

貼り合せ予測法

国 沢 清 典*

1. はじめに

地域的に広範囲に広がった現象は、これを巨視的に眺めることによって予測に利用できる場合がある。たとえば1人当たり個人県民所得をみると

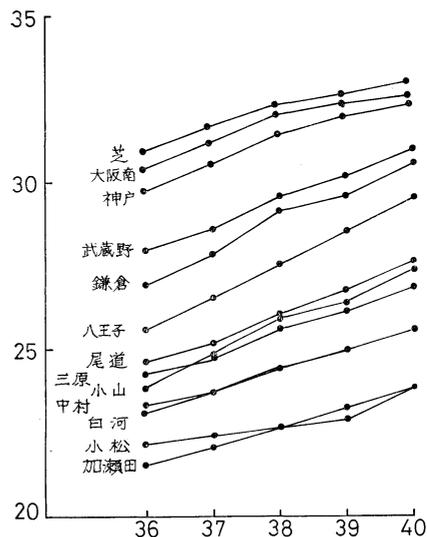
都府県	所得 (円)	年次
東京	3270	(昭和39)
大阪	3254	(昭和40)
神奈川	3301	(昭和42)
兵庫	3241	(昭和43)
石川	3256	(昭和44)
秋田	3223	(昭和45)
茨城	3294	(昭和46)

単位 100 円

となっている。昭和39年の東京 327,000 円、大略1年おくらせて大阪、大阪より2年おくらせて神奈川、神奈川より1年おくらせて兵庫、更に兵庫より1年おくらせて石川、更に翌年が秋田、その翌年が茨城となっている。これはわが国における経済成長の府県別時差を示していることになる。気象については筆者は何も知らないが、恐らくこれに類した現象が見られることであろう。例えば大阪での雨は翌日の東京の雨を意味しているときよくいわれている。このような現象は上の府県別個人所得の例ばかりでなく、GNPの国際比較などでもよく見られる。次の図は新生、いこい、ハイライト、ピース(R)についての喫煙者の販売支所別平均購買価格(10本当りについて喫煙しているたばこの価格)を示したものである。これは全国500余りの支所の中からいくつかを抽出して経年的変化の模様を示したものである。武蔵野の36年は鎌倉の37年、八王子の38年6月に相当していることがわかる。また第1図より消費水準の上昇が都会から地方に及んでいることがよくわかる。

しかし、個々の支所例えば芝が今後どのように伸びてゆくか、また加瀬田の今後はどうなるかを判断する場合、単にこれらの折線を将来に向けて引きのばすにはデータもすくないし、またとてもそのような勇気ももてないであろう。

すくないデータをどのように利用して、将来の予測を



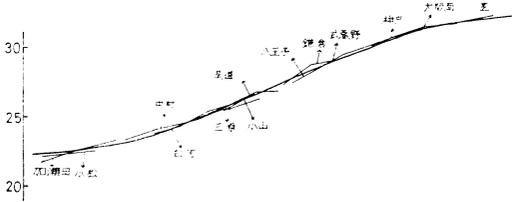
第1図

たてることができるか。最近、道路交通の広域制御システムの開発が検討されているが、これには交通の予測が必要であり、一つの道路の交通量のみのデータを眺めていても、広域制御のための予測はできない、個々の主要道路の交通量を貼り合せながら、巨視的立場に立って次の時刻の予測をする必要がある。このためには、全体としてある一定の傾向を持っていなければならない。交通量の場合ならば、朝のラッシュアワーならば、都心に向う車が圧倒的に多い。夕方になれば都心から放射的に都心を離れる車が多いであろう。このように個々の車は自由なる行動をしているが、全体として統一のある行動をしているかにみえることが大切である。個々の車がバラバラな行動をしておれば、以下の理論は適用されない点を断っておく必要がある。恐らく気象関係の情報にしても、巨視的判断が大切であり、各地域がバラバラの行動をとる現象はすくないと思われる。集中豪雨などがこの類の中に入るかも知れないが。

2. 貼り合せ予測法

第1図についても一度考えてみよう。芝、大阪南、

* K. Kunisawa 東京工業大学



第2図

神戸で一群、武蔵野、鎌倉、八王子で一群、尾道、三原、小山で一群、中村、白河、小松、加瀬田等で一群と分類してみると、それぞれ各群の折線にはそれぞれの特長を持っていることに気付く。いま、これらの折線を平行移動して一本の曲線を作ることができないかと考えてみる。この場合の平行移動は時間軸に平行に移動する。例えば八王子を固定して、それより上位のものは右側に平行移動して、時間的に先行さす。八王子より下位のものには左側に平行移動して時間的に遅行さす。このような操作を行って一本の曲線が得られることに注意する(第2図)。このような操作を貼り合せ法とよんでいる。

しかし上述の方法は如何にも原始的であるので、これをコンピュータにやらせることを考えなければならない。今回は貼り合せの結果、成長曲線が得られる場合について述べてみよう。またこの場合が非常に多いので利用価値もあると思われる。

いま N 地域のデータについて、時点 $t=1, 2, \dots, n$ (n が3または4で十分) に於けるものを

$$X_i(t) \quad i=1, 2, \dots, N; \quad t=1, 2, \dots, n$$

であらわす。次に $X_i(t)$ の逆数 $1/X_i(t)$ をとり、これを x 軸に、 $1/X_i(t+1)$ を y 軸にとり平面上に点

$$\left(\frac{1}{X_i(t)}, \frac{1}{X_i(t+1)} \right)$$

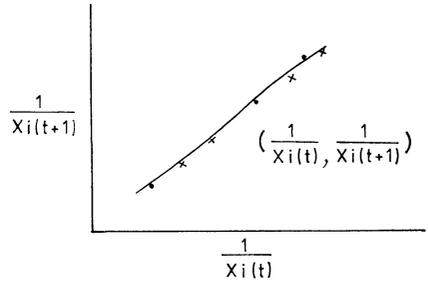
をプロットする。これらの打点をすべての i と t について実施する。その結果、直線に回帰できることが、貼り合せ法によって成長曲線が得られるための必要かつ十分な条件である。つまり

定理 点

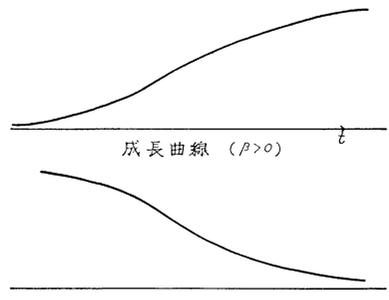
$$\left(\frac{1}{X_i(t)}, \frac{1}{X_i(t+1)} \right)$$

が切片が正、勾配が0と1との間にある直線上に載るならば、 $X_i(t)$ は成長曲線上に載る。また逆も成立する。つぎに点

$$\left(\frac{1}{X_i(t)}, \frac{1}{X_i(t+1)} \right)$$



第3図



第4図

が切片負で、勾配が1より大なる直線上に載るならば、 $X_i(t)$ は衰退曲線(成長曲線を上下逆にした曲線)上に載る。逆もまた真である。

いま

$$C = \frac{1 - \text{勾配}}{\text{切片}}, \quad \beta = -\log(\text{勾配})$$

とおけば、成長曲線(衰退曲線)(第4図)は

$$X(t) = \frac{C}{1 + e^{-\beta t}} \quad \begin{matrix} \beta > 0 \text{ 成長曲線} \\ \beta < 0 \text{ 衰退曲線} \end{matrix}$$

となる。実際にコンピュータに与えるアルゴリズムでは成長であろうと衰退であろうとそれをいちいち判定することなく、

$$\left(\frac{1}{X_i(t)}, \frac{1}{X_i(t+1)} \right)$$

の直線回帰をやればよい。

ここで直線回帰のやり方について述べておこう。 $X_i(t)$ を i 地域を時刻 t のデータとし、

$$\left(\frac{1}{X_i(1)}, \frac{1}{X_i(2)} \right) \quad i=1, 2, \dots, N$$

$$\begin{pmatrix} \frac{1}{X_i(2)}, & \frac{1}{X_i(3)} \\ \vdots \\ \frac{1}{X_i(n-1)}, & \frac{1}{X_i(n)} \end{pmatrix} \quad i=1, 2, \dots, N$$

の合計 nN 個の点がプロットされる。これらを直線回帰してこの回帰直線の勾配 (r)、切片 (S)は

$$\sum_{i=1}^N \sum_{t=1}^n \left(\frac{1}{X_i(t)} r + S - \frac{1}{X_i(t+1)} \right)^2$$

を最小にする(最小二乗法) r, S を求め、これによって定数 C, β は

$$C = \frac{1-r}{S}, \quad \beta = -\log r$$

で求まることとなる。

しかし、 $X_i(t)$ の逆数をとる性格上、 $X_i(t)$ が小さいところでは上記の打点はバラツキが大きくなる傾向があって、これが回帰の定数 r, S の決定に影響する恐れがあるから、次のようなウェイト w を与えて直線回帰する方法が考えられる。このようなウェイト回帰法は従来の回帰に新しい問題を投げかけるものであり、問題によっては、非常に重要となるであろう。ウェイトはここで述べるような単純なものから、もっと多くの周辺因子を取り入れるようなウェイトであってよく、最適なウェイトを発見することはこのような回帰予測の重要な問題となることは明らかである。

さてわれわれの場合は $X_i(t)$ によって定まるウェイト $w(X_i(t))$ を考えて、

$$\sum_{i=1}^N \sum_{t=1}^n w(X_i(t)) \left(\frac{X_i(t)}{1} r + S - \frac{1}{X_i(t+1)} \right)^2$$

を最小にする r, S を求める方法である。このウェイト回帰は實際上、ウェイトのとり方に機動性をもたせ、非常に良好な結果を得ることが多い。

コンピュータのアルゴリズム上注意しなければならない点をもう一つ述べておかねばならない。

$$C = \frac{1-r}{S} > 0$$

ならば文句なしに、成長(衰退)曲線に載るわけであるが、実際にコンピュータにかけた場合には

$$C = \frac{1-r}{S} \leq 0$$

となることも起こりうるわけである。この場合は $S=0$ とおいて、直線回帰の最小二乗法をやりなおすということにするとよい。

3. 予測値をどのように計算するか

目的は各地域 i の過去の t 時刻のデータを積み上げて、全地域一本の各地域の未来値を予測することである。この場合、あまり過去のデータをインプットすると、構造の異なったデータをインプットすることになるからできるだけ最新のデータを各地域について2ないし3位投与すればよい。これが実は貼り合せ予測法の最大の武器である。個々の地域のデータだけでは予測には、データが少なすぎるが、全地域よりかき集められたデータは相当なものになる。毎日、新聞紙上やテレビに出る全国天気図は、ある一つの時点、たとえば午前6時の全国の気圧配置図である。しかしこれを午前4時、5時、6時と貼り合せながら眺めることができるならば、予測にとって大きな情報を与えることができるであろう。

さて話をもとにもどして予測法を具体的に述べよう。いま将来の T 時刻の i 地域の値を予測しようとする。 T 時刻の X_i の予測値を $\hat{X}_i(T)$ と書き、次の公式によって求める。

(i) $C=(1-r)/S>0$ の場合(直接回帰, ウェイト回帰共通)であり、 $\beta=-\log r$, 対数は自然対数)いま現時刻を T_0 として、リードタイムを $T-T_0=l$ とする。

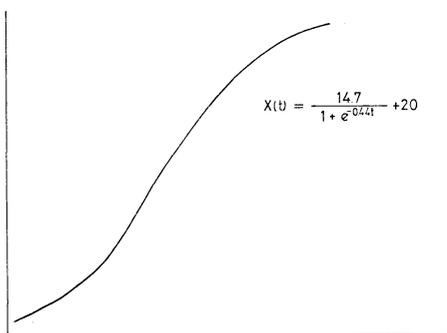
$X_i(T_0) \leq C$ の時、

$$\hat{X}_i(T_0+l) = \hat{X}_i(T) = \frac{C}{1+e^{-\beta(t_0+l)}}$$

ここに t_0 は

$$X_i(T_0) = \frac{C}{1+e^{-\beta t_0}}$$

をみたす t_0 である。



第5図

$X_i(T_0) > C$ のとき,

$$\hat{X}_i(T) = \hat{X}_i(T_0 + l) = C$$

とする.

(ii) $C \leq 0$ の場合,

直接回帰法では

$$\hat{X}_i(T) = \left[\frac{\sum_{i=1}^N \sum_{t=1}^n \left(\frac{1}{\hat{X}_i(t)} \right)^2}{\sum_{i=1}^N \sum_{t=1}^n \left(\frac{1}{\hat{X}_i(t)} \cdot \frac{1}{\hat{X}_i(t+1)} \right)} \right]^i \cdot X_i(t_0)$$

ウェイト法では

$$\hat{X}_i(T) = \left[\frac{\sum_{i=1}^N \sum_{t=1}^n w(X_i(t)) \frac{1}{(\hat{X}_i(t))^2}}{\sum_{i=1}^N \sum_{t=1}^n w(X_i(t)) \frac{1}{\hat{X}_i(t)} \frac{1}{\hat{X}_i(t+1)}} \right]^l \cdot X_i(t_0)$$

この方法で第1図で示した喫煙者のたばこ平均全銘柄の購買価格を計算した(値上げ以前). これをかかげておく(第5図). この場合の時間軸についてはもはや時間の概念をもっていないで単に「ずらし尺」の意味しかない.

最後に貼り合せ法については, 国際比較による予測法などにも適用して成功しているが, 気象関係についても参考になればと思って筆をとった次第である.

電子計算機と天気予報*

新 田 尚*

1. はしがき

シンポジウムが終了後, シャべった内容をまとめるように依頼されて, もう一度今回のシンポジウム全体をふり返って考えてみた. 大規模気象現象の予測について, 学会の立場からいろいろ話題を提供し, それを材料にして各方面から議論を盛上げてゆくというのが本来の趣旨であったかと思う. 事前の話題提供者の打合せでも, その点で意志統一がなされていたと考えていたつもりだったが, 実際は

- i) 全体的にみて science の立場に重点をおくのか, engineering の立場に重点をおくのか 明確でなかった.
- ii) 特に筆者分担の「電子計算機と天気予報」というテーマの趣旨を, 筆者自身十分消化していなかった.
- iii) 全般に討論が中途半端であった.

という印象を筆者は抱いた.

そこで, 本稿では実際に当日会場でシャべった内容に若干加筆修正して, 筆者なりに上記の欠点を補っておきたいと思う. なお, 筆者は今回の予測の問題を engineering の立場に重点をおいて, 10年ぐらい先の将来をみとおしながら, なるべく学問的に扱ったつもりである.

2. 電子計算機の性能と将来のみとおし

今後10年ぐらいの間に, 電子計算機の性能がどの程度進歩するかを一応確認しておきたい. 勿論, これは予想

第1表 代表的な電子計算機の処理能力比較表
(MIPS: million instructions per second)

電子計算機	処理能力
IBM-360/65	1 MIPS
360/91	5
360/195	8~9
360/95-371/165	15
CDC-6500	1.5
6600	2.5
7600	8
HITAC-8800	3~4
VESNA	1
TI-ASC	60

(註) 他の国産機種については, HITAC との比較で考えて頂きたい. また, TI-ASC は1974年3月現在まだ本格的運用に入っていない.

であるからはずれられるかもしれないが, 大規模気象現象の予測の研究を行うためには電子計算機は欠かせぬ道具であるので, まずそれに関するみとおしをたてておきたいと思う. 第1表に現在世界中で用いられている代表的な電子計算機の処理能力を示した. これをみると, 現用の計算機は大体 5 MIPS 前後といえよう. 最近の hardware と software の進歩の度合からみて, 1980年に現実的に使用可能な電子計算機の処理能力は, おおよそ10~25MIPS とみこまれる. この性能の向上によって, 1

* Electronic computer and weather forecast

** T. Nitta 気象庁予報部電子計算室