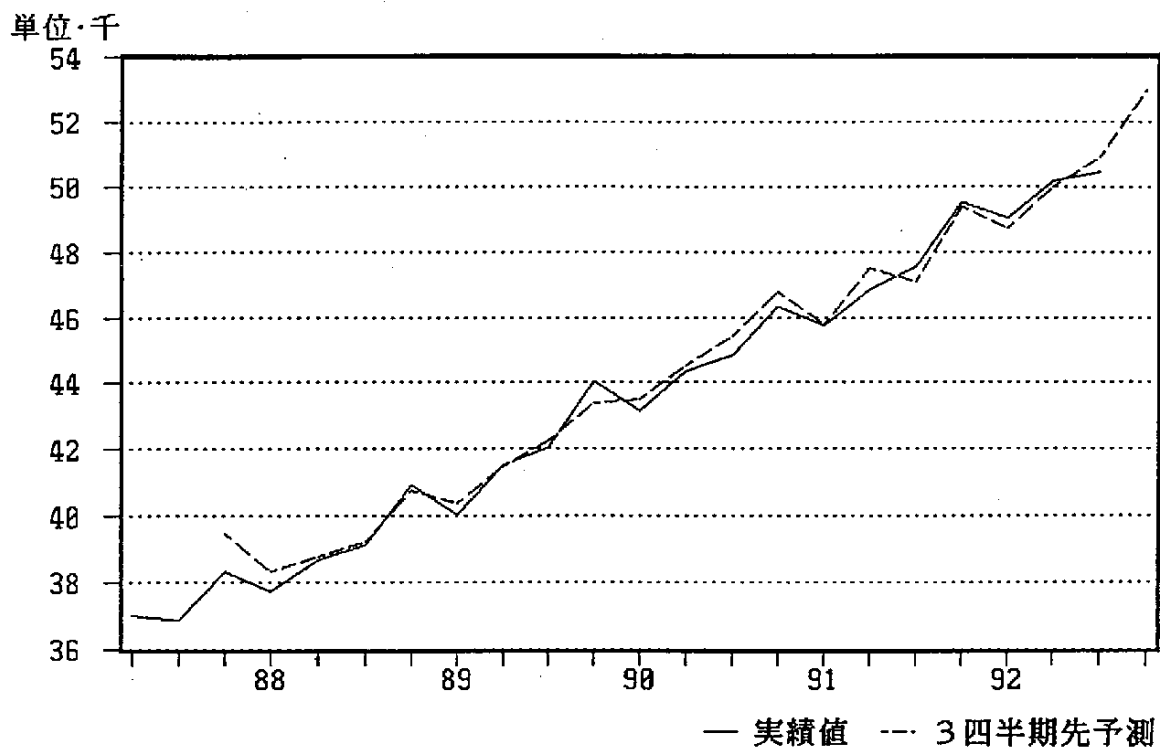
d-① 在庫

単位・千



3-22図 同時に行う予測

d-② 在庫



3-22図 同時に行う予測

d-③ 在庫

c. 予測手法のまとめ

ニューラルネットで考えられる2とおりの予測方法について比較検討してきたが、予測手法としては、今回の例では四半期先を予測し、その予測結果から次の結果を予測していく方法が一番良い方法であると考え。また、この方法の特質から予測誤差は次に引き継がれるため、四半期データであれば1年先、年次データであれば3年先までを予測することが限界であると考え。

(5) ニューラルネットを用いた経済予測手法のまとめ

ニューラルネットを用いた予測手法を検討してきた。予測を行う方法をまとめると以下のようなになる。

- ① 予測に用いるデータは、比を用いること。
- ② 学習に用いるデータは、長期のデータを用いるとよいと言うものではない。四半期データであれば、5年間程度にすること。
- ③ 学習に用いるデータは最新のデータを用いること。
- ④ 予測手法は、四半期データであれば四半期先を予測し、その結果をもとに次の四半期を予測する方法が良い
- ⑤ 予測手法により誤差が次に引き継がれるため、四半期データであれば約1年先、年次データであれば約3年先が予測の限界

ニューラルネットを用いた予測手法をまとめたが、多分に試行錯誤のところが多いこと、予測結果があたるか、または外れるかについての説明ができていないこと、さらに、計量経済モデルでいうところの同時決定ができないなど多くの課題が残されている。しかし、ニューラルネットによる予測も今始まったところであり、各所でこのような取組が行われ、そのノウハウが蓄積されるとともに、このような問題点も少しずつ説明されていくものと思われる。

4. コネクショニズムの経済学

— ニューラルネットワークが示唆する新しい経済学 —

4. コネクショニズムの経済学——ニューラルネットワークが示唆する新しい経済学——

(1) はじめに

もし、機械が新しい経済学をつくりだしたとしたら、それは経済学の200年の近代史の中で画期的な出来事になるだろう。

今日、コンピュータが人間の思考と思考の環境に決定的な影響を与えていることは明かである。経済学においても、コンピュータが普及しなかったならば、経済が時々刻々生み出す膨大な情報、特にその数量化された情報としての統計データを駆使するマクロ計量分析、産業連関分析などは今日の水準まで発展することはありえなかったと断言できる。しかし、この場合もあくまでもコンピュータ抜きにも存在可能だった理論がコンピュータの助けによってより発展したというのにすぎない。問題は、このコンピュータの発展がまったく新たな経済学、革新的な経済学を生み出すなどということが、いったいありうるのかということである。

コネクショニズムとはニューラルネットワークという機械とともに生まれてきた、新しい学問的な立場をあらわしている。本稿は、この新たな学問的立場が経済学に与えているインパクトの大きさを正當に評価しようという一つの試みである。

ニューラルネットワークは単にコンピュータ技術の発展が生みだした新しい機械にとどまらない。それは、経済学がこれまで当然のこととして反省することもなく受け入れていた方法原理の特殊性を暴き出した。すなわち、経済主体、あるいはより一般的に経済は、集中化された情報を担ったシンボル、記号に対して論理的にモニター可能な反応を行うことによって活動しているということである。ここで集中的情報シンボルというのは、所得、賃金、雇用、利潤、資本など、およそ経済学的概念といわれるものすべてである。また、モニター可能というのは、単純化すれば「説明可能」といってもよい。したがって、簡単に、経済学は「経済主体と経済の活動が経済学的用語によって説明可能なものとしておこなわれている」ことを前提にして成立している、といってもよい。ニューラルネットワークはこの大前提に疑問を提出しているのである。

前述の経済学の方法原理は、複雑化する一方の経済の中で時代遅れになりつつある。その最も象徴的なあらわれは、大規模マクロ計量モデルの衰退である。数100本、あるいは1,000本を超えるような構造方程式を連立させて構成されるマクロ計量モデルは近代の経済学が生みだした怪物である。大規模で複雑な経済を、従来の経済学の方法原理の上に大規模で複雑なモデルを構成することによっては近似できないことが誰の目にも明らかになってきているのである。

ニューラルネットワークが示唆する新しい経済学は、「経済主体と経済は非記号的な、分散化された情報の直感的な総合判断のもとに行動する」ことを前提にする。判断と行動が説明可能なものであることを前提としない。これは、ただし行動と判断をシミュレート可能な機械が存在することを要求する。そして、この機械と人間のインターフェイス機能としてののみ、「モニター機能としての説明原理」が必要となるのである。以下では、この新しい経済学の基本的な方法について若干の

解説を試みる。

(2) 均衡の可能性と「直感する経済人」

経済学にとって最も重要な概念の一つである均衡の概念を材料にして議論をはじめることによろう。経済学における均衡の概念は、ある経済活動の状態がその活動に関わる主体にとって積極的に変更する必要のない状態として定義される。ただし、ここでの経済活動はある一定の範囲内に画されているものと考えられる。その範囲外のもの、与件として与えられる環境となっている。たとえば、ある消費者の均衡という場合、消費対象となる価格が与えられているもとでの消費選択の均衡であるとか、購入、販売価格が与件となっている場合の企業の均衡生産水準とか、外国と資源の環境が与えられた場合の、一般的経済均衡であるといった形で、均衡概念は用いられる。

均衡概念は、単に経済の調和的状态の存在可能性を暗示するという意味で利用されてきているわけではない。均衡は経済学に常に要求される「説明」のために、最も多用されてきた概念の一つなのである。市場における価格の変化が需給の均衡を達成することを否定する経済学者も、何らかの経済問題、たとえばインフレーションや失業といった経済問題を「説明」するために用いたモデルには需給の均衡があらかじめ与えられている場合がほとんどである。いたるところ不均衡であり、関係する多数の主体が状態の変更を必要としている状態しか用いずに経済問題が説明されることは希である。均衡は、経済学の説明原理にしっかりはめ込まれた概念なのである。

経済学にとってこれほど便利な均衡概念も、均衡の達成それ自体が経済の現実からかけ離れているとして、根強い批判が行われてきた。それらの中で、最も興味深いものが塩沢由典氏によって行われた「合理性の限界」としての均衡論批判である。塩沢氏は、消費者の効用最大化や企業の利潤最大化の主体均衡も一般均衡も実際的な時間内で到達し得ない、均衡を求めるための「計算」が最適解に到達し得ないと主張する。塩沢氏は、単純化した消費者効用の最大化問題についてコンピュータのアルゴリズムを定式化し、商品数が30を超えると実際的な時間内で解けなくなってしまうことを示している。計算時間が、 n を商品数とすると実数の n 乗のオーダーで爆発的に増大するからである。さらに、これを最適解ではなく、近似解を求めることで満足したとしても、人間の計算能力ではそれを実現することは現実的ではないと主張する。塩沢氏はこの結論の上に、最適原理ではなくH.A.サイモンの「満足原理」を経済学が受け入れることを要求し、それと両立しうような経済学の構築を課題として提示している。

塩沢氏自身の意図はどうあれ、この議論で一つ興味深い点は、経済学の理論的主張をするためにコンピュータが本質的な媒介物として用いられていることである。機械が経済学的主張の決定的な背景となっていることである。経済的均衡は基本的に関係する主体の最適解として与えられる。しかし、その最適解は計算可能ではない。なぜなら、コンピュータですら実際的な時間内で解けないのだから。コンピュータが、経済学批判の手段として用いられているのである。

塩沢氏の主張は、塩沢氏の前提のもとでは受け入れざるを得ないように思える。経済主体の行動

が満足原理によるものであるか最適原理によるものであるかは二つの原理定義の仕方にもよる。消費者は、たとえスーパーマーケットにおける買い物においても、そこで実際に買い物をする商品の数はもちろん、そのスーパーで目に触れる商品の数よりもはるかに多数の商品を潜在的には考慮しながら最適な消費をしていると考えることは、ただちに誤りだとはいえない。消費者はさまざまな宣伝手段によって、ほとんど無数の商品とそれらのおよその価格を潜在的には意識していると考えすることはただちに問題であるとはいえない。そして、スーパーでの買い物の際、効用水準に関して無差別ないくつかの買い物の組み合わせは存在したとしても、多くの場合は空間的、時間的な制約の範囲内で最も望ましい選択をしたと考えながら消費行動を終えているのである。消費者の消費選択は、多くの場合、他の選択に代替しがたいという意味で最適なのである。

塩沢氏の、最も問題ある前提は消費者もまた「計算する人である」という前提で、最適原理の可能性を吟味していることである。確かに、消費において計算は重要な役割を果たすが、不可欠の計算は財の購入の際の総額を「計算」することくらいである。コンピュータと同じような逐次計算のアルゴリズムで予算制約と効用水準の変化を「計算」したりはしない。実際、消費者にしろ企業にしろ、すべての経済主体に関して、行動の裏付けのために計算が果たしている役割ははるかに小さい。一般に、「経済人は計算する人ではない」のである。

塩沢氏は、最適原理を放棄しても、経済人を「計算する人」とする見方は放棄しない。もしこれを放棄すれば立論そのものが無効になるから、「計算する人」の維持は不可欠である。そのかわり、消費行動の例でいえば一つの商品についてある一定の満足を超えていれば予算の許す範囲内で購入するというものである。これは、余りにも消費者を無能にしている。消費者が、先に指摘したように、商品の量的組み合わせ、つまりバスケットの選択を意識していることは明かであり、そのもとで多くの場合、最適であると理解して消費バスケットを選択しているのである。このように、消費者行動を極端に単純なものとみなさざるを得なくなったのも経済人を「計算する人」と考えることにこだわったためである。

なぜ、「経済人は計算する人ではない」のか。実は、塩沢氏の主張の最も積極的意味は、この問に対する解答を与えてしまっていることにある。「計算する人」は、余りに愚図でありのろまな人になってしまうのである。計算というのは「非常に遅い」あるいは「時間のかかる」情報の処理方法なのである。この点でわれわれは一種の錯覚に陥っている。現在のコンピュータは「計算する機械」である。それは、人間の計算という非常に特殊な思考の機能をシミュレートしている。特殊なものであっても、この計算という情報処理をコンピュータは人間から比べるととてつもない速さでそれを実現している。したがって、われわれはコンピュータの計算をすべて速いものと思い込んでしまっている。しかし、コンピュータは計算が速いだけであって、本来人間が計算以外の方法で処理しているものを、シミュレートさせればとてつもなく遅い。人間の全体的な情報処理の中で、計算は最も遅い情報処理系なのであるから。

人間の備えている情報処理系の中で、最も支配的であるのは「直感」である。われわれは、直感

を何か不正確で、避けることが望ましいもののように考えがちであるが、それは特殊な情報処理問題に限られている。人間は生活の中で最も直感を多用し、長い訓練によって基本的に最適な判断を得ようになっている。何も経済学の例を持ち出さなくてもよい。人間が地上を移動するルートを判断する場合、計算しているとは多くの場合、逆に時間を浪費してしまう。もちろん、すべてが最適な判断とはならない。時にはまちがい、時には最適に近い判断で満足しなければならないことがある。しかし、多くの場合、最適であることが人間の日常生活の定常性を実現しているのである。

経済活動においてもこの点はいかならない。経済という複雑な系が時に重大な混乱を引き起こすものの、これだけ持続しているのは、経済主体の判断がある程度の誤りは避けられないものの、経済主体に与えられた情報の範囲では基本的に最適な判断をしているからである。そして、その判断の根拠が「訓練された直感」なのである。したがって、現実の経済人は「計算する人」ではなく「直感する人」なのである。このような議論をふまえれば、均衡を説明原理として用いることは、均衡は計算によってではなく訓練された直感によって実現されているという、仮定をおくことによって有効性を持ちうるといえなくもない。しかし、より重要なのは「直感する経済人」を前提とした新しい経済学を構築することである。

(3) 逐次記号処理と並列分散処理

ここでは、機械そのものについていくつか必要な議論をしよう。コンピュータがあくまで「計算する機械」でしかなかったとしたら、これ以上の議論は余り意味をもちえないだろう。すなわち、いくら「直感する経済人」を前提にするといったところで、それは事実をただ承認しただけである。あくまで経済学が経済的事実、問題に対する説明を直接供給するだけのものであり、コンピュータはその説明のための手段としての「計算する機械」であり続けるなら、「直感する経済人」という命題はなんら活かされることはない。しかし、「直感する機械」がニューラルネットワークによって実現する可能性があらわれてきたことによって「直感する経済人」を前提とした経済学が示唆されるにいたった。それは機械の存在をはじめから前提にした経済学である。

まず、「直感する機械」としてのニューラルネットワークの情報処理原理を現行のコンピュータの原理と対比させておこう。

現行のコンピュータの処理原理を逐次記号処理とあらわすことができる。逐次とは、一連の個別処理の順序が結果に影響を与える処理系であることを意味する。処理が一つの順序を持った流れとして行われるといってもよい。記号処理とは情報が事前に記号の中に集約されていて、記号が処理過程で常に同じ情報単位として扱われていることである。現行のコンピュータは、処理過程で手順がメモリー上に与えられ、厳密にその手順にそって処理されることを要求する。ここに、人間の「説明」という情報の処理法、あるいは「計算」という情報の処理法と類似の特徴がはっきりとあらわれている。意味、ないしは記号の順序が決定的な役割を果たしているのである。また、コンピュータは0, 1から構成される二進値によって情報が表現されるが、結局この二つの記号の一定

のまとまった組み合わせが多数の区別された記号を構成し人間はあらかじめその記号に意味を与えることによって、処理の目的を実現するための手順を構成することができるようになるのである。

これに対してニューラルネットワークの処理原理は、近似的ではあるが並列分散処理とあらわすことができる。すなわち、処理において順序が支配的な役割を果たさない。処理過程のある部分では、継起的に処理が行われてもよいし、さまざまな処理が時間的なばらつきがあったまま一斉に行われてもよい。逆に、極端に言えば処理に秩序がなければ無いほど、結果の信頼性が高まる。そして、処理の入り口では、もちろん意味のある情報が入力されるが処理の過程では情報はまったく分散されたものとなって処理される。記号処理の場合は、処理の過程においても記号としての情報単位が生きたままになっているのとは決定的に異なっている。そして、出口においても一度、意味と情報の再構成が行われる。

この並列分散処理が、脳のハード的な機能のシミュレーションであることにも言及する必要がある。脳は100億を超えるニューロンのネットワークとして構成されている。一つ一つのニューロンが情報処理単位であるが、ニューロンは基本的に興奮している状態とそうでない状態とをとりうるだけである。したがって、情報処理の単位が担う情報は最も低レベルのものである。人間は情報を記号化するが脳はそのための特殊なハードウェアを持ち合わせていない。人間は記号を用い、計算し、説明するがそのためのハードウェアを持ち合わせていないのである。直感的な判断はこのニューロンのネットワークの並列分散処理に直接依存すると考えられている。ニューロンとニューロンを結び付けるシナプスの可塑性によって、訓練によってより正確な判断をするネットワークに構成され、直感的な判断の最適性を実現するようになっているのである。ハードウェアにより接近した処理であるので、直感とは脳の中の「速い処理系」として実現されている。では、人間がもう一方で行う、逐次記号処理、それによる計算や説明はどのように行われるのか。それは、他に方法がない以上、このネットワークのより高度な作用として実現されている。したがって、それは脳の中ではきわめて「遅い処理系」として実現されることになるのである。

この「速い処理系」と「遅い処理系」の違いを簡単に説明しておこう。いまたとえば、「き」という文字が書かれていたとする。われわれはこの文字がなんの文字であるかを即座に判断する。ここに一つの脳による情報の処理が行われたことは明かである。いま、われわれが「この文字は『き』という文字である」と判断したのを、ひらがなを知らない外国人が「なぜその文字は『き』であるのか」と尋ねたとしよう。そこで、「説明」しなければならない。「横の線が二本あって、上からその真ん中あたりに……」と説明する。この「き」という文字の説明はひとつの文字認識のアルゴリズムになっている。しかしわれわれは、このような手間のかかるアルゴリズムで文字を判断しているわけではないのである。はじめの直感的な文字認識は、脳の「速い処理系」によって実現され、説明のアルゴリズムによって可能となる認識方法は「遅い処理系」によって実現しているのである。この「遅い処理系」は、常に対応する「速い処理系」を持っているとは限らない。「速い処理系」では認識や、判断が困難な場合には「遅い処理系」のみが駆動されることもある。

ニューラルネットワークはこの「速い処理系」をシミュレートする機械なのである。

ここで、このニューラルネットワークの機械的能力の例を一つだけ示しておこう。それは、前節で述べた塩沢氏の「計算可能性」の問題とうまく対応した例である。ホップフィールドとタンクは、巡回セールスマン問題に対して、相結合型のニューラルネットワークが効率的に近似解を出しうる可能性を示して、大きな注目を浴びニューラルネットワークのブームへの一つのきっかけとなった。そしてこの巡回セールスマン問題は、セールスマンが与えられた都市を最も効率的に無駄なく巡回するルートを求めるという問題である。それは塩沢氏が示した効用最大化問題と同じレベルで、巡回する都市の増大とともに必要な計算量が爆発的に増大する問題である。

脳の機能をシミュレートするというニューラルネットワークは、未だ完成された機械となっていない。完全な完成をみるものであるかどうかすら必ずしも分かっていない。しかし、経済という、経済学の対象の複雑さの中で、それを構成する多数のさまざまな主体が「計算」によってではなく、「直感」によって最適性を実現しようとしている状況の中で、逐次記号処理という「遅い処理系」によってシミュレートしようというのは、自ずと限界がある。ニューラルネットワークという機械の発展に歩調を合わせながら、機械上で経済をシミュレートして、同時に機械自身と人間の情報処理によって、逐次記号処理という「遅い処理系」を「速い処理系」のモニターとして同時に機能させながら、経済現象の理解に迫ることは、経済的現実の発展が必要とするものになっていくだろう。

(4) 経済構造の自己組織化

ここでは、これまで提起した観点から経済学で多用されている経済モデルについて分析を加えよう。

経済モデルにおける記号は変数によって代表されるが、変数を関係づける関数もまた記号である。変数は意味を持っているが、変数に意味をもたせる知的作業は多くの情報を変数に帰属させる作業でもある。同様に関数もまた記号である。関数は部分的には一種に手続きをあらわすが、基本的には変数と同様に、分析者が意図するしないに関わらず、大量の情報を帰属させることを意味している。ほとんどの場合、分析者は変数間の関係を関数化することによって、漠然と意識している以上の情報をこの記号に体化させていることになるのである。

最も簡単な二つの経済変数間の関係で考えてみよう。 c という経済変数がたとえばマクロ的な消費水準をあらわすとしよう。すると、この記号 c に対してさまざまな情報が帰属させられたことになる。マクロ的消費概念の意味全体が帰属させられている。もう一つの変数を y として、これを国民可処分所得としよう。この記号には、国民可処分所得という意味、すなわち情報が帰属させられたことになる。この二つの変数を関係づける記号 $f(\cdot)$ を導入しよう。この関数記号は、 c と y がある一定の束縛のものに関係づけられていることを示す記号である。いわゆるマクロ消費関数である。このなんら特定化されていない関数 $f(\cdot)$ の導入自体がすでに「この二つの変数が関係づけら

れる」という情報を帰属させていることは明かだろう。この二つの変数が無関係、互いの変化を拘束しないということは否定されているのである。これらの記号化は、部分的なものではあれ経済構造を特定化していることを意味している。

関数の導入は、ここまでの段階ではまだ自由度を多く残したものではあるが、経済理論はさらにこの関数を特定化し、構造をより鮮明なものにする。まず、たとえばこの関数は連続的であるという、特定化が加えられる。すなわちやや曖昧な言い方ではあるが、二つの変数が相互に連続的に変化しうる関数である、という形での特定化である。さらに、この関数は微分可能であると特定化される。変化が常になめらかに進行するということを意味している。さらに、この微分係数の符号が特定化される。消費関数の場合、この微分係数は正であると想定される。ほとんど最後に実際の関数形が、係数の自由度のみを残して特定化される。たとえば、

$$c = \alpha + \beta y \quad 0 < \beta < 1, \quad (1)$$

$$c = \alpha y^\beta \quad 0 < \beta < 1, \quad (2)$$

などのように特定化される。

経済学者はここまでの特定化が行き過ぎでないことを、係数の自由度を残していることで漠然と信じている。しかし、コネクショニズムの立場からいえば、途方もなく莫大な特定化を行っているのである。たとえば次のようなことを考えてみよう。任意の可処分国民所得の変化 Δy を考える。すると(1)の場合は、常に対応する消費の変化 Δc は $\beta \Delta y$ で与えられる。(2)の場合はすこし複雑だが、 Δc は $\alpha (\Delta y)^\beta$ で与えられる。なにか自明なことを言っているようであるが、一方の変化に対するもう一つの経済変数の変化量が常に一定の規則性を持って与えられることである。この意味で、関数をこのように特定化することは、無限の規則性、あるいは無限の情報をあらかじめめ込むことに等しいのである。確かに、無限の規則性を想定しているのはわれわれの脳であるが、これはきわめて高度な脳の抽象化作業として実現している。一つ一つのニューロンは、単純に興奮か否かという情報、すなわち0,1の二値情報しか処理しない。したがって、この二値情報がいくら100億個存在したとしても、無限の規則性を表現することは不可能なのである。もしたとえば規則性が有限であったとしても、このような定量的な規則性を、ニューロンのネットワークで表現することはそもそも困難なのである。

関数を特定化するという作業は、経済構造の全体的な特定化のほとんど99%を行うに等しいと経済学者は考えるべきなのである。残りの1%をパラメータの実際の値によって決定することでしかないのである。大規模なマクロ計量モデルが、現実の経済の動きを追うことが難しいのは、実態経済を単にパラメータの部分にしか反映できない、すなわちたとえて言えば、モデルの構造のたった1%部分にしか現実経済を反映できないからに他ならない。マクロ計量モデルが有効性を保持でき

ないのはほとんど自明であると言わなければならない。

脳の機能との対比で明らかになることは、関数関係による構造の特定化が逐次記号処理という脳のより高次の、すなわちハードウェアであるニューラルネットワークのより総合的な機能として実現することである。このことが要求される必然性は、われわれが、経済モデルが常に可能な限り、直接に説明可能な存在でなければならないことを要求するからである。いいかえれば、モデルの構造の「どの部分」をみても、分析者自身ばかりでなく他の人にとっても「直接に説明可能であること」が要求されるからである。たとえば、1,000本の方程式群から構成されているマクロモデルを考えてみよう。そのいかなる一つの方程式を持ってきても、意味を持っている。すなわち、説明可能である。経済学においては、もしある一つの方程式の意味が「説明不可能」であったとすると、たとえ他の999本の式が意味を持っていたとしても、ただちにこのモデルは無効であると宣言されるだろう。分析者もしたがって、はじめから、意味のある式しか用いないのである。

今日の経済モデルの問題点をこのようにとらえれば、自ずとコネクショニズムの経済モデルの特徴が明らかになる。それは、事前的な構造化を最小限にとどめる。すなわち、与えられる情報そのものに基づいて経済の構造そのものも機械の中に自己組織化することである。自己組織化とは機械としてのニューラルネットワークであり、多数のネットワークユニット(=ニューロン)間の結合の強弱パターンの中に構造をビルトインすることである。このコネクショニスティックな経済構造の構成で、構造は部分的であれ全体的であれ直接に説明可能ではない。通常の経済モデルの記号の数から比べるとはるかに多数のユニットから構成されるであろうネットワークは、全体はもちろん、いかなる部分をとってもそれによって構造が直接「説明される」ことはない。

しかし、そのことは説明がいかなる意味でも不可能な「ブラックボックス」モデルとなるということの意味していない。ここで、再びモニターという機能が重要な意味を持ってくる。構造を体化させたネットワークは、脳でいえば「速い処理系」に対応し、モニター機能は「遅い処理系」に対応する。すなわち、ニューラルネットワークに対して、機械のもう一つの機能としてモニター機能を走らせ、人間とのインターフェイスを実現するのである。このモニター機能によってモデルの「説明機能」が再現することになる。したがって、コネクショニズムの経済モデルは、完全に機械上で与えられ、構造を体化させたネットワークとそれに対するモニターシステムという二つの情報処理系によって実現することになる。

経済モデルが、機械上にのみ与えられることに抵抗を感じるものがあるかもしれないが、これはそれほど不思議なことではない。たとえば、単に理論分析のためだけに用いるモデルといえども、さまざまな記号を用いて「紙の上」に書かれる。これが単に「機械上」に与えられるということに、かわっただけのことである。一般にわれわれは紙の上にモデルが与えられたことによって、概念や意味を操作し分析的にモデルを用いる。ニューラルネットワークは、単に構造をあらわしたネットワークだけでは、分析的使用が困難であるが、説明手段としてのモニター機能を同時に走らせることによって、紙上のモデルと同様に分析的使用が可能になるのである。

ただし、経済学的な「理解」の意味の一定の変更は避けられないかもしれない。経済モデルは、経済現象、経済問題を理解するための道具でもある。このような道具として経済モデルをみたときに、要求されることは、与件と、結果とを媒介するさまざまな因果の序列が説明可能になることである。計量モデルの大規模化は、ある程度この要求に対して時にはあきらめざるを得ないことを例示した。しかし、たとえ精密で複雑な数学的な経済モデルを用いたとしても因果の系列が明確に説明されることは、理解の重要な要素であることにはかわりない。コネクショニズムの経済モデルは、このような形で理解の手段として機能しないかもしれない。「モニターの不完全性」が生じる可能性が高い。特に、複雑なモデルをネットワークに体化すればするほどそうなるだろう。それは、われわれの脳の直感的判断の多くが、モニター不可能、すなわちなぜその判断がされたかを説明し尽くせないことに対応しているといえ、分かりやすい。それは、日本語の「理解」という概念に、「了解」という意味が含まれていることに似ている。われわれは、機械そのものをこの意味で理解することを要求されることになるかもしれないのである。

(5) 新しい経済学へ

コネクショニズムが必要とされる背景は次の点にある。経済主体の判断は基本的に直感に依存し、経済主体と経済のシステムとしての運行を計算する主体として近似することは極端に誤差を大きくすることである。直感とは、情報処理原理としては並列分散処理であり、複雑な経済システムのシミュレーションは経済分析上のこの処理原理を導入することによって行われなければならないのである。それはまた、構造をあらかじめ理論的に関数関係の中に表現してしまうのではなく、ニューラルネットワークという並列分散処理機械の中に自己組織化することが必要であることを示している。したがって、コネクショニズムの経済学は、もともと機械の進歩によって触発され、機械の進歩に強く依存した形でのみ発展可能である。

この新しい経済学は未だ出発点に立ったばかりである。ニューラルネットワークの経済分析への応用例も極端に少ない。しかし、ニューラルネットワークの機械的な発展と、応用分析例の蓄積は、拡大的に発展の可能性を人々に示して行くだろう。

参考文献

- [1] Ackley, D.H., G.E. Hinton and T.J. Sejnowski, 1985, A Learning Algorithm for Boltzmann Machines, Cognitive Science, Vol.9, pp.147-169.
- [2] 甘利俊一, 1978, 『神経回路網の数理』, 産業図書.
- [3] 甘利俊一, 1989, 『神経回路網モデルとコネクショニズム』, 東京大学出版会.
- [4] 甘利俊一他, 1988, 「コグニティブサイエンスは何を問うか」, 『哲学』4, 哲学書房.
- [5] 麻生英樹, 1988, 『ニューラルネットワーク情報処理』, 産業図書.
- [6] Bartlett, F.C., 1932, Remembering: A Study in Experimental and Social Psychology,

Cambridge Univ. Press. (「想起の心理学」, 宇津木保, 辻正三訳, 誠信書房, 1983年刊)

- [7] Fahlman, S.E., G.E. Hinton and T.J. Sejnowski, 1983, Massively Parallel Architectures for AI: NETL, THISTLE, and Boltzmann Machines, Proc. Natl. Conf. AI, AAAI, pp. 109-113.
- [8] Hinton, G.E. and T.J. Sejnowski, 1983, Optimal Perceptual Inference, IEEE Computer Conference on Computer Vision and Pattern Recognition, Proceedings CVPR, pp. 448-453.
- [9] 廣松渉, 1988, 「AI問題についての偶感」, 『哲学』4, 哲学書房.
- [10] 廣松渉, 1987, 「AI問題によせて——意識観の脱構築のために——」, 『現代思想』6.
- [11] Hopfield, J.J., 1982, Neural Networks and Physical Systems with Emergent Collective Computational Abilities, Proc. Natl. Sci. USA, Vol. 79, pp. 2554-2558.
- [12] Hopfield, J.J. and D.W. Tank, 1985, "Neural" Computation of Decisions in Optimization Problems, Biol. Cybern, Vol. 52, pp. 141-152.
- [13] McCulloch, W.S. and W. Pitts, 1943, A Logical Calculus of the Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133.
- [14] 中野馨他, 1990, 『ニューロコンピュータの基礎』, コロナ社.
- [15] 日本情報処理開発協会・関西情報センター, 1992, 『地域経済予測サポートシステムに関する調査研究報告書』.
- [16] Rumelhart, R.C. and A. Ortony, 1977, The Representation of Knowledge in Memory, In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.), Schooling and the Acquisition of Knowledge, Hillsdale, NJ: Erlbaum, pp. 99-135.
- [17] Rumelhart, D.E. and McClelland, 1986, Parallel Distributed Processing, Vol. 1, 2, MIT Press. (『PDPモデル——認知科学とニューロン回路網の探索』, 甘利俊一監訳, 産業図書, 1989年刊)
- [18] 佐伯胖, 1989, 「認知科学とコンピュータ」, 『情報処理』, Vol. 30, No. 5.
- [19] 佐和隆光, 佐伯胖, 1987, 「経済のリアリズムAIのリアリズム」, 『AIジャーナル』.
- [20] 塩沢由典, 1990, 『市場の秩序学』, 筑摩書房.
- [21] 鷲田豊明, 1992, 「ニューラルネットワークによる時系列データの特性認識」, 『経済理論』, 第249号, pp. 56-81.

5. 時系列データと経済予測

5. 時系列データと経済予測¹⁾

(1) はじめに

時系列データによる計量経済モデル分析や予測手法は、一方にはケインズ・モデルに代表される同時体系の連立方程式で記述される構造モデルを推定し予測するといったアプローチがあり、他方にはボックス・ジェンキンス(Box and Jenkins)²⁾によって代表されるARIMA(Auto Regressive Integrated Moving-Average)モデルやそれを複数変数に拡張したVAR(Vector Auto Regressive)モデル³⁾などによる時系列分析がある。

ニューラルネットを適用した地域経済予測の有効性については既に前年と今年の報告書においても述べられているが、この経済予測の分野における先行した経済時系列モデルとの比較検討がニューラルネット適用の意義を深める上でも非常に重要になってくる。

そこで本章では、時系列データを用いた経済構造や予測分析手法について、まず時系列(経済)データについての特性を整理し、次に今では伝統的ともいえるボックス・ジェンキンスの時系列分析モデルを述べる。そして、1980年代以降、すなわちボックス・ジェンキンスの後、今日までマクロ経済データ相互の因果関係や構造把握を前進させてきた計量経済理論について簡単に解説をおこない、今後のニューラルネットを用いた地域経済予測に関する研究課題の示唆を与える。

(2) 時系列データの特徴

① 定常過程と非定常過程

時系列データについて考えてみると、そこでは、時間に関する推移が定常過程にしたがうものと非定常過程にしたがうものとに分けられる。ある時系列データ(x_t)が厳密な意味において定常的(Stationary State)であるとは、

$$E(x_t) = \mu < \infty$$

$$\text{Var}(x_t) = \gamma(0) < \infty$$

$$\text{Cov}(x_t, x_{t-s}) = \gamma(s)$$

¹⁾本章の中心部分は、最近刊行された以下の計量経済学の教科書における時系列分析の記述に主として基づいている。

G.S.Maddala, *Introduction to Econometrics*, 2nd Edition, Macmillan, 1992.

W.E.Griffiths, R.C.Hill and G.G.Judge, *Learning and Practicing Econometrics*,

John Wiley & Sons, 1993.

W.H.Greene, *Econometric Analysis*, 2nd Edition, Macmillan, 1993.

²⁾G.E.P.Box and G.M.Jenkins, *Time Series Analysis, Forecasting and Control*,

Revised Edition, Holden-Day, 1976.

³⁾G.C.Tsiao and G.E.P.Box, 'Modelling Multiple Time Series with Application,'

Journal of the American Statistical Association, Vol.76, pp.802-816, 1981.

C.A.Sims, 'Macroeconomics and Reality,' *Econometrica*, Vol.48, pp.1-48, 1980.

これらの式の意味は、時系列データの平均が時間に関わらず一定の μ 、分散も自己共分散も時間に無関係な一定の値を示すということである。

非定常過程は、発散過程(Explosive Process)のものとトレンドを除去したり差分などをとることにより定常過程に変換できるものとに分かれる。後者は、Homogeneous Non-Stationary Processと言われ、トレンド型定常過程(Trend Stationary Process)と差分型定常過程(Difference Stationary Process)の2つに分類される。

いま、時系列 x_t が

$$x_t = f(t) + u_t \quad (1)$$

というタイム・トレンドによって説明されるものとする。ここで、攪乱項 u_t は平均が0で分散が σ^2 の定常系列とする。(1)式を線形で特定化する場合、

$$x_t = \alpha + \beta t + u_t \quad (2)$$

となり、これを最小自乗法で推定するならば、最小自乗法による誤差である $\hat{u}_t = x_t - \hat{\alpha} - \beta t$ はトレンドが除去されたものとなる。すなわち、 $\sum \hat{u}_t = 0$ 、 $\sum t \hat{u}_t = 0$ である。このようにしてトレンドが除去される系列をトレンド型定常過程(TSP: Trend Stationary Process)という。

回帰分析の代わりに(2)式の差分をとることによってトレンドを除去するならば、 x_t が

$$\begin{aligned} \Delta x_t &= x_t - x_{t-1} \\ &= \beta + u_t - u_{t-1} \\ \Delta^2 x_t &= (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) \\ &= u_t - 2u_{t-1} + u_{t-2} \\ &= \Delta^2 u_t \end{aligned}$$

というように二階差分をとることによってトレンドを除去できることになる。

他方、

$$x_t = x_{t-1} + \beta + u_t \quad (3)$$

あるいは

$$x_t - x_{t-1} = \beta + u_t \quad (4)$$

によって生成される時系列を考える。ここでも u_t は平均0で分散 σ^2 の定常系列である。この場合、 x_t の一階差分は、平均 β での定常系列となる。このモデルは、ドリフト付きのランダム・ウォークのモデルとしても知られている。また、 $\beta = 0$ のときは単にランダム・ウォークのモデルといわれる。(3)式を初期値 x_0 から加えていくと、

$$x_t = x_0 + \beta t + \sum_{j=1}^t u_j \quad (5)$$

となり、これは攪乱項が非定常で分散が $t\sigma^2$ で時間と共に大きくなること以外は(2)と同じ形である。

このように非定常な時系列データ数を分類する必要性は、回帰分析において非定常変数が存在

すると推定値の分布形状が通常の t 分布のようなものにしないからである。このため、(3) 式のようなランダム・ウォークを示す非定常変数を分析する場合は、差分（階差）をとって(4)式のような定常過程に変換することが望ましい。

② 非定常性の検定

一般に d 回の差分（階差）をとることによって定常系列が得られる非定常の原系列は、 d 個の単位根 (Unit Root) をもつことが数学的に証明される。

Nelson and Plosser (1982)⁴⁾ は、(2) を TSP: Trend-Stationary Process、(3) を DSP: Difference-Stationary Process とよんでいる。両方のモデルとも線形のトレンドを示すが、トレンドを除去する適切な方法は異なっている。時系列データがどちらの系列に属するのかを検定するために、Nelson and Plosser は、Dickey and Fuller (1979)⁵⁾ によって示された検定を用いている。これは、

$$x_t = \alpha + \rho x_{t-1} + \beta t + \varepsilon_t \quad (6)$$

というモデルを推定することを構成しており、これは、

$$\rho = 1, \beta = 0 \rightarrow \text{Difference-Stationary Process}$$

$$x_t = \alpha + x_{t-1} + \varepsilon_t$$

$$-1 < \rho < 1 \rightarrow \text{Trend-Stationary Process}$$

ということになる。したがって、 $-1 < \rho < 1$ という仮説に対して、 $\rho = 1, \beta = 0$ という仮説を検定しなければならない。ここでの問題は、 $\rho = 1$ のときに通常の最小二乗分布理論を用いることができないということである。Dickey-Fuller によると、 ρ の最小二乗推定は DSP 仮説の下で 1 の回りには分布せず、むしろ 1 より小さな値の回りを分布する。しかしながら、マイナス方向の偏りは、サンプル数が増加するにつれて消滅する。Dickey-Fuller は、この仮説検定のための有意点を表にしている。Nelson and Plosser (1982) は、US 経済の長期系列に Dickey-Fuller 検定を適用し、失業率以外の全ての系列において DSP 仮説が受け入れられることを見だしている。彼らは、多くの経済時系列には DSP モデルがより適切であり、(2) の誤差 u_t が高い自己相関があると仮定する場合にのみ、TSP モデルが適切となる、とした。

検定を実施する際には(6)式の両辺から x_{t-1} を引いた形の

$$\Delta x_t = \alpha + (\rho - 1) x_{t-1} + \beta t + \varepsilon_t \quad (7)$$

を推定するが、より高次の自己回帰係数の存在も考慮した検定ができるように

⁴⁾ C.R. Nelson and C.I. Plosser, 'Trend and Random Walks in Macroeconomic Time Series,' *Journal of Monetary Economics*, Vol. 10, pp. 139-162, 1982.

⁵⁾ D.A. Dickey and W.A. Fuller, 'Distribution of Estimators for Autoregressive Time Series with a Unit Root,' *Journal of the American Statistical Association*, Vol. 74, pp. 427-431, 1979.

$$\Delta x_t = \alpha + (\rho - 1) x_{t-1} + \beta t + \sum \rho_j x_{t-j} + \varepsilon_t \quad (8)$$

を用いる Augmented Dickey Fuller 検定⁶⁾ も考案されている。

結合帰無仮説

$$H_0: \beta = 0, \rho = 1$$

に対する検定統計量は、通常の F 統計量である。

$$F = \frac{SSE_R - SSE_U}{J(SSE_U) / (d.f.)}$$

ここで、 SSE_U は(6)あるいは(7)式の係数制約のない場合の OLS 推定から得られる誤差の二乗和であり、 J は仮説の数、 $d.f.$ (自由度) = $T - 3 - n$ 、そして SSE_R は帰無仮説が真ならば保たれるところの $\beta = 0$ とした係数制約付きの OLS 推定から得られる誤差の二乗和である。検定統計量は通常の F 分布にしたがうが、臨界値は通常の $F(J, T - 3 - n)$ 分布からは得られない。これについては、Dickey and Fuller (1981) の p.1063 に表が示されている。

③ 見せかけの相関(Spurious regression)⁷⁾

非定常過程の系列同士においては、互いに無関係であっても高い相関関係を示す場合がある。たとえば、 ε_t と η_t が互いに独立であれば

$$x_t = y_t + \varepsilon_t$$

$$y_t = x_t + \eta_t$$

で生成される系列 x_t と y_t は互いに独立な確率変数のはずである。しかしながら、回帰分析

$$y_t = \alpha + \beta x_t + u_t$$

を OLS でおこなうと、 $\beta = 0$ という帰無仮説を棄却する可能性が高くなり、同時に誤差項の系列相関を検定するダービン・ワトソン係数も小さくなるのである。したがって、トレンドをもった経済系列同士で回帰分析を行って、ダービン・ワトソン係数が小さくなるような場合には、決定係数が高くても「見せかけの相関」の存在を考慮する必要がある。

(3) 伝統的な時系列分析

時系列分析には、大きく分けて、周波数領域(frequency-domain)による方法と時間領域(time-domain)による方法の 2 つがある。前者による手法はスペクトル解析と呼ばれている。ここでは、後者の時間領域による方法について説明し、事例を示すことにする。

⁶⁾ D.A. Dickey and W.A. Fuller, 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root,' *Econometrica*, Vol. 49, pp.1057-1072, 1981.

⁷⁾ C.W.J. Granger and P. Newbold, 'Spurious Regressions in Econometrics,' *Journal of Econometrics*, Vol. 2, pp.111-120, 1974.

① 時系列モデル

a. 純粋なランダム過程

これは x_t という時系列が相互に独立で同一の分布をもった確率変数であることを意味している。すなわち、平均値と分散が一定で、自己共分散関数: $\gamma(k)$ は、

$$\gamma(k) = \text{cov}(x_t, x_{t+k}) = 0 \quad \text{for } k \neq 0$$

であり、自己相関関数は

$$\rho(k) = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases}$$

となる。これはホワイト・ノイズとも呼ばれている。

b. ランダム・ウォーク

これは株価を記述するときにしばしば用いられるモデルである。 $\{\varepsilon_t\}$ を平均が μ で分散が σ^2 という純粋なランダム過程とする。このとき、 $\{x_t\}$ が

$$x_t = x_{t-1} + \varepsilon_t \quad (9)$$

という過程にしたがうならば、ランダム・ウォークという。このプロセスは、 x_t の初期値を $x_t = \varepsilon_t$ と与えると、

$$x_t = \sum_{i=1}^t \varepsilon_i$$

と解ける。これより

$$E(x_t) = t\mu$$

$$\text{Var}(x_t) = t\sigma^2$$

であることがわかり、 x_t の平均も分散も時間と共に変化するので、このプロセスは非定常であることになる。ただし、この一階差分は定常過程となる。

c. 移動平均過程(Moving-Average Process)

$\{\varepsilon_t\}$ を平均が μ で分散が σ^2 という純粋なランダム過程とする。このとき、 $\{x_t\}$ が

$$x_t = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_m \varepsilon_{t-m} \quad (10)$$

は、 m 次の移動平均過程といわれ、 $MA(m)$ と示される。全ての ε は観察できない変数であるので、 $\beta_0 = 1$ となるように基準化しても一般性を失わない。

$$E(\varepsilon_t) = 0 \quad \text{for all } t$$

であるので

$$E(x_t) = 0$$

となる。また、 ε_t は系列相関はないので、

$$\text{Var}(x) = \left[\sum_{i=0}^m \beta_i^2 \right] \sigma^2$$

となる。

MA(m)は、ラグ・オペレータを使うと、

$$\begin{aligned}x_t &= (1 + \beta_1 L + \beta_2 L^2 + \cdots + \beta_m L^m) \varepsilon_t \\ &= \beta(L) \varepsilon_t\end{aligned}$$

と表せる。このLのm多項式はm個の根をもつので、

$$x_t = (1 - \pi_1 L) (1 - \pi_2 L^2) \cdots (1 - \pi_m L^m)$$

とも表現可能で、ここで $\pi_1, \pi_2, \dots, \pi_m$ は、方程式

$$Y^m + \beta_1 Y^{m-1} + \cdots + \beta_m = 0$$

の根である。このモデルを推定すると、 $[\beta(L)]^{-1}$ が収束するという仮定の下で誤差 $\varepsilon_t = [\beta(L)]^{-1} x_t$ を計算することができる。この収束条件は、可逆性条件(invertibility condition)と呼ばれており、 $-1 < \pi_i < 1$ であることが必要十分条件である。このことは、

$$1 + \beta_1 z + \beta_2 z^2 + \cdots + \beta_m z^m = 0$$

というzについてのm次方程式の根が、すべて単位円の外側にあることと同値である。この場合、MA(m)モデルは、AR(∞)モデルとしても記述できるのである。

例としてm=2の場合を考えてみる。すなわち、MA(2)モデルは、ラグ・オペレータを使って、

$$x_t = (1 + \beta_1 L + \beta_2 L^2) \varepsilon_t$$

と表せる。これは

$$x_t = (1 - \pi_1 L) (1 - \pi_2 L) \varepsilon_t$$

とも書くことができ、ここでの π_1 と π_2 は、特性方程式

$$Y^2 + \beta_1 Y + \beta_2 = 0$$

の根である。収束する条件： $-1 < \pi_i < 1$ より

$$\left| \frac{-\beta_1 \pm [\beta_1^2 - 4\beta_2]^{1/2}}{2} \right| < 1$$

となり、したがって、

$$\beta_1 + \beta_2 > -1$$

$$\beta_2 - \beta_1 > -1$$

$$-1 < \beta_2 < 1$$

$$\Leftrightarrow \beta_2 = \pi_1 \cdot \pi_2$$

このような移動平均モデルは、しばしば時系列データにおいてトレンドを除去するのに用いられる。たとえば、

$$x_t = a_0 + a_1 t + a_2 t^2 + \varepsilon_t$$

において、

$$\Delta x_t = x_t - x_{t-1}$$

$$= a_1 + 2a_2 t - a_2 + \varepsilon_t - \varepsilon_{t-1}$$

$$\Delta^2 x_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$$

$$= 2a_2 + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}$$

のように2階差分によってトレンドを除去でき、移動平均モデルによって記述されることになる。

ところで、移動平均モデルの次数、すなわち「m」の決定については、自己相関関数 $\rho(k)$ の値が $k > m$ のところでは0になるという性質を利用する。

d. 自己回帰過程(Autoregressive Process)

ε_t を前と同じ前提で

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_r x_{t-r} + \varepsilon_t$$

によって示されるモデルを r 次の自己回帰モデル、AR(r)という。この式をラグ・オペレータを使って表すと、

$$\begin{aligned} x_t &= \frac{1}{1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_r L^r} \varepsilon_t \\ &= \frac{1}{(1 - \pi_1 L)(1 - \pi_2 L) \cdots (1 - \pi_r L)} \varepsilon_t \end{aligned}$$

となり、ここで $\pi_1, \pi_2, \dots, \pi_r$ は、方程式

$$Y^r + \alpha_1 Y^{r-1} + \cdots + \alpha_r = 0$$

の根である。

自己回帰モデルAR(r)の次数 r は、Yule-Walker方程式を順次解いて得られる偏自己相関関数の値がほとんど0となる前のところの次数を選択する。すなわち、自己回帰過程が r 次であるとすると、偏自己相関係数の値は r 次より高い次数ではほとんど0か0に近い値となる。

e. 自己回帰移動平均過程(Autoregressive Moving-Average Process)

ARモデルとMAモデルを結合したモデルは、ARMAモデルと呼ばれている。一般的に、ARMA(p, q)モデルは、

$$x_t = \alpha_1 x_{t-1} + \cdots + \alpha_p x_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \cdots + \beta_q \varepsilon_{t-q}$$

と定義される。ラグ・オペレータを用いて表現すると、

$$\phi(L)x_t = \theta(L)\varepsilon_t$$

となり、ここで

$$\phi(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_p L^p$$

$$\theta(L) = 1 + \beta_1 L + \beta_2 L^2 + \cdots + \beta_q L^q$$

である。

f. 階差型自己回帰移動平均過程(Autoregressive Integrated Moving-Average Process)

経済で用いる時系列データは非定常であることが多く、その場合は原系列の階差をとり、定常過程に変換した自己回帰移動平均のモデルを用いる。

$$\Delta = 1 - L$$

と定義すると、

$$\Delta x_t = x_t - x_{t-1}$$

$$\Delta^2 x_t = (x_t - x_{t-1}) - (x_{t-1} - x_{t-2})$$

.....

のように表現できる。

いま、 $\Delta^d x_t$ で定常系列が得られたとすると、ARMA(p,q)で表せる $\Delta^d x_t$ は、ARIMA(p, d,q)と呼ばれる。

どこまでの階差あるいは差分をとるかという問題は、どこまでの差分をとった系列が定常状態となったと判断できるか、ということである。自己相関関数が、幾何級数的に逓減していくようなパターンとなるまで差分をとる、すなわちそこでdを決める。

② 時系列モデルの推定方法

ARモデルの推定は、誤差の自乗和を最小にするOLSで行うことができる。しかしながら、MAモデルについては、ARモデルのような誤差の自乗和を単純に計算することはできない。移動平均誤差の共分散行列に制約を付け最尤推定法を適用する方法がOsborne (1976) やDavidson (1981) によって提示されているが、ここではBox-Jenkins (1976) によるgrid search procedure についてMA(2)のモデル

$$x_t = \mu + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2}$$

を例にとって示す。

任意の初期値、たとえば、 $\mu = x_0$ 、 $\varepsilon_0 = \varepsilon_{-1} = 0$ を与えると、

$$\varepsilon_1 = x_1 - \mu$$

$$\varepsilon_2 = x_2 - \mu - \beta_1 \varepsilon_1$$

$$\varepsilon_t = x_t - \mu - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} \quad t \geq 3$$

のように連続的に誤差 $[\varepsilon_t]$ を求めることが可能となり、この自乗和を最小にするように $[\beta_1, \beta_2]$ を求めることができる。

ARMAモデルの推定については、

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2}$$

をラグ・オペレータを使って表すと、

$$(1 - \alpha_1 L - \alpha_2 L^2) x_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2}$$

あるいは

$$x_t = \frac{1}{1 - \alpha_1 L - \alpha_2 L^2} (\varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2})$$

となる。ここで、

$$z_t = \frac{1}{1 - \alpha_1 L - \alpha_2 L^2} \varepsilon_t$$

とおき、この式の両辺に $[1 - \alpha_1 L - \alpha_2 L^2]$ を掛けてやると、

$$z_t - \alpha_1 z_{t-1} - \alpha_2 z_{t-2} = \varepsilon_t$$

を得る。一方、

$$x_t = z_t + \beta_1 z_{t-1} + \beta_2 z_{t-2}$$

あるいは

$$z_t = x_t - \beta_1 z_{t-1} - \beta_2 z_{t-2}$$

である。

grid search の手続きは、初期値を $z_0 = z_{-1} = 0$ とし、収束条件を満たす任意の $[\beta_1, \beta_2]$ に対して、

$$z_1 = x_1$$

$$z_2 = x_2 - \beta_1 z_1$$

$$z_t = x_t - \beta_1 z_{t-1} - \beta_2 z_{t-2}$$

のように z の連続的数値を得ることができる。そして、パラメータ (α_1, α_2) を推定するために、この z の値を用いて OLS を実行できる。

パラメータ $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ の推定値が得られると、先の

$$z_t - \alpha_1 z_{t-1} - \alpha_2 z_{t-2} = \varepsilon_t$$

から誤差を

$$\varepsilon_t = z_t - \alpha_1 z_{t-1} - \alpha_2 z_{t-2}$$

のように求めることができる。この誤差は、後で述べる ARMA モデルによる予測の際に用いられる。

③ ボックス・ジェンキンスのアプローチ

ARMA モデルによる時系列分析は通常、ボックス・ジェンキンスのアプローチといわれている。これは、彼らがこのモデルの創始者だという訳ではなく、このモデルを一般に広めた功績によってそう呼ばれているのである。

a. 推定の手順

1. 定常状態を得るために差分をとる。どの程度までの差分をとるかは、通常はコレログラム (correlogram) を見て判断する。コレログラムとはラグ k の自己相関係数 $\rho(k)$ を k に対してプロットしたグラフである。これが大きく低下するところの差分階数で定常状態が得られる。
2. AR 要素と MA 要素の次数を決める。MA の次数はコレログラムがある点から以降はゼロになり、AR の次数は幾何級数的に遞減していく。ARMA の場合は、コレログラムがしばらく

くして大きく低下するパターンを示す。

3. ある暫定的なモデルが2. によって得られると、ARMAモデルを推定する。
4. モデルの適合性をチェックする。Ljung and Box (1978) は、(推定された) 残差の全ての階数の自己相関でチェックすることを示唆している。ARMA(p,q)モデルにおいて、これが適切に定式化されているならば、 r_k をサンプルNでラグkの自己相関係数とおくと、

$$Q = N(N+1) \sum_{k=1}^K r_k^2 / (N-k)$$

が、自由度が $[K-p-q]$ の χ^2 分布にしたがう。

5. モデルの定式化が妥当であれば、予測のステップに進むが、そうでなければ再び2. のステップに戻り、モデルの再定式化をおこなう。

b. 予測手法

ここでは簡略化のため、ARMA (2, 2) のモデルについて考える。 $\{x_t\}$ についてn個の観測値があり、k期先の値： x_{n+k} を予測するものとする。この予測値を、 $\hat{x}_{n,k}$ と表す。最初の添え字のnは予測がなされるスタート時点を意味し、後の添え字は予測される先の期間を意味している。

まず、 $k=1$ から出発する。時点nにおいて x_{n+1} の予測値を求める。

$$x_{n+1} = \alpha_1 x_n + \alpha_2 x_{n-1} + \varepsilon_{n+1} + \beta_1 \varepsilon_n + \beta_2 \varepsilon_{n-1}$$

の式において、 ε_{n+1} をその期待値： $E(\varepsilon_{n+1}) = 0$ に置き換えると、 x_{n+1} の予測値 $\hat{x}_{n,1}$ が、

$$\hat{x}_{n,1} = \hat{\alpha}_1 x_n + \hat{\alpha}_2 x_{n-1} + \hat{\beta}_1 \hat{\varepsilon}_n + \hat{\beta}_2 \hat{\varepsilon}_{n-1}$$

と求まる。ここで、 $\hat{\cdot}$ は推定値であることを示している。次に、 $k=2$ の場合は、

$$x_{n+2} = \alpha_1 x_{n+1} + \alpha_2 x_n + \varepsilon_{n+2} + \beta_1 \varepsilon_{n+1} + \beta_2 \varepsilon_n$$

の式において、 ε_{n+2} と ε_{n+1} をそれぞれその期待値と置き換え、さらに $k=1$ で予測された $\hat{x}_{n,1}$ を用いて、

$$\hat{x}_{n,2} = \hat{\alpha}_1 \hat{x}_{n,1} + \hat{\alpha}_2 x_n + \hat{\beta}_2 \hat{\varepsilon}_n$$

によって x_{n+2} の予測値を求める。以下、この手順を繰り返していく。

(4) 最近の時系列分析—VAR、単位根、共和分—

Box-Jenkinsのアプローチは20年近く前に開発されたものであり、これは時系列の定常系列を得るための変換とARMAモデルの推定に力点がおかれている。しかしながら、差分をとることによって、長期の動向に関する貴重な情報が失われている可能性を含んでいる。

ここでは、非定常系列を分析するのに、この10年間に新たに展開された3つのトピックスについて解説する。1つ目はVARモデルで、これは複数の時系列間の相互関係を分析する出発点として非常に有用なモデルである。2つ目は単位根であり、そこではBox-Jenkinsのコレログラムを見することで時系列データの差分をとるという方法が、単位根の検定という形で定式化される。単位根

の考えは、一階差分で定常系列になる非定常時系列を研究するものである。3つ目は共和分である。共和分の理論は、変数間の長期的トレンドの間の関係、これはBox-Jenkinsでは差分で取り去られたものであるが、これを説明しようとするものである。

① ベクトル自己回帰

いくつかの時系列変数があるとき、それら相互の依存関係を考慮する必要がある。このための1つの方法は、全ての変数にラグをもった同時方程式体系を推定することである。このようなモデルは、「dynamic simultaneous equations model」と言われている。しかしながら、この定式化には2段階の手順がある。まず、変数を外生変数と内生変数に分類する。第2に、識別するために何らかの制約をパラメータに付加する。Sims (1980)によると、この2段階手法は多くの恣意的な決定を含むので、その代わりにVAR(vector autoregressive)アプローチを示唆している。これは、まさにARモデルの複数変数時系列への一般化に他ならない。VARモデルはOLS手法が適用できるので推定が非常に簡単である。

y_{1t} と y_{2t} の2つの時系列変数を考える。貨幣供給量の増加率とGNPデフレータの上昇率であるとして、一期のラグをもったモデルは、定数項を省略して、

$$y_{1t} = \alpha_{11} y_{1,t-1} + \alpha_{12} y_{2,t-1} + \varepsilon_{1t}$$

$$y_{2t} = \alpha_{21} y_{1,t-1} + \alpha_{22} y_{2,t-1} + \varepsilon_{2t}$$

行列表示だと、

$$\begin{bmatrix} 1 - \alpha_{11}L & \alpha_{12}L \\ -\alpha_{21} & 1 - \alpha_{22}L \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

この解は、

$$\begin{aligned} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} &= \begin{bmatrix} 1 - \alpha_{11}L & \alpha_{12}L \\ -\alpha_{21} & 1 - \alpha_{22}L \end{bmatrix}^{-1} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \\ &= \frac{1}{\Delta} \begin{bmatrix} 1 - \alpha_{22}L & \alpha_{12}L \\ -\alpha_{21} & 1 - \alpha_{11}L \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \end{aligned}$$

ただし、

$$\begin{aligned} \Delta &= (1 - \alpha_{11}L)(1 - \alpha_{22}L) - (\alpha_{12}L)(\alpha_{21}L) \\ &= 1 - (\alpha_{11} + \alpha_{22})L + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})L^2 \\ &= (1 - \lambda_1L)(1 - \lambda_2L) \end{aligned}$$

ここで、 λ_1 と λ_2 は

$$\lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}) = 0$$

の根である。

安定性の条件が満たされると、 y_{1t} と y_{2t} を ε_{1t} と ε_{2t} のカレントとラグ変数でもって表現できることになる。これは、インパルス・レスポンス(応答)関数と言われている。 y_{1t} について

だと、

$$\begin{aligned} y_{1t} &= \Delta^{-1} [(1 - \alpha_{22}L)\varepsilon_{1t} + \alpha_{12}L\varepsilon_{2t}] \\ &= \varepsilon_{1t} + \alpha_{11}\varepsilon_{1,t-1} + (\alpha_{11}^2 + \alpha_{12}\alpha_{21})\varepsilon_{1,t-2} + \dots \\ &\quad + \alpha_{12}\varepsilon_{2,t-1} + \alpha_{12}(\alpha_{11} + \alpha_{22})\varepsilon_{2,t-2} + \dots \end{aligned}$$

2変数で互いに1期のラグをもった単純なモデルを例にVARモデルを説明した。しかし、実際には、移動平均誤差を考慮していないので、自己回帰はおそらく予測に対してもっと多くのラグを持たなければならないであろう。そうでなければ、一変量のARMAの方が望ましいといえる。たとえば、4つの変数のそれぞれに対して6つのラグをもったモデルを考えると、各方程式は24のパラメータを推定しなければならなくなり、全体で96個のパラメータを推定する必要がある。このようなパラメータ制約を付けないVARモデルは予測に対して非常に有用であるとは今日まで見いだされておらず、VARモデルに何らかの制約を付加することが示唆されてきている。

予測において特に有用であると見つけられてきたものの1つにBayesian Vector Autoregressive(BVAR)がある。これはVARモデルの係数に先験的な分布を置くものである。⁸⁾ Littermanによると、BVARは同時体系の構造モデルよりも予測に関して優れていることが示されている。また、VARモデルに対して経済理論から先験的な制約をおくStructural Vector Autoregressive(SVAR)が、Blanchard等によって提示されている。⁹⁾

② 単位根の問題

もし系列がDSP(Difference Stationary Process)ならば、何等かのショックの影響は永久的であることがわかる。たとえば、

$$y_t = y_{t-1} + \varepsilon_t$$

において、ある時点 y_T で ε_T にCのジャンプが生じたとする。すると、その後の系列である $y_T, y_{T+1}, \dots$ もすべてCだけ増加する。したがって、ショックCの効果は永続的である。

もし、系列が

$$y_t = \alpha y_{t-1} + \varepsilon_t, \quad -1 < \alpha < 1$$

ならば、ショックは時間と共に減衰していくことになる。このように y_{t-1} の係数が1であるか

⁸⁾ R.B.Litterman, 'Forecasting with Bayesian Vector Autoregression: Five Years of Experience,' *Journal of Business and Economic Statistics*, Vol.4, pp.28-35, 1986.

⁹⁾ O.J.Blanchard, 'A Traditional Interpretation of Macroeconomic Fluctuations,' *American Economic Review*, Vol.79, pp.1146-1164, 1989.

O.J.Blanchard and D.Quah, 'The Dynamic Effects of Aggregate Demand and Supply Disturbances,' *American Economic Review*, Vol.79, pp.655-673, 1989.

どうか、すなわち単位根をもつかどうかは、経済政策においても非常に重要な意味をもってくる。

単位根の検定で、まず、

$$y_t = \alpha y_{t-1} + \varepsilon_t$$

というモデルを考える。ここで、 ε_t はホワイト・ノイズである。ランダム・ウォーク ($\alpha = 1$) のとき、この方程式のOLS推定値は0の方向へ偏りをもった α の推定値を生じる。しかしながら、 α が1より小さくて1に近いときにもまた、OLS推定値は0の方向へ偏りをもった α の推定値を生じるのである。

Dickey-Fuller 検定を考えるために、次のモデルを考えよう。

$$y_t = \beta_0 + \beta_1 t + u_t$$

$$u_t = \alpha u_{t-1} + \varepsilon_t$$

ここで、 ε_t は平均0の定常過程である。これらの式の誘導形は

$$y_t = \gamma + \delta t + \alpha y_{t-1} + \varepsilon_t \quad (II)$$

となる。ここで、

$$\gamma = \beta_0 (1 - \alpha), \quad \delta = \beta_1 (1 - \alpha)$$

である。この方程式は、 $\alpha = 1$ ならば単位根をもつといわれる。

Dickey-Fuller 検定は、 ε_t がホワイトノイズ誤差であるという仮定の下で(II)における $\alpha = 1$ という仮説を検定することに基づいたものである。そこでは3つの検定統計量が存在する。

$$K(1) = T(\hat{\alpha} - 1)$$

$$t(1) = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})}$$

$$F(0, 1)$$

ここで、 $\hat{\alpha}$ は(10)式のOLS推定値であり、 $SE(\hat{\alpha})$ は $\hat{\alpha}$ の標準誤差である。また、 $F(0, 1)$ は、(II)における結合仮説 $\delta = 0$ と $\alpha = 1$ を検定するための通常のF統計量である。これらの統計量は標準正規、t、F分布をもたない。K(1)とt(1)に対する臨界値はFuller (1976)¹⁰⁾の $\delta = 0$ の場合に表にされている。F(0, 1)に対する臨界値は、Dickey and Fuller (1981)に表で示されている。

③ 共和分(Cointegration)

a. 共和分の定義

経済時系列の多くが単位根をもつ非定常過程であると、それらは互いに完全に独立な動きを

¹⁰⁾ W.A.Fuller, *Introduction to Statistical Time Series*, Wiley, New York, 1976.

見せる場合もあるが、所得と消費や輸出と輸入のように長期的には同一方向に着かず離れずといった動きを見せる場合がしばしば存在する。Granger (1981)¹¹⁾は、複数の変数が互いに乖離することを防ぐように推移していく変数間の関係が存在する場合、そのような特殊な関係を共和分(Cointegration)と名付けた。

計量経済学における重要なことの1つは、短期の変動を長期均衡において表現し統合することの必要性である。短期の不均衡をモデル化する伝統的なアプローチは、部分調整モデルである。このモデルを拡張したものに誤差修正モデル(ECM: Error Correction Model)がある。これは、過去の期間の不均衡もまた組み込んだモデルである。短期変動の動学分析は、まず差分をとるなどして変数のトレンドを除去することから始められる。しかしながら、このプロセスでは長期における変数間の関係について貴重な情報を失うことをも意味している。共和分(cointegration)の理論は、Granger (1981)によって考案され、Engle and Granger (1987)¹²⁾によって精緻化されたもので、短期変動の動学を長期均衡で表現、統合する手法である。

Δy_t が定常的な時系列であるならば、 y_t は一階の和分(integrated)あるいは $I(1)$ と書ける。したがって、定常時系列は $I(0)$ と表せる。ランダム・ウォークは、 $I(1)$ 系列の特殊例である。なぜならば、もし y_t がランダム・ウォークであるならば、 Δy_t はランダム系列かホワイト・ノイズであるからである。ホワイト・ノイズは定常系列の1つの特殊例である。もし Δy_t が $I(1)$ にしたがうならば、 y_t は $I(2)$ となる。また、もし $[y_t \sim I(1)]$ かつ $[u_t \sim I(0)]$ ならば、それらの合計は、 $[Z_t = y_t + u_t] \sim I(1)$ となる。

Granger and Weiss (1983)によると、共和分(cointegration)の定義は、【時系列データ x_t と y_t について、 $y_t \sim I(1)$ であり $x_t \sim I(1)$ とする。このとき、 $[y_t - \beta x_t] \sim I(0)$ であるような係数 β が存在するならば、 y_t と x_t は共和分の関係にある】といわれる。このことは、 y_t と x_t は $CI(1, 1)$ ということによって示される。より一般的には、

$$y_t \sim I(d) \text{ かつ } x_t \sim I(d)$$

であり、

$$b > 0 \text{ で } [y_t - \beta x_t] \sim I(d-b)$$

ならば

$$\{y_t, x_t\} \sim CI(d, b)$$

である。したがって、共和分の関係にあるときには、 y_t と x_t の間には長期の均衡関係が存在することになる。

¹¹⁾ C.W.J.Granger, 'Some Properties of Time Series Data and Their Use in Econometric Model Specification, *Journal of Econometrics*, Vol.16, pp.121-130, 1981.

¹²⁾ R.F.Engle and C.W.J.Granger, 'Co-Integration and Error Correction: Representation, Estimation, and Testing,' *Econometrica*, Vol.55, pp.251-276, 1987.

このことが意味していることは、回帰方程式

$$y_t = \beta x_t + u_t \quad (12)$$

は、 x_t と y_t が時間の経過につれてお互いにそうは離れて徘徊しない、という理由から意味をもつということである。だから、それらの間には長期的均衡関係が存在するのである。もし x_t と y_t が共和分関係にない、すなわち、 $[y_t - \beta x_t = u_t] \sim I(1)$ なら、それらは時間と共に離れて徘徊していく。したがって、この場合の両者の間には長期均衡関係は存在しない。このとき、 y_t を x_t に回帰して得られる y_t と x_t の関係は、「見せかけ」である。

Box-Jenkinsの手法においては、時系列が非定常的ならば定常となるまでその差分（階差）をとりARMAモデルをフィットさせることをおこなう。時系列が y_t と x_t のように2つの場合も同様である。この差分をとるという操作は、系列におけるトレンドや長期的な動きを取り除いてしまう。しかしながら、ここで関心のあることは、 y_t と x_t におけるトレンド相互の関係を説明することである。このことは y_t の x_t への回帰をすることによってでもできるが、この回帰は長期的な安定的関係が存在しないならば意味をなさない。 y_t と x_t が共和分関係にあるかどうかという質問をすることは、 y_t と x_t におけるトレンド間に何らかの長期的な安定関係があるか否かを尋ねていることなのである。

季節調整の場合も同様である。 y と x から季節要因を取り除いてそれを分析するよりも、 y と x において相互に季節性があるかどうかを尋ねてもよい。この場合は、一階差分を意味しているのではなく、月次データなら12階差分を意味しているのである。

共通のトレンドを言うとき、deterministic と stochastic の区別をしなければならない。時間に対して回帰分析をおこないトレンドを除去することはdeterministic trend を仮定しており、差分をとることはstochastic trendを仮定しているのである。「共和分」の概念は、common stochastic trend の考えに言及している。しかし、これは唯一の共通トレンドではない。common deterministic trends を考えることもできる。

deterministic と stochastic trends が共に存在しているモデルで

$$x_t = T_t + S_t + \mu_t + \eta_t + \varepsilon_t$$

ここで、 T_t ：時間の多項式で表現される deterministic trend

S_t ：季節ダミーなどで表現される deterministic seasonal

μ_t ： $I(1)$ 過程で表現できるような stochastic trend

η_t ：四半期データだと $(1-L^4)$ が定常となるような stochastic seasonal

と表現できる。

b. 共和分の存在検定

y_t と x_t がともにランダム・ウォークとすると、それらはともに $I(1)$ である。そして、一階差分の式、

$$\Delta y_t = \alpha \Delta x_t + \lambda (y_t - \beta x_t) + \nu_t \quad (13)$$

は妥当な方程式である。なぜなら、 $\Delta y_t, \Delta x_t, (y_t - \beta x_t), v_t$ は、すべて $I(0)$ だからである。方程式(12)は y_t と x_t の間の長期的関係が考慮されており、方程式(13)は短期的動学を記述している。Engle and Granger (1987) は(11)をOLSで推定し、 β の推定値を(13)に代入し、パラメータ α と λ を推定するという2段階推定を示唆している。しかしながら、この2段階の手続きは、 y_t と x_t が共和分の関係にあるという仮定に依っている。それ故、より一層、共和分に対する検定が重要になってくるのである。

Engle and Granger は(12)をOLSで推定し、残差 u_t を得、そしてDickey-Fuleer の検定をその残差に対して適用するのであるが、この検定が帰するところは、

$$u_t = \rho u_{t-1} + e_t$$

において、 $\rho = 1$ という仮説すなわち、

$$H_0: u_t \sim I(1)$$

を検定することである。

本質は、 y_t と x_t は「not cointegrated」であるという帰無仮説を検定することである。 $y_t \sim I(1)$ で $x_t \sim I(1)$ であり、だから u_t が $I(1)$ にならないということを見ようとしていることに気を付けよう。

以上のことからEngle and Granger (1987) の方法によると、その手順は以下のようになる。

1. x_t と y_t が共に $I(1)$ であることをチェックするために単位根検定を適用する
2. y_t を x_t に回帰して残差項、 $\hat{u}_t = y_t - \beta x_t$ を計算する
3. $\hat{u}_t$ に対して単位根検定を適用する。

もし、 x と y が共和分の関係にあれば、 $u = y - \beta x$ は $I(0)$ となる。他方、共和分の関係になれば、 u は $I(1)$ となろう。単位根検定は u に対して適用されるので、帰無仮説は、

$$H_0: u \text{ は単位根をもつ、あるいは、} x \text{ と } y \text{ は共和分の関係にない}$$

$$H_1: x \text{ と } y \text{ は共和分の関係にある。}$$

ところが、Bewley (1979), Wickens and Breusch (1988)¹³⁾ によって示されたように、Engle and Granger のこの2段階推定方法は必要がないともいえる。これは短期と長期のパラメータを同時に推定することで、 β のより効率的な推定値を得ることができるからである。(12)を変形して、

¹³⁾ R.A.Bewley, 'The Direct Estimation of the Equilibrium Response in a Linear Dynamic Model,' *Economics Letters*, Vol.3, pp.357-361, 1979.

M.R.Wickens and T.S.Breusch, 'Dynamic Specification, the Long Run and the Estimation of Transformed Regression Models,' *Economic Journal*, Supplement, pp.189-205, 1988

$$y_t = \beta x_t + \frac{1}{\lambda} \Delta y_t - \frac{\alpha}{\lambda} \Delta x_t - \frac{u_t}{\lambda} \quad (14)$$

を得る。 Δy_t は誤差項 u_t と相関しているであろうから、(14)は操作変数法で推定せねばならない。 Δy_t と Δx_t の係数は短期の変動を記述している。もし、 $y_t \sim I(1)$ で $x_t \sim I(1)$ であると、 Δy_t と Δx_t は u_t のように $I(0)$ である。短期変動についての誤った定式化は、長期的パラメータの推定値に対して多くの影響は及ぼさない。たとえば、(14)において、たとえ Δx_t が除かれていても、 β の推定値は一致推定値である。このことの直感的な理由は、 Δy_t と Δx_t は y_t と x_t よりも低次の階数であるからである。

c. 共和分とVARモデル

時系列変数間の共和分の存在は、VARモデルの構造推定にも重要な意味をもってくる。たとえば、変数間に共和分の関係が存在しているに関わらず、それを考慮しないで構築したVARモデルは定式化の誤りを犯していることになる。したがって、VARモデルの各変数に対して単位根の検定をおこない、その結果から差分（階差）をとった定常系列で分析するときには、それらの間の共和分の存在を確かめる必要がある。

VARモデルにおけるいくつかの変数が相互に共和分の関係にあるならば、VARモデルにおけるパラメータ制約を課すことができる。制約のないVARモデルよりも予測の精度が高いことが示されており、今後はBVARモデルとの比較検討が必要である。

(5) おわりに

本章では、経済時系列データの取扱いについて、今日では伝統的ともいえるBox-Jenkinsのアプローチとそれ以降の最近までの計量経済学の理論的進展について解説をおこなった。

今後、それらを生かして地域経済指標の動向を正しく把握し的確に予測して政策に生かすためには、今後の展望として、以下のような課題が考えられよう。

まず、モデルの構造的問題として、計量経済モデルについては、時系列データの充実にともなって、トレンドをもった時系列変数間の関係をここで述べたような最近の計量経済理論の成果を積極的に適用して、より現実に近い因果関係をもったモデルを構築していく努力が必要である。ニューラル・ネットワークを適用したモデルについては、それが計量経済学とは異なりパラメトリックな推定をとまわらないため、ラグ次数の決定など予測に際して用いる変数間の因果関係の構造について、何らかの経済学的な行動仮説のようなバックグラウンドを見いだす必要がある。3年間の本調査研究では、内外で行われているニューラル・ネットワーク理論の社会科学への適用研究に比べて、予測モデルの構造決定および予測手法についてかなりのアドバンテージを持っていることは、様々な関連文献調査やヒアリング等からも立証されよう。しかしながら、今後、より客観的な評価の高いモデルへと展開していこうするならば、このニューラル・ネットワーク理論の適用に際しての経済学的な意味付けは必須となってくる。

次に、地域経済指標の将来予測の精度に関しては、

- ① Boxs-Jenkins型の時系列モデルによる予測
- ② 同時方程式体系での計量経済モデルによる予測
変数間の「共和分」の存在の検定
- ③ 制約付きのVARモデルでの予測
- ④ ニューラル・ネットワークの理論を適用して予測

の比較検討が重要となってくる。すでに過去の本報告書において②と④の比較検討はおこなわれニューラル・ネットワークの理論を適用して予測の精度が非常に高いことも示されたが、①も含めての比較検討をおこなうこと非常に興味深いことである。

最後に、②と③、④の並列的な比較とは別に、ニューラル・ネットワークの理論に本章で述べたような最近の計量経済理論の一部適用、もしくは一部導入の可能性を考えることも1つの課題といえよう。

禁無断転載

平成5年3月発行

発行所 財団法人 日本情報処理開発協会

東京都港区芝公園3丁目5番8号

機械振興会館内

Tel 03-3432-9372

04-R009

