

# 進化するARIMAモデル

## —時系列解析による予測値推計の高精度化・自動化—

ITの進展により時系列データが多く蓄積されている。時系列データを予測に用いることにより、在庫などさまざまな計画の精度（妥当性）を高めることが可能だが、予測モデルの構築や更新にコストがかかることが課題になる。本稿では、野村総合研究所（NRI）などが開発・適用している自動予測手法であるARIMAモデルを紹介する。

### 蓄積される時系列データ

時系列データの例としてなじみが深いのは、商品を購入した日時と数量を記録したPOS（販売時点情報管理）データであろう。最近では、デバイスやセンサーで取得したデータや、携帯情報端末などから取得したデータの蓄積も進んでいる。

時系列データに基づいて将来値を高精度に予測できれば、例えば流通業においては販売量の予測精度が上がり、仕入れ量の調整によって在庫の適正化が可能となる。時系列データによる予測は、このほかにもさまざまな用途が考えられる。

### 代表的な2種類の予測手法

予測の手法には、変化の要因を推定し、その要因についてのデータに基づいて予測する手法と、過去の実績値と将来値との関連性を推定し、過去の実績値に基づいて予測する手法の2つがある。

前者の手法には重回帰分析があり、広く用いられているために実用的な精度を確保しやすいという利点がある。しかし、将来値予測に用いる説明変数（統計情報、気象情報な

ど）と実績値の相関を分析し、実績値を説明する変数と係数を推定する際に人間の判断を必要とするために、予測対象の数に比例してモデル構築の人的負荷が増大する。このほか、説明変数に用いるデータの取得・メンテナンスにコストを要し、環境変化により因果関係が変わった場合にモデルの再調整が必要になるという欠点もある。

後者の手法にはARIMA（AutoRegressive Integrated Moving Average：自己回帰和分移動平均）モデルがある。ARIMAモデルで予測に用いる変数は予測対象の実績値のみであり、重回帰分析と比較するとデータの取得・メンテナンスのコストが低い。また、重回帰分析が将来値の平均値を予測するような手法であるために変動が平均化されてしまうのに対し、ARIMAモデルは変動のパターンを予測する手法であり、日々変動する値を予測する用途に適している。

しかしながら、ARIMAモデルはあまり活用が進んでいない。その理由には、自己回帰の次数（時系列の列数）や移動平均の係数の推計に時間を要すること、多変量解析と同様にモデルの構築に人的負荷が大きいことが挙げられる。このほか、実績値に極端な値（特

野村総合研究所  
コンサルティング事業本部  
IT事業推進部  
上級コンサルタント  
今井 恒 (いまいこう)  
専門はシステム開発、地理情報  
システム、データ分析



図1 販売数量予測の例



異値)が含まれていた場合に予測精度が低下すること、新商品の販売量予測など、実績データがない場合や実績データとの因果関係が全くない事象については適用できないことも欠点として挙げられる。

### 改良されるARIMAモデル

NRIは、ARIMAモデルを用いた時系列解析による予測エンジンの開発・改良を行っている。最近では、推計アルゴリズムの進化により、予測計算ごとに自動的に次数・係数を十分に実用的な処理速度で推計することが可能となった。予測計算ごとにモデルを新たに自動構築することが可能となったため、環境の変化によりモデルの調整が必要になった場合も容易に対応できる。

この予測エンジンでは予測計算ごとにモデルの再構築を行うため、多数の対象について1秒ごとに予測を行うような用途では、コス

トも含めてまだ実用的ではない。しかし、流通の受発注数量の予測のように1日ごとに予測計算を行う場合や、数分ごとに予測計算を行う程度であれば十分に実用的である。

### 予想される各分野での活用拡大

図1は、開発した予測エンジンを用いて過去70日間の販売数量に基づいて次の7日間の販売数量の予測を繰り返した例である。中央付近で実績と予測の乖離が見られるが、その前後はかなり一致しており、期間を通して予測値に基づく仕入れを行えば在庫の削減が可能と考えられる。

ARIMAモデルは処理時間が改善され精度も比較的高いため、今後、各分野での活用が期待される。NRIでは既存システムの予測エンジンをARIMAモデルを用いて高度化するなど、流通や金融への適用を進めており、さらに他分野へも拡大していく予定である。■