

東京大学工学部 学生員 水野 明久

東京大学工学部 学生員 若谷 佳史

1. はじめに

動的モデルを構築した時 その有効性は予測の適合度によって評価される。一般にモデルは事後シミュレーションによりテストされるが、その際、得られた予測値と実績値の時系列から 全体としての予測誤差の大きさをどのように測定するか が問題となる。本研究では 種々の測定尺度の内 Theil の不一致係数をとりあげ、動的モデルの精度分析を行なった。そして不一致係数の有効性を検討した。

2. 予測誤差の測定尺度

予測誤差の測定に関する主な尺度は次のようなものがある。(予測値 P_t 実績値 A_t)

(1) 平均絶対値 予測誤差	$\frac{1}{N} \sum_{t=1}^N P_t - A_t $
(2) 平均絶対百分率 誤差	$\frac{1}{N} \sum_{t=1}^N \left \frac{P_t - A_t}{A_t} \right $
(3) 平均二乗誤差 の平方根	$\sqrt{\frac{1}{N} \sum_{t=1}^N (P_t - A_t)^2}$
(4) 相関係数	$R = \frac{\sum (P_t - \bar{P})(A_t - \bar{A})}{\sqrt{\sum (P_t - \bar{P})^2 \sum (A_t - \bar{A})^2}}$
(5) Theil の 不一致係数	$U = \frac{\sqrt{\frac{1}{N} \sum (P_t - A_t)^2}}{\sqrt{\frac{1}{N} \sum P_t^2} + \sqrt{\frac{1}{N} \sum A_t^2}} \quad (0 \leq U \leq 1)$

(1)~(4)については 詳しい説明は必要ないと思われるが、(4)の相関係数は予測精度を分析するには不備な点がある。相関係数は、予測値 P_t と実績値 A_t の間に

$P_t = \alpha + \beta A_t$ が成立することを意味するに過ぎないからである。予測が完全であるためには $\alpha = 0$, $\beta = 1$ である必要がある。この点を考えて Theil は (5)の不一致係数を定義した。次にこの不一致係数 U の性質を説明する。 U の分子の平方根の中は

$$\frac{1}{N} \sum (P_t - A_t)^2 = (\bar{P} - \bar{A})^2 + (S_P - S_A)^2 + 2(1 - r) S_P S_A$$

($\bar{P}$, $\bar{A}$ は予測値、実績値の平均値, S_P , S_A は) ①
各々の標準偏差, r は P , A の相関係数

右辺の第一・二・三項を U の分母の 2 乗でわったものを U_M^2 , U_A^2 , U_C^2 とかくと

$$U^2 = U_M^2 + U_A^2 + U_C^2 \quad \cdots \cdots (2)$$

よって 不一致係数 U が 予測値平均と実績値平均のずれ (U_M), 両者の散らばり方の差 (U_A) 相関係数が 1 より小さいことによる誤差 (U_C) に分けられたことになる。又 (2) 式を U^2 で除すと三者の比率 U_M^2/U^2 , U_A^2/U^2 , U_C^2/U^2 が求まり 全体の誤差中の構成割合が表わされる。

$$U_M^2/U^2 + U_A^2/U^2 + U_C^2/U^2 = 1 \quad \cdots \cdots (3)$$

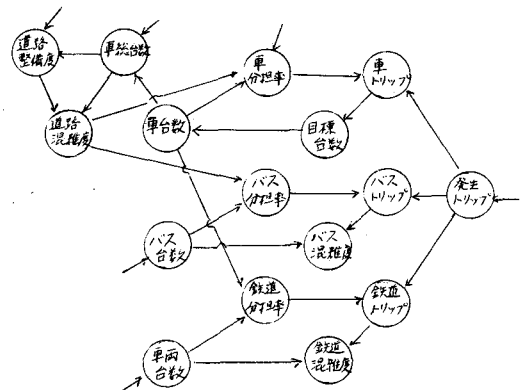
このような指標を算出することにより 予測精度、誤差の原因が分析できる。Theil は U_M^2/U^2 , U_A^2/U^2 については規則的な誤差とし、改善の余地があると考え $U_C^2/U^2 = 1$ に近い程有効な予測としている。

3. 動的モデルの概略

予測を目的とする動的モデルの例として交通モデルを構築し、交通分担セクターにおける構造パラメータを変化させ 予測精度の変動を調べた。モデルは三つのセクター (1) 都市活動セクター (2) 交通分担セクター (3) 大気汚染セクター から構成されている。交通分担セクターのフロー図は 図-1 のようである。

変化した構造パラメータは 自動車台数増加メカニズムに関わるものである。

図-1 交通分担セクターフロー図



$$\left\{ \begin{aligned} \text{采期目標台数} &= \frac{\text{今期車トリップ}}{\text{1台当りの輸送人員}} \times \alpha \\ \text{車増加数} &= (\text{采期目標台数} - \text{今期車台数}) \times \beta \end{aligned} \right.$$

α は増加係数、 β は抑制係数である。この α ・ β 両者の値の変化による予測精度への影響を検討し、このモデルで最適な α ・ β を決定した。

4. 不一致係数を用いた精度分析

(1) UとRの比較

図-2からわかるようにUの値はケースⅢ($\alpha=1.35$, $\beta=0.47$)、ケースⅡ($\alpha=1.345$, $\beta=0.46$)、ケースⅠ($\alpha=1.345$, $\beta=0.45$)の順に小さく、精度が良くなっている。一方、相関係数Rは、Uに比べケース毎の差が小さい。(図-3)つまりUの方がRに比べて構造パラメータの変化による感度が大きく、予測精度を把握しやすいことがわかる。

予測と考えた時Rは必ずしもよい評価指標ではないことは前に示した。バストリップ(BPT)を例にあげる。R<0.4でありRから判断すれば予測不可能と判断されるがU値をみればBPTは他変数と比べ中位程度の精度を持っている。これを実績値-予測値のグラフ(図-4)で検討するとバラツキは大きい。Uの変動量が小さいためU値が小さくなっていることがわかる。この場合相関は悪くても予測値として許容されることが考えられる。

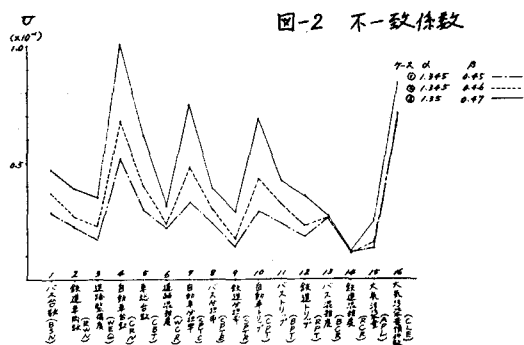


図-2 不一致係数

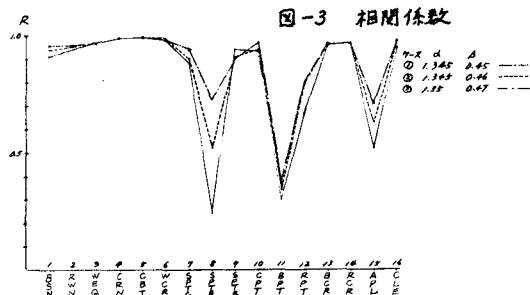


図-3 相関係数

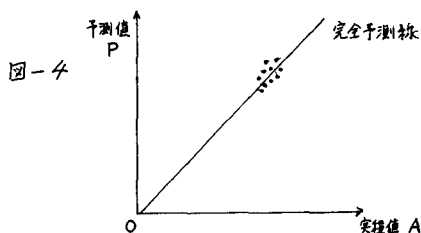


図-4

(2) 誤差の構成割合の検討

U^M U^S U^Cを三角座標に表わしたものが図-5である。ケースⅠとケースⅢと比較する。予測精度の良いケースⅠの方が各変数ともU^S=1の方向にかたっており有効な予測となっている。ケースⅢにおいては三角座標内のバラツキが大きく規則的な誤差が含まれていると考えられる。

ケースⅠでCRN(4) CBT(5) WCR(6)のU^Sが大きいのは、実績値と予測値の分散の大きさが異なっている結果であろう。又RPT(12) RWN(2)は他変数に比べU^Mの割合が大きく平均値としての予測にあたっていないことを示している。

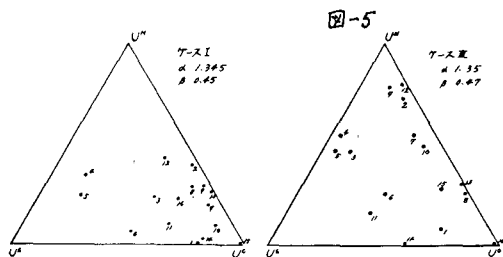


図-5

5 おわりに

予測を目的とした動的モデルを構築し、Theilの不一致係数を用いて予測適合度の評価を行なった。又Uは予測誤差の原因分析に役立ち、構造改善の材料となることを述べた。今回は構造パラメータの変化によるモデルの予測適合度の検討であったがさらに拡張して全体システムの中のスブセクターと多面的にとらえいくつかモデル化し、それらと比較検討する場合にも有用であろう。Theilの不一致係数は、予測誤差をグラフ表現によって視覚的に認識するに加えて、誤差を定量的におさえる指標として、又構造改善に対する情報を与えるものとして一層の利用が望まれる。

参考文献 「経済の予測と政策」 H.Theil 岡本訳
「日本の経済予測」 森一夫