

web 検索履歴データを用いた外国為替市場のボラティリティ予測

家治川 博†

佐藤彰洋‡

京都大学大学院情報学研究科†

京都大学大学院情報学研究科‡

1. 序論

これまで ARCH モデルなど、外国為替市場のボラティリティを予測する手法が数多く提案されてきた[1]。一方で IT 技術の発達により様々な情報が入手可能になり、いかにその情報を利用するかが重要となっている。

Smith[2]は GARCH モデルを用いて過去の分散を推定し、これと Google trends の検索履歴の 1 週間と 4 週間平均を用いて回帰を行い、ボラティリティの予測方法を提案している。また、時系列が与えられたとき、ある時点で傾向が変わることが考えられるが Cheong[3]らは尤度比を用いて傾向の変化を捉える方法を提案している。

本稿では Smith の方法にインターバンクでの通貨ペアの取引回数を指標として組み込み、Cheong らの方法で時系列を分割する回帰モデルを用いる。

2. モデル

今回提案する時系列分割回帰モデルの概要を以下に示す。まず与えられたデータ $x_{i,t}, y_t$ ($i=1, \dots, p$; $t=1, \dots, N$) に対して次の(1)式のように回帰式を仮定し、その回帰誤差 z_t を計算する。

$$y_t = \beta_0 + \sum_{i=1}^p \beta_i x_{i,t} + z_t \quad (1)$$

与えられた時系列を 2 つに分ける点 τ についてそれぞれ誤差分布の尤度を計算する。時系列全体で計算した尤度と点 τ で分かれた左右の時系列の尤度の和との尤度比を計算し、最も尤度比が大きいときの点 τ を傾向の変化点の候補として採用する。ここでデータ数が M の尤度関数 L は以下のように計算できる。

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = \begin{bmatrix} M & \sum x_{1t} & \sum x_{2t} & \cdots & \sum x_{pt} \\ \sum x_{1t} & \sum x_{1t}^2 & \sum x_{1t}x_{2t} & \cdots & \sum x_{1t}x_{pt} \\ \sum x_{2t} & \sum x_{1t}x_{2t} & \sum x_{2t}^2 & \cdots & \sum x_{2t}x_{pt} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum x_{pt} & \sum x_{1t}x_{pt} & \sum x_{2t}x_{pt} & \cdots & \sum x_{pt}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_t \\ \sum x_{1t}y_t \\ \sum x_{2t}y_t \\ \vdots \\ \sum x_{pt}y_t \end{bmatrix} \quad (2)$$

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^M (y_t - \hat{\beta}_0 - \sum_{i=1}^p \hat{\beta}_i x_{i,t})^2}{M} \quad (3)$$

$$L(\hat{\beta}_0, \dots, \hat{\beta}_p, \hat{\sigma}^2) = \left(\frac{1}{2\pi\hat{\sigma}^2} \right)^{\frac{M}{2}} \exp\left(-\frac{M}{2}\right) \quad (4)$$

ただし、候補点において計算された尤度比は標本誤差を含んでおり、この有意性はブートストラップ分布から算出できる[4]。次の図 1 のように尤度比のブートストラップ分布を算出し、閾値と信頼区間の大小関係から有意性を確認する。これを再帰的に分割していく。

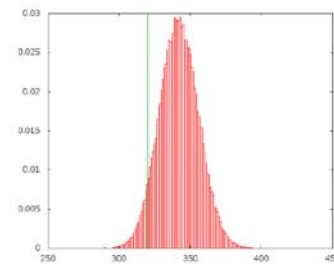


図 1: ブートストラップ分布と 5%信頼区間

3. シミュレーション

提案した時系列分割回帰モデルがどのような結果を示すのか、簡単な例で示す。

回帰式が $0 \leq t < 105$ で $y_t = a_0x_t + b_0 + z_t$, $105 \leq t < 200$ で $y_t = a_1x_t + b_1 + z_t$ で表される単回帰の場合で確認する。全体の線形回帰が図 2 であり、提案手法によるものが図 3 である。ここでは提案手法により回帰係数を再現することができている。

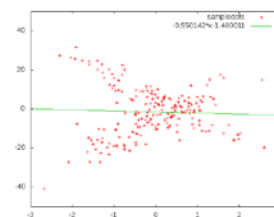


図 2: 全体での回帰分析結果

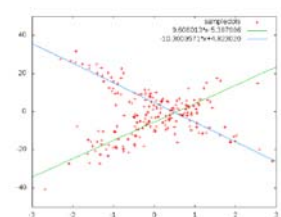


図 3: 分割後の回帰分析結果

4. データ

Volatility Prediction of Foreign Exchange Market Using Web Search Queries

†Hiroshi Kajikawa, Kyoto University

‡Aki-Hiro Sato, Kyoto University

本稿で扱う外国為替データは Federal Reserve System

(<http://www.federalreserve.gov/releases/H10/hist/>) から、検索クエリは Google trends (<http://www.google.com/trends/>) からそれぞれダウンロードした。回帰分析の説明変数として Google の検索クエリ (計 42 ワード) の対数値 G_w 及びその 4 週間平均、GARCH(1, 1) モデルによって計算された外国為替のボラティリティ h_t (対 USD の 7 通貨) と ICAP から抽出した銀行間取引回数の対数値 I_t を設定する。目的関数には実際に取引の行われた価格 r_t から算出した対数収益率 ε_t を用いる。ここで、対数収益率は $\varepsilon_t = \log(r_t) - \log(r_{t-1})$ で定義される。

データは 2007 年 5 月 28 日 (月) ~ 2010 年 12 月 31 日 (金) の期間に設定し、目的関数に用いる対数収益率は一日進んだものを設定した。また、Google の検索クエリ、インターバンク通貨取引は 1 週間単位のデータであるため、今回は抜けている値は直前の値を使った。

以下に今回のモデルで使用した回帰式を示す。

$$\log(\varepsilon_{t+1}^2) = \beta_0 + \beta_1 h_t^2 + \beta_2 \log(G_{w-1}) + \beta_3 \frac{1}{4} \sum_{i=1}^4 \log(G_{w-i}) + \beta_4 \log(I_t) + z'_t$$

ここで、 z'_t が回帰誤差である。

5. 結果・考察

提案手法を用いて実際に外国為替のボラティリティ予測を行う。AUD, CAD, CHF, EUR, GBP, JPY, NZD の 7 通貨と 3 つの区分に分けた特徴のある検索ワード (“関係のありそうなもの”、“関係のなさそうで検索回数が多いもの”、“関係のなさそうで検索回数が普通のもの”) の組み合わせで分析を行った。

	AUD	CAD	CHF	EUR	GBP	JPY	NZD
USA	5	6	10	5	5	6	6
foreign Exchange	2	11	17	5	4	3	5
Money	2	7	1	5	4	2	5
Oil	1	9	1	1	5	5	13
economics	3	9	10	6	7	5	8
bull market	1	3	6	7	5	2	3
but	6	6	10	5	4	2	4
and	1	10	3	5	5	2	7
apple	2	19	4	11	7	4	7
car	3	10	4	4	6	1	6
dog	1	5	6	2	4	5	6
water	3	5	9	5	4	3	10
guitar	1	4	10	5	4	3	9
sunny	5	7	6	11	5	4	5

表 1: 検索ワードごとの時系列分割回数

上の表 1 に示すように、利用する検索ワードによって時系列の分割回数が少し変わるという結果が得られた。同じ検索ワードでも通貨ごとに分割回数が違うことから、検索ワードと通貨には相性があり、時系列全体を通して説明能力がある場合やない場合、部分的に説明能力が高まる場合があると考えられる。

また、通貨によって分割の平均回数が異なる。分割回数の多い通貨は傾向が頻繁に変わる通貨と考えられる。

この手続きのもとで分割された時系列において、現在時刻に最も近いデータ列のみを使うことで、時系列のもつ直近の傾向を反映した回帰を用いた予測ができると期待される。

通貨ごとの予測に有用な検索ワードを見つける方法として、分割を実行した際に、現在時刻に最も近い分割のデータ列が短すぎず、かつその分割における回帰係数の値が大きい検索ワードが説明能力の高いものと考えられる。

6. 参考文献

1. R. F. Engle, “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, *Econometrica*, Vol. 50, (4) (1982), pp. 987–1007
2. G. P. Smith, “Google Internet search activity and volatility prediction in the market for foreign currency”, *Finance Research Letters*, Vol. 9, (2) (2012), pp. 103–110
3. S. A. Cheong, R. P. Forna, G. H. T. Lee, J. L. Kok, W. S. Yim, D. Y. Xu, and Y. Zhang, “The Japanese Economy in Crises: A Time Series Segmentation Study”, *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 6, (2012–5) (2012)
4. B. Efron, “Bootstrap Methods: Another Look at the Jackknife”, *The Annals of Statistics*, Vol. 7 (1) (1979), pp. 1–26