

時系列モデルによる高速道路の交通量と所要時間の変動の分析

山梨大学 工学部土木環境工学科 学生会員 ○鈴木 康平
山梨大学 大学院医学工学総合研究部 正会員 佐々木 邦明

1. はじめに

1) 研究背景

近年, ETC の普及により高速道路利用者の継続的なデータが得られている. これを用いて, 交通量と所要時間の時間的な変動を見ていく. 交通量と所要時間の変動幅の変化が, 高速道路の速達性を安定的に果たす能力である, 時間信頼性に影響を与えている.

2) 研究目的

時間と共に変動する現象に対して, 時間の順序で測定・観測したデータを時系列データという. 時系列データの変動を統計的に分析し, 特徴を捉え, 現象の解明と将来の変動を予測・制御しようとするのが時系列分析の主な目的である.

時系列データには, 平均や分散などの統計的性質が時点に依らず一定である定常時系列と時間の経過と共にこれらの性質が変化する非定常時系列がある.

時系列分析を用いて交通量と所要時間のデータをモデル化し, 各種料金政策や道路の建設・改良が行われた際の変化を見ていく. モデルの比較を通じて, どの様な手法が道路事業を評価する上で望ましいか, 時間信頼性がどう変化していくのかを明らかにしていきたい.

2. 分析対象データ

分析には, 以下の二つのデータを用いた.

1) 交通量

2006/4/1~2010/2/28 の中央自動車道八王子 IC→河口湖 IC 間 (下り) の日交通量

2) 平均所要時間

2006 年 5 月の一ヶ月間の河口湖 IC→八王子 TB 間 (上り) の 15 分ごとの平均所要時間

キーワード: 時系列分析, 高速道路, 信頼性

山梨大学土木環境工学科交通工学講座

(〒400-8511 甲府市武田 4-3-11 Tell/Fax 055-220-8671)

E-mail: t11c038@yamanashi.ac.jp

3. 時系列モデル

定常時系列データの分析には AR モデル, MA モデル, ARMA モデルが用いられ, 非定常時系列データの分析には ARIMA モデル, SARIMA モデルなどが用いられる.

1) AR (自己回帰) モデル

確率変数列 $\{R_1, R_2, R_3, \dots, R_n\}$ の要素 R_t と R_{t-1} の間に何らかの関係があり, この関係性を定式化すると以下のように表される.

$$R_t = \mu + \varphi_1 R_{t-1} + \varepsilon_t \cdots (1)$$

(μ : 定数項, φ_1 : パラメータ, ε_t : イノベーション)

1 時点前の自分自身を説明変数としており, 1 次の自己回帰モデルと呼び, AR(p) と表記する. ここで p は次数である. ε_t はホワイトノイズ (平均が 0, 分散がある一定値, 自己共分散が 0) を仮定しており自己相関性はないため, 過去の情報とは無関係に新たな情報を確率的に与える部分である.

R_t を平均所要時間とすると, R_t は 1 時点前の平均所要時間で表され, 時間信頼性は時間変動する誤差項の分散で表せるため評価が可能.

2) MA (移動平均) モデル

$$X_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \cdots (2)$$

(q: 次数, θ : パラメータ, ε : 誤差項)

当日の値は, 過去の誤差の影響を受けることを表現している. MA(q) と表記する. q は次数である.

3) ARMA (自己回帰移動平均) モデル

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \cdots (3)$$

(p, q: 次数, ε_t : 誤差項, φ, θ : パラメータ)

時点 t の値 X_t を過去の X の値と過去の残差で説明するモデルであり, AR モデルと MA モデルを組み合わせたモデルで, 比較的小さな次数でも複雑な現象を表現できる. また, ARMA(p,q) と表記する.

4) ARIMA (自己回帰和分移動平均) モデル

時点 t における実測値を X_t とすると, (4) 式のように

に表される.

$$\phi(L)\nabla^d X_t = \theta(L)\varepsilon_t \cdots (4)$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p \cdots (5)$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q \cdots (6)$$

ϕ_i , θ_i は係数, L はラグ演算子($L^n X_t = X_{t-n}$), ∇ は連続階差を示す差分演算子($\nabla^n X_t = X_t - X_{t-n}$), ε_t は残差項をそれぞれ表す. また, (4)式は演算子を定義することにより, L , ∇^n を一般の変数とみなして計算することができ, (7)式のように要約される.

$$ARIMA(p, d, q) \cdots (7)$$

(p, q : 次数 d : 階差)

時系列データの階差を取ることで, データを定常化し ARMA モデルを適用するモデルである.

5) SARIMA(季節 ARIMA) モデル

ARIMA モデル同様に, 時点 t における実測値を X_t とすると, (3)式のように表される.

$$\phi(L)\phi(L^s)\nabla_s^D \nabla^d X_t = \theta(L)\theta(L^s)\varepsilon_t \cdots (8)$$

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p \cdots (9)$$

$$\phi(L^s) = 1 - \phi_1 L^s - \phi_2 L^{2s} - \cdots - \phi_p L^{sp} \cdots (10)$$

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \cdots - \theta_q L^q \cdots (11)$$

$$\theta(L^s) = 1 - \theta_1 L^s - \theta_2 L^{2s} - \cdots - \theta_q L^{sq} \cdots (12)$$

ϕ_i , θ_i は係数, ∇_s は季節階差を示す差分演算子($\nabla_s^n X_t = X_t - X_{t-n}$), を表す. また, ARIMA モデル同様に(8)式は演算子を定義することによって L , ∇^n , ∇_s を一般の変数とみなして計算でき, (13)式のように要約される.

$$SARIMA(p, d, q)(P, D, Q)_s \cdots (13)$$

(p, q, P, Q : 次数 d, D : 階差 s : 周期)

時系列データの階差を取り, ARMA モデルを適用する ARIMA(p, d, q)と, 季節階差(周期的変動)に関する ARIMA(P, D, Q)を組み合わせたモデルである.

6) GARCH モデル

時間変動する誤差分散を表現するモデルである.

$$X_t = \mu + \varepsilon_t \cdots (14)$$

$$\varepsilon_t = \sigma_t v_t \cdots (15)$$

$$v_t \sim N(0, 1) \cdots (16)$$

$$E(\varepsilon_t^2) = \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \cdots (17)$$

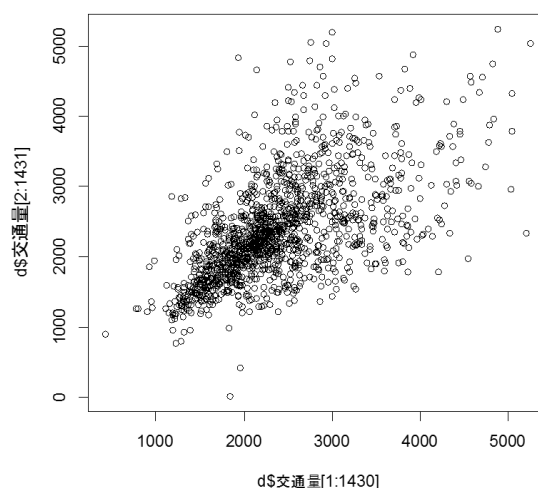
(p, q : 次数, v_t : 標準化残差, μ, ω : 定数項)

(17)式をボラティリティと言い, 残差の分散を時

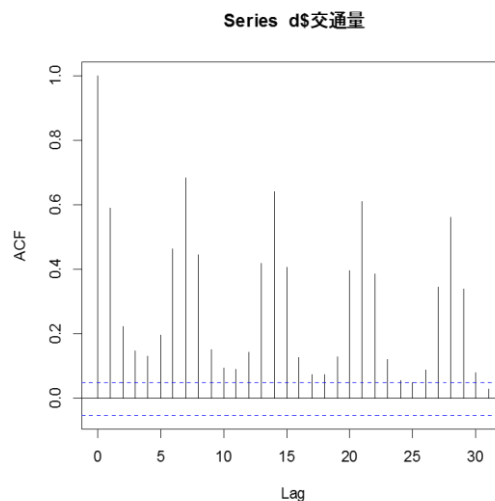
系列で表現するため, ばらつきを評価できる.

4. 交通量の分析

自己相関の散布図から正の相関が見られ, 統計的検定からも自己相関関係を有すると言える. 自己相関係数をグラフで見えていくと, 時間の経過と共に減少する傾向があり, 周期的な変動をしていることが分かる.



図—1



図—2

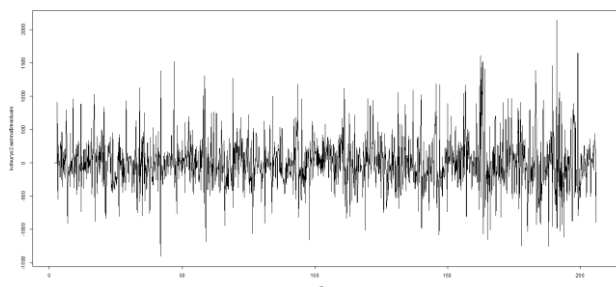
交通量のデータは平均や分散が時間と共に変化する非定常時系列であり, 一週間, 一年といった周期が見られることから SARIMA モデルで表現していく. 周期 $s=7$ (曜日), d, D は階差を取る($=1$)か取らない($=0$)かなので $d=1, D=1$ となる. 次数の組み合わせは $p \times q \times P \times Q$ 通りと無数にあるが, 限られたデ

一タからより多くのパラメータを推定すると誤差が大きくなりやすく、パラメータ数は1〜3ぐらいが望ましい（Box-Jenkinsのケチの原理より）とされている⁴⁾ため $3 \times 3 \times 3 \times 3 = 81$ 通りの中から選択する。今回はパラメータの t 値の有意性や、AIC から SARIMA(1,1,2)(3,1,3)7 を選択。各推定値は次の表の通りである。

表—1

SARIMA(1,1,2)(3,1,3)7	推定値	標準誤差		
ϕ_1	0.2004	0.0749		
θ_1	-0.6361	0.0755		
θ_2	-0.2454	0.0598		
Φ_1	-0.3299	0.0332		
Φ_2	-0.9179	0.0252		
Φ_3	0.0604	0.0303		
θ_1	-0.5641	0.0218	残差の分散	155996
θ_2	0.6313	0.0191	最大対数尤度	-10533.42
θ_3	-0.9359	0.0269	AIC	21086.84

最大誤差は2141.822(台)、一日当たりの平均誤差は393.8573(台)である。



図—3

決定係数は0.710922と、状態空間モデルに比べて精度は劣るものの、データの7割を表現できている。次の表は左が時系列モデル、右が先行研究の状態空間モデルである。先行研究¹⁾で同データを状態空間モデルで表現しており、推定されたモデルの比較を行った。

表—2

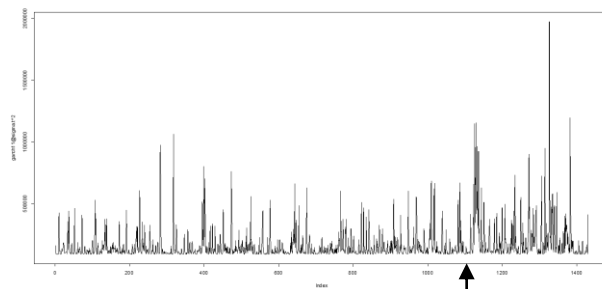
最大誤差	2141.822	最大誤差	-40.7169
一日当たりの平均誤差	393.8573	一日当たりの平均誤差	4.6152
誤差の標準偏差	393.9821	誤差の標準偏差	6.6971
自由度調整済み決定係数 r^2	0.710922	r^2	0.99994

GARCHモデルの次数はAIC等の情報量基準によって選択されるが、多くの実証研究において次数を大きくしてもあまり説明力が改善されないことが示されている⁵⁾。従って標準的なGARCH(1,1)モデルについて分析していく。得られた推定値は以下の通り。

表—3

	推定値	標準誤差		
μ	-1.339	9.051		
ω	77440	7945		
α_1	0.4035	0.0561		
β_1	0.1407	0.0582	最大対数尤度	-10494.12

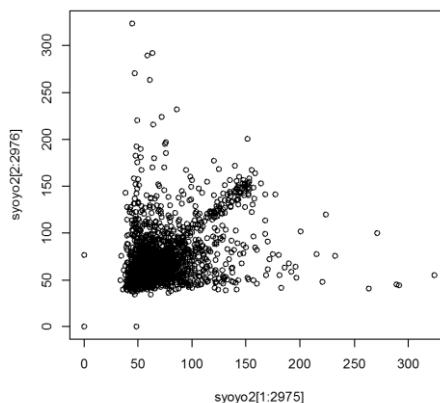
ボラティリティは次のグラフの通り。2009/3/28(表の矢印)に休日特別割引が導入されて以降、ボラティリティの変動が大きくなっていることが分かる。



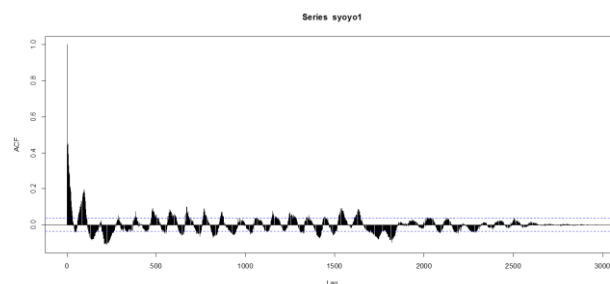
図—4

5. 平均所要時間の分析

交通量と同様に自己相関の散布図から正の相関が見られ、統計的検定からも自己相関関係を有することが分かる。自己相関係数も時間経過と共に減少し、周期的な変動が見られた。



図—5



図—6

6. まとめ

交通量に関して、休日特別割引以降にボラティリティの変動が大きくなることが確認できた。このことから、割引導入後の高速道路利用者の増加によって交通量のばらつきが大きくなっていることが分かった。また、所要時間も交通量と同様に周期性を持ち、時間経過と共に自己相関が減衰していく傾向が見られた。

今後の課題としては、所要時間のモデルの作成と比較、分析を進め、時間信頼性の変化について見ていく必要がある。また、データの期間を変えて道路の建設・改良の前後での時間信頼性の変化を見て、道路事業の評価に繋げていく必要がある。

参考文献

- 1) 中沢航太：IC 間利用履歴データを用いた高速道路の短期的需要予測手法に関する研究，山梨大学大学院修士論文，2012。
- 2) 名取優太：状態空間モデルによる高速道路の需要変動特性の分析，山梨大学卒業論文，2013。
- 3) 青木義充，横内大介：時系列データ分析，技術評論者，2014。
- 4) 田中孝文：R による時系列分析入門，シーエーピー出版，2013。
- 5) 三井秀俊：東日本大震災による日本の株式市場の構造変化の検証，産業経営研究 第 34 号 (2012)，pp.31-41。
- 6) 川崎智也，松田琢磨，花岡伸也：東アジア積米国揚コンテナ荷動き予測における SARIMA モデルの適用性，日本物流学会誌第 21 号 平成 25 年 5 月，pp.167-174。