

高頻度金融時系列を用いた外国為替市場の大規模分析

佐藤彰洋

京都大学大学院情報学研究科数理工学専攻

外国為替市場の短時間での構造変化を定量化・可視化するために 24 種類の通貨からなる 46 種類の通貨ペアに対して、スペクトル距離を用いた実証分析を行った。分析の結果、気配値頻度時系列とベストアスクレート時系列の市場全体での類似性構造の変化に弱いながらも関連があることが見出された。更に、通貨交換レートの類似性構造の変化が短時間で大規模に発生することを確認した。

Exhousive analysis of the foreign exchange market with high-frequency financial data

Aki-Hiro Sato

Department of Applied Mathematics and Physics,
Graduate School of Informatics, Kyoto University

In order to quantify and visualize states of market participants with high resolution empirical analysis by using spectral distances is conducted. Empirically it is confirmed that total spectral distance of quotation frequencies is related to that of best ask rates. Furthermore it is found that similarity structures of exchange rates vary drastically in time.

1. はじめに

近年の情報通信技術 (ICT) の進歩と普及によって、多くの人間の社会活動に関するデータが蓄積されている。これらのデータを用いることによって、社会活動の様子を網羅的に巨視的観点から分析することが可能になってきている。

金融市場は電子取引システムを導入するようになり、注文と決済がコンピュータを用い電子化されている。その結果、原理的には世界中のほとんどの金融市場に対して、コンピュータネットワークを通じて注文を瞬時に提示し、取引を実行することができるようになってきている。更に、電子取引システムの普及の結果、これまで人間が決済していたときには分らなかった注文と取引の成立に関する詳細なデータ (高頻度経済時系列データ) がコンピュータ上に蓄積されるようになってきた。

この高頻度時系列データを用いることによって、金融市場全体での注文と取引の成立がどのように生じているかを詳細なデータに基づき理解ができるようになりつつある。そして、これまで不明確で不確実であった市場参加者の振舞いを、データに基づいた実証論的な方法により、高精度に理解できる余地が生れてきた^{1)~5)}。更に、情報提供会社を通じ、個人情報保護を考慮した高頻度経済時系列データが販売されるようになり、誰もが高頻度経済時系列データを入手することができる環境が整いつつある。

一方、高頻度時系列データは 1 日当たり数百 MB から数十 GB になることもあり得るが、近年のコンピュータ処理能力の向上とストレージの大容量化によって市販のパソコン程度でもソフトウェアさえ整えば網羅的に高頻度時系列データの分析が可能となってきている。このように、高頻度時系列データの使用と網羅的分析は多くの人々が実行できる環境が整いつつある。

本稿では外国為替市場の高頻度時系列データを用い、外国為替市場で取引されている通貨ペアの網羅的分析を行い、短時間の構造変化を定量化し可視化することで、市場の状態 (=市場参加者の状態) を俯瞰する数理的手法について考察する。

2. 外国為替市場

新聞やニュースにおいて報道されているように、日本円と米ドルの交換レートが毎日変化していることは周知の事実であるが、その時表示されている価格はどのような意味があるのだろうか。例えば 107.82-107.94 という表示について考えてみよう。前の 107.82 という数字であるが、これは 1 ドルに対して 107.82 円の買い注文価格がその瞬間最高値であることを意味している。一方、後ろの 107.94 という数字は売り注文価格の最安値が 1 ドル 107.94 円であることを示している。これら 2 つの価格を見ることによって、現在の妥当な取引価格が 1 ドル 107.82 円から 107.94 円の間にあることが推測される。

このようなトレーダーの注文気配のデータが近年入手可能となってきており、このデータを用いることで、外国為替市場の様子がわかり、ひいては市場の状態がデータに基づき見えてくると期待できる。

本分析では、CQG 社が提供する注文気配値のデータベースである Time & Sales データ (TSD) を用いて分析を行った。TSD は 1 分の解像度でブローキングシステムに提示された売り注文価格 (ask) と買い注文価格 (bid) が記録されているデータである。分析対象とした通貨は図. 1 に示した 24 通貨からなる 46 通貨ペアである。

外国為替市場の状態を特徴付ける指標として、気配値データ列から抽出される単位時間当りの気配値の到着頻度 (気配値頻度) と単位時間内の最小の売り気配値 (ベストアスクレート) を用いた。気配値頻度は

$$A_j(k) = \frac{1}{\Delta t} C_j(k\Delta t; (k+1)\Delta t), \quad (1)$$

と定義される。ここで $C_j(t_1; t_2)$ は j 番目の通過ペアの t_1 から t_2 の間の気配値の出現回数である。ベストアスクレートは

$$R_j(k) = \begin{cases} \min_{k\Delta t \leq t' \leq (k+1)\Delta t} \{ask_j(t')\} \\ (A_j(k) \neq 0) \\ R_j(k-1) \\ (A_j(k) = 0) \end{cases}, \quad (2)$$

と定義される。ここで、 $ask_j(t)$ は j 番目の通過ペアの時刻 t における売り注文レートを意味する。本分析では $\Delta t = 1[\text{min}]$ とした。

3. 手 法

サンプリング間隔 Δt でサンプルされた M 次元の多変量時系列 $x_j(k)$ ($j = 1, \dots, M; k = 0, 1, \dots, L-1$) を考える。

多変量時系列の分析手法として、VAR モデルを用いた分析方法⁷⁾ や相関関数を用いた分析方法が提案されている。VAR モデルでは M 次元多変量時系列を分析する場合、 M^2 個のパラメータを推定する必要がある。時系列の次元が増加するにつれて M^2 個のパラメータを推定するための計算時間は急激に増加する。

また、相関関数を用いた方法では時系列の非正常性の影響を強く受け、相関の有意性を確認することが難しい場合があることが知られている⁸⁾。ここでは、より勘弁に次元数 M の大きな多変量時系列に対して計算が可能な方法としてスペクトル距離を用いる方法を考察する^{9),10)}。この方法は、時系列のパワースペクトルの形状を比較することにより、時系列の類似性を定量化する手法である。

時系列 $x_j(k)$ のパワースペクトルを Hanning 窓

$$w(k) = \frac{1}{2} \left(1 - \cos\left(\frac{2\pi k}{N-1}\right) \right), \quad (3)$$

を用いたベリオドグラム推定法により推定する。

$$P_j(f_n, t) = \frac{1}{N^2} \left| \sum_{k=0}^{N-1} w(k) x_j(k+t) e^{-2\pi i k \frac{f_n}{N}} \right|^2 \quad \left(f_n = \frac{n}{N\Delta t} \right), \quad (4)$$

ここで、Nyquist 周波数 $f_c = 1/(2\Delta t)$ である。ここで、局在化間隔は人間の活動時間のおおよその単位である 8 時間とし、 $N = 480$ と設定した。そして、規格化パワースペクトル

$$p_j(f_n, t) = \frac{P_j(f_n, t)}{\sum_{n=1}^{N-1} P_j(f_n, t)}, \quad (5)$$

を周波数領域における確率密度関数と見做し⁶⁾、確率密度関数の類似性を定量化する指標として用いられる Jensen-Shannon ダイバージェンス (JS)¹¹⁾、

$$J(t) = H\left(\sum_{j=1}^M \pi_j p_j, t\right) - \sum_{j=1}^M \pi_j H(p_j, t), \quad (6)$$

と Kullback-Leibler ダイバージェンス (KL)¹²⁾

$$KL_{lm}(t) = \sum_{n=1}^{N-1} p_l(f_n, t) \log \frac{p_l(f_n, t)}{p_m(f_n, t)} \quad (7)$$

の値により時系列間の類似性とし、外国為替市場から得られる多変量時系列データを用いて市場状態の定量化を行う。

4. 分析結果

分析対象とした 46 通貨ペアの気配値頻度時系列とベストアスクレートに対して 480 分の局所化した規格化パワースペクトルを用いて JS と KL より時系列の類似性を定量化した。図. 2 は 2007 年 6 月 4 日から 2007 年 6 月 8 日までの時系列を用いて計算した JS の時系列である。時間は時系列を局所化した時刻を表現している。気配値頻度時系列の類似度とベストアスクレートの類似度の時系列はしばしば同調し、気配値頻度時系列の類似性と、ベストアスクレートの類似性とは一定の対応関係が存在する。また、ある時刻に類似性が極端に変化する時期が確認される。気配頻度時系列の JS、 $J_A(t)$ とベストアスクレートの JS、 $J_R(t)$ の相互相関関数を

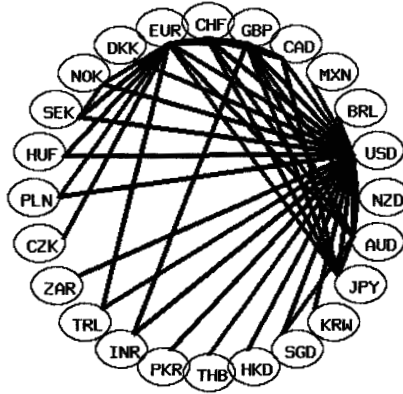
$$C_{AR} = \frac{\langle J_A J_R \rangle - \langle J_A \rangle \langle J_R \rangle}{\sqrt{\text{Var}\{J_A\}} \sqrt{\text{Var}\{J_R\}}}, \quad (8)$$

により定義する。2007 年 6 月 4 日から 2007 年 6 月 8 日までの JS の値は $C_{AR} = 0.357$ となった。このことから両者の気配値頻度時系列とベストアスクレートの変化パターンは弱いながらも関連していると考えられる。

図. 3 は 2007 年 6 月 5 日 0:00 から 8:00(UTC+2) の時系列から計算された KL である。JS で確認された

図 1 ISO 4217 に立脚した通貨表記と通貨名称および分析に用いた 46 通貨ペアのネットワークによる表現。ノードは通貨、2 つのノード間を結ぶリンクはデータに含まれる通貨ペアを表現している。

Fig. 1 Currency abbreviations based on ISO 4217 and currency names. Graphical illustration of currency paris included in the database. Each node represents currency, and link between two nodes currency pair.



NZD	ニュージーランド・ドル	AUD	オーストラリア・ドル	SGD	シンガポール・ドル
HKD	香港・ドル	KRW	大韓民国・ウォン	JPY	日本・円
THB	タイ・バーツ	INR	インド・ルピー	PKR	パキスタン・ルピー
ZAR	南アフリカ・ランド	TRL	トルコ・リラ	EUR	欧州連合・ユーロ
DKK	デンマーク・クローネ	CZK	チェコ・クルナ	PLN	ポーランド・ズウォティ
HUF	ハンガリー・フォリント	CHF	スイス・フラン	GBP	イギリス・ポンド
SEK	スウェーデン・クローナ	NOK	ノルウェー・クローネ	CAD	カナダ・ドル
MXN	メキシコ・ペソ	BRL	ブラジル・レアル	USD	アメリカ合衆国・ドル

こと同様に、気配値頻度時系列の類似性とベストアスクレート時系列の類似性に一部対応関係が見られる。しかしながら、その相関はあまり強くない。更に、JS と KL の平均値との間には以下の比例関係が成立する¹³⁾。

$$J(t) \propto \frac{1}{M^2} \sum_{l=1}^M \sum_{m=1}^M KL_{lm}(t). \quad (9)$$

5. 結 論

本稿では外国為替市場の高頻度時系列データを用いた市場状態の定量化手法にスペクトル距離を用い、外国為替市場で取引される 24 通貨からなる 46 通貨ペアの気配値行動頻度時系列とベストアスクレートの総合的な類似性構造の時間変化を調べた。その結果、外国為替市場の類似性構造が時間的に変化している様子をとらえることができた。今後の課題として、超高次元多変量時系列の類似性構造の特徴抽出方法の開発および長期間にわたる外国為替市場構造変化の様子を調べることにより、類似性構造の時間変化の性質を調べることが挙げられる。

謝 辞

本研究を行う上で宗像豊哲氏(京都大学)、五十嵐顕人氏(京都大学)には貴重なコメントを戴きました。謝意を表します。

参 考 文 献

- 1) トニス・ヴァーガ著, 新田功, 永原裕一訳: *複雑系と相場*, 白桃書房(1999).
- 2) Dacorogna, M.M., Gençay, R., Müller, U., Olsen, R.B., and Pictet, O.V.: *An introduction to high-frequency finance*, Academic Press (2000).
- 3) Mantegna, R. and Stanley, H.: *An Introduction to Econophysics Correlations and Complexity in Finance*, Cambridge University Press (2000).
- 4) 高安秀樹, 高安美佐子: *エコノフィジックス 市場に潜む物理法則*, 日本経済新聞社(2001).
- 5) デイディエ・ソネット著, 森谷博之監訳: *[入門] 経済物理学*, PHP 研究所(2003).
- 6) Bein, B.: *Entropy, Best Practice & Research Clinical Anaesthesiology*, Vol. 20, No. 1, pp. 101-

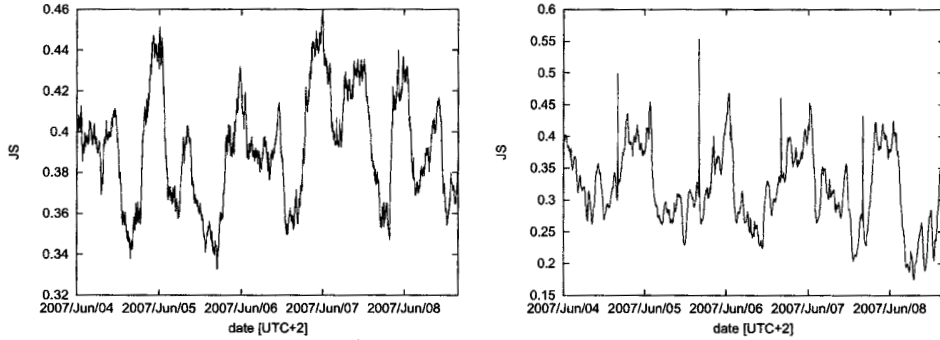


図2 気配値頻度時系列(左)とベストアスクレート(右)の46通貨ペア間の Jensen-Shannon ダイバージェンスの時系列.

Fig. 2 Jensen-Shannon spectral divergence among 46 currency pairs of quotation frequencies (left) and best ask rates (right) by using the data for a period from 0:00 to 8:00 on 5th June 2007.

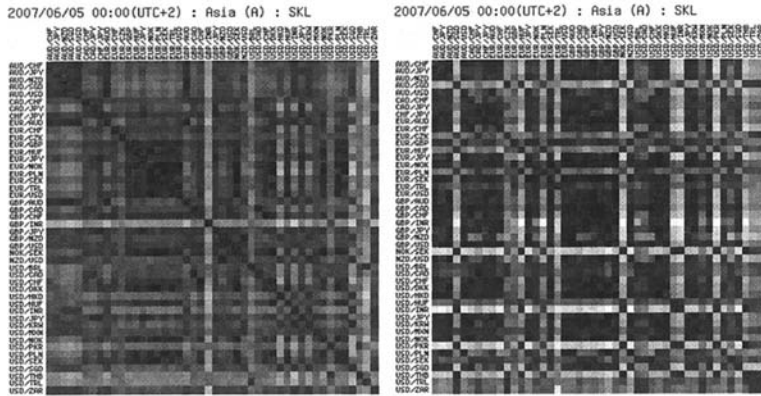


図3 気配値頻度時系列(左)とベストアスクレート(右)の46通貨ペア間の Kullback-Leibler ダイバージェンスの値. 黒/白は2つの時系列間の類似度が高い/低いことを意味する.

Fig. 3 Kullback-Leibler spectral divergence matrices among 46 currency pairs of quotation frequencies (left) and best ask rates (right) by using the data for a period from 4th June to 8th June 2007. A black/white pixel represents a similar/disimilar relation.

109 (2006).
 7) 赤池弘次, 北川源四郎編: **時系列解析の実際 I/II**, 朝倉書店 (1995).
 8) Rybski, D., Havlin, S., Bunde, A.: Phase synchronization in temperature and precipitation records, Vol. 320, *Physica A*, pp. 601-610 (2003).
 9) Sato, A.-H., and Oshiro, J.: Quantifying similarity between markets with application to high frequency financial data, *Journal of the Physical Society of Japan*, Vol. 75 pp.0845005-1-4 (2006).
 10) Sato A.-H.: Frequency analysis of tick quotes on the foreign exchange market and agent-based

modeling: A spectral distance approach, *Physica A*, Vol. 382, pp. 258-270 (2007).
 11) Lin, J.: Divergence Measures Based on the Shannon Entropy, *IEEE transactions on information theory*, Vol. 37, No. 1, pp.145-150 (1991).
 12) Veldhuis, R., and Klabbbers, E.: On the computation of the Kullback-Leibler measure for spectral distance, *IEEE transactions on speech and audio processing*, Vol. 11, pp. 100-103 (2003).
 13) 佐藤彰洋: 多体 Jensen-Shannon スペクトル距離に基づいた市場参加者行動特性の定量化, *JWEIN2007 講演論文集*, pp. 48-55 (2007).