

ベクトル自己回帰モデルによるコミュニティサイクル貸出回数の時系列分析

| | | |
|-------------------------|------|--------|
| 秋田工業高等専門学校専攻科環境システム工学専攻 | 学生会員 | ○千田 真吾 |
| 秋田工業高等専門学校環境都市工学科 | 正会員 | 長谷川裕修 |
| 株式会社ドーコン交通事業本部交通部 | 正会員 | 松田 真宜 |
| 室蘭工業大学大学院工学研究科くらし環境系領域 | 正会員 | 有村 幹治 |

1. はじめに

近年、都市部を中心として、自転車の共同利用事業（コミュニティサイクル、以下 CC）が展開されている。CC は配置されている専用の自転車貸出・返却場所となる駐輪ポート間であれば、どこでも自由に貸出・返却が可能な、自転車の共同利用事業である。CC の導入により、都心モビリティの向上、放置自転車の抑制、自動車交通からの転換による CO₂ 排出の削減などの効果が期待されている。

我が国における CC を巡る環境は、海外と比較すると、まだ整っていないとはいえない。導入の可能性や安定した運用方法について我が国独自の研究が必要とされている。CC の安定した運用、効率的運用のためにはより正確にその利用動態を把握することが必要となる。現在、いくつかの CC 事業においては IC カードや携帯端末による個人認証と決済がなされ、利用者の利用記録が記録されている。CC の利用記録を用いた研究としては、佐藤らのコミュニティサイクルシステムの利用実態とステーション配置に関する研究¹⁾、丸山らのデータマイニングアプローチによるコミュニティサイクルの利用動態の抽出²⁾などがある。これら既往の研究の多くは CC の利用実態を把握することを目的としているが、時系列的にその把握を試みた研究は少ない。

以上を踏まえて本研究は、札幌市の CC「ポロクル」利用状況の時系列的な把握を目的として、「ポロクル」の利用記録データおよび札幌市の気象データを用いて、ベクトル自己回帰モデル（以下 VAR モデル）を構築する。

2. 使用データの概要と前処理

ポロクルは IC カードまたは携帯端末により、個人認証と決済がなされ、各利用者の利用動向が記録さ

れる。本研究では個人を特定できる属性を消去し、貸出・返却の時間及びポートを含むデータセットを用いた。ポロクルの営業期間は冬期間を含まない 4 月から 11 月まで、営業時間は 7:30~21:00 までとなっている。本研究では 2011・2012 年度の「ポロクル」利用履歴データ全 162,570 レコードから、1 時間ごとの全ポート合計での『貸出回数』データを集計した。一日当たりの『貸出回数』は 7 時台から 20 時台までの 14 個であり、2 年度分で計 6202 個のデータを分析に用いた。

気象が CC 貸出回数に与える影響を考慮するために気象庁の気象統計情報より、札幌市の気象データとして 1 時間ごとの気温・降水量のデータを入手した。『貸出回数』と気温・降水量の分布を散布図で確認すると、気温が 5 度以下になるときと、降雨があるときに『貸出回数』が低下していることがわかった。このことから『貸出回数』に負の影響を与える場合を 1、それ以外を 0 としたダミー変数を作成し、それぞれを『気温ダミー』・『降水ダミー』とした。

VAR モデルによる時系列分析に先立ち、『貸出回数』の周期性を自己相関係数によって確認した。その結果、日内変動、隔日での変動、週間変動による周期性があることが分かった。また、『貸出回数』の時系列的な分布図から、全体的に増加傾向にあり、2011 年度に比べて 2012 年度は大きく増加していることがわかった。『貸出回数』の経年的な増加を考慮するために『営業年数』を 2011 年度を 1、2012 年度を 2 とし、付値した。

3. VAR モデルによる CC 貸出回数の時系列分析

(1) VAR モデル構築

CC の利用状況を時系列的に把握するために VAR モデルを構築する。VAR モデルとは過去の時系列デ

ータから未来の値を予測するのに用いるモデルの一つで、自己回帰モデルを多変量に拡張したものである³⁾。複数の時系列データを内生変数ベクトルとして用いることで、予測の精度を上げ、変数間の動的な関係を見るのに適したモデルである。(1)式はn変量の内生変数ベクトル y_t についてVARモデルを構築した式である。ある時間tの内生変数ベクトル y_t が定数ベクトルcと自身のp期の過去の値に回帰するモデルとなる。

$$y_t = c + \sum_1^p a_k y_{t-k} + \varepsilon_t, \varepsilon_t \sim W.N.(\Sigma) \quad (1)$$

ここで、 y_t :t時の予測値ベクトル, c:定数ベクトル, a_k :係数行列, y_{t-k} :t-1期からt-p期までの実測値ベクトル, p:期数, ε_t :ホワイトノイズベクトル

本研究ではCCの利用状況のデータとして『貸出回数』、予測精度を上げるため『気温ダミー』、『降水ダミー』、『営業年数』を利用した。これらの4変量が内生変数ベクトルとなる。今回のモデルはトレンドという増加・減少などの変動傾向と、定数項で表現されるモデル全体のずれのどちらも考慮するモデルを選択した。何期前の値に回帰したモデルとするのかを決定する期数pは10営業日である140期を上限として、AICが最小となる140を用いた。

(2)モデルの評価

構築した4変量のVARモデルに対して、Granger因果性の検定を行った。この検定はある変量が時系列的に予測に有用であるかどうかを検定するものである。検定結果を表1に示す。P値が5%以下のとき、その変数はモデルの予測に有用であることを表す。また、F値が大きく、P値が小さいほど、予測に有用な変数である。表1から4変量は全てモデルの予測に有用であり、その中でも『貸出回数』が最も有用であり、以下『降水ダミー』、『営業年数』、『気温ダミー』の順に有用であることがわかる。

時系列分析によって利用状況の把握を試みることに有効な手段であることを示す一つの判断材料として、同じ変数データを用いた回帰モデルとVARモデルのAICを比較した。比較に用いる回帰モデルは目的変数を『貸出回数』、説明変数を『気温ダミー』、『降水ダミー』、『営業年数』として線形回帰モ

デルを構築した。表2に回帰モデルとVARモデルのAICを示す。表2からVARモデルの方がAICが小さく、今回使用したデータの予測に相当であると判断できる。

表1 Granger因果性の検定結果

| | F値 | P値 |
|-------|------|---------|
| 貸出回数 | 11.6 | 2.2E-16 |
| 気温ダミー | 1.5 | 2.8E-09 |
| 降水ダミー | 4.1 | 2.2E-16 |
| 営業年数 | 2.2 | 2.2E-16 |

表2 AICの比較

| | AIC |
|--------|-------|
| 回帰モデル | 55519 |
| VARモデル | -6882 |

4. おわりに

本研究では札幌都心部で展開されているCC事業「ポロクル」の利用記録及び気象庁の過去の気象データを用いてベクトル自己回帰モデルを構築し、利用状況の時系列的な変動の把握を試みた。結果、CCの貸出回数は日内変動、隔日での変動、週間変動による周期性を持ち、気温の低さ、降雨の有無、営業年数が影響していることが明らかとなった。

この結果は各ポートでの利用状況の分析やより発展したモデルの構築などに役立てることができると考えている。最終的に各ポートの利用状況の把握や短期的な予測などに展開することで、より効率的なCC事業の運用に役立てることができると考えている。

参考文献

- 1) 佐藤 仁美, 酒井 良輔, 三輪 富生, 森川 高行: コミュニティサイクルシステムの利用実態とステーション配置に関する研究, 土木学会論文集 Vol.69, No.5, pp.563-570, 2013
- 2) 丸山 翔大, 松田 真宜, 長谷川 裕修, 有村 幹治: データマイニングアプローチによるコミュニティサイクルの利用動態の抽出, 土木学会論文集 Vol.70, No.5, pp.671-680, 2014
- 3) 沖本竜義: 経済・ファイナンスデータの計量時系列分析, 朝倉書店, 2010