

## 第 10 回進化計算学会研究会 に参加して

鳥山直樹

Naoki TORIYAMA

電子情報学専攻修士課程 1年

### 1. はじめに

2016年3月17-18日に、第10回進化計算学会研究会において、ポスターセッションで研究成果「GPによる Realized Volatility の推定に関する評価」の発表を行った。

### 2. 内容

株式や為替などの価格変化の度合を表す指標にボラティリティがある。ボラティリティは投資リスクを表す指標であり、その変化の特性を解析し予測することは、リスク管理をするうえで重要となる。こうしたボラティリティの変化を表すモデルとして ARCH モデルや GARCH モデル等の様々な自己回帰モデルがあるが、使用するモデルによって同じデータに対して推定されるボラティリティが異なる。これに対し RV (Realized Volatility) は、日中の例えば5分ごとのリターンの2乗を足し合わせたもので計算される値であり、その値はモデルに依存しない。そのため RV を推定する様々な手法が考案され、GP を使って RV を推定する手法<sup>1)</sup>も提案されている。ここで GP とは Genetic Programming の略称であり、スペースの都合上詳細は割愛するが、設計者が考案したパラメータ (ノード) をうまく用いることで対象問題を解決することのできる学習モデルである。本研究では、RV の推定をするために9個パラメータを作成し、GP と同じ学習モデルである NN、経済の理論モデルとして使われている HAR モデルといった既存モデルとの性能比較を行った。また GP の大きな特徴として、学習したデータを解析することによってどのようなパラメータが使用頻度が高かったかなどの分析を行うことができ

Table 1 終端ノード

$RV(x)$ :  $x$  日前の RV (1-5 日)  
 $RV_{oc}$ : 前日の始値と終値の差の絶対値  
 $RV_{hl}$ : 前日の高値と安値の差の絶対値  
 $Ret^2$ : 前日のリターンの2乗  
 $RV^2$ : 前日の RV の2乗

る。よって RV の推定に有効となるパラメータの分析を行った。

本研究では、日経平均株価指数の2014年10月1日から2014年12月30日のデータを対象とし、1日の5分足データを用いて RV の推定を行った。

RV の推定に用いた GP では、非終端ノードを  $\{+, -, *, /, \sin, \cos, \exp, \log, \sqrt{\quad}, \sqrt[3]{\quad}\}$  とし、終端ノードを  $\{RV[x], RV_{oc}, RV_{hl}, Ret^2, RV^2\}$  (Table 1) とした。

### 3. 結果

#### 3.1 性能の比較

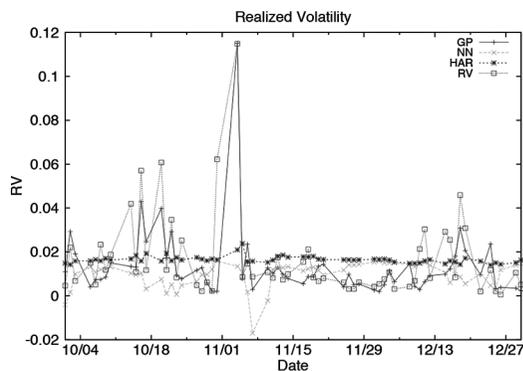


Fig. 1 各手法の RV の推定

Table 2 各手法との性能の比較

	GP	HAR	NN
Fitness	990.7	978.	973.5

Fig. 1 は各モデルで RV を推定した結果である。HAR モデルでは、全体的な変化の平均をとるような結果になり、NN では小さい変化は捉えることができたが、所々に出てくる大きな変化を捉えることが困難であった。GP では、全体の変化を捉えるよ

**Table 3** 終端ノードの出現頻度（最悪解）

	AverageFitness	$RV\ 1$	$RV\ 2$	$RV\ 3$	$RV\ 4$	$RV\ 5$	$Ret^2$	$RV^2$	$xoc$	$xhl$
Depth 10	777.5	10	12	2	6	12	10	7	8	2
Depth 15	612.8	8	4	8	20	7	6	8	12	6
Depth 20	683.8	25	16	23	18	24	22	10	26	14

**Table 4** 終端ノードの出現頻度（最良解）

	AverageFitness	$RV\ 1$	$RV\ 2$	$RV\ 3$	$RV\ 4$	$RV\ 5$	$Ret^2$	$RV^2$	$xoc$	$xhl$
Depth 10	988.4	28	22	11	14	8	9	8	12	13
Depth 15	990.4	31	48	23	15	22	25	22	28	16
Depth 20	990.7	43	60	31	26	31	54	38	44	39

うな精度の良いマッチングが出来ており、Table 2 が示すように、評価値も高い値となっている。

### 3.2 解析

次に、GP で生成されたモデルの解析を行う。GP では対象問題を解決するための重要な終端ノードがあると考え、今回単純な方法として終端ノードの出現数を数える。今回は、深さ制限（GP で解を作る際の広さ）を 10, 15, 20 の 3 パターン用意し、それぞれ 10 回試行した。各試行の最終世代において、Table 2 は最も評価値が低かった解、Table 3 は最も評価値が高かった解の終端ノードの個数をカウントしたものである。

Table 2 では、各パターンによって出現している終端ノードにばらつきがあった。深さ制限が 10, 20 のとき、 $RV\ 1$  の数が多いが、深さ制限 15 の場合は  $RV\ 4$  が最も多く、平均評価値も一番低い結果になった。一方 Table 3 では、 $RV\ 1$  や  $RV\ 2$  使用している解が多かった。これより今回使用した対象問題では、前日、2 日前の  $RV$  が性能の高い解を生成するための重要な終端ノードである可能性が高い。

## 4. まとめ

ボラティリティの特性を解析し、予測することは株式や為替などのリスク管理をするうえで重要なも

のである。本研究では GP を用いて、日経平均株価指数の  $RV$  の推定を行った。また HAR モデルや NN と性能を比較し、GP が最も良い性能を示した。次に GP の有効ノードの分析をした。今回の対象問題において、1 日前・2 日前の  $RV$  を使用することが、高性能の予測モデルを生成するために必要なノードである可能性が高いことを示した。

GP で複雑な問題を解決するためには様々な入力データが必要になるが、その問題に対してキーとなる入力データを発見し、終端ノードとしての出現頻度を調整することが、より素早く高性能な GP をつくるために必要だと考えられる。

## 5. おわりに

ポスターセッションによる活発な議論により、更に研究に対するモチベーションを高めることができた。今回の経験を今後の学会発表等の様々なことに活かしていきたいと思う。発表を行うにあたってご指導いただいた小野景子講師、ならびに研究室の皆様深く感謝致します。

### 参考文献

- 1) Zheng Yin, Anthony Brabazon, Conall O'Sullivan and Michael O'Neill, "Realised volatility forecasting: A genetic programming approach"