

明 細 書

発明の名称：

時系列データの変化点検出方法及びプログラム、未来の時系列データ値の確率密度分布予測方法及びプログラム

技術分野

[0001] 本発明は時系列データの変化点検出方法及びプログラム、未来の時系列データ値の確率密度分布予測方法及びプログラムに関する。

背景技術

[0002] 今日、様々な分野において、各種の時系列データが用いられている。例えば、国際金融における外国為替の変動や、証券取引における株価などが時系列データとして用いられている。他にも、例えば半導体製造などの製造業において、歩留まりトレンドの管理などが行われる。

[0003] こうした時系列データは、一般に、トレンドの変化点を検出し、その変化に応じた行動及び対策を取るために用いられる。例えば為替取引における円／ドル取引においては、値動きの動向、特に大きくトレンドが変化する点を検出し、売買方針を適宜修正することが求められる。こうした時系列データでは、過去の時系列データの移動平均を用いて長期的なトレンドを見積もることが行われる。身近な例として、円／ドル取引では、為替チャートに25日間や13週間などの移動平均線が表示される。また、こうした時系列データの長期トレンド分析においては、移動平均を中心とするランダムウォークを適用する手法が広く用いられている。

[0004] ところが、為替取引などでは、値動きに連動した売り注文又は買い注文が集中するなどの事象が生じることで、短期的に急激な値動きを生じる場合がある。この場合、単に過去のデータの移動平均を中心とするランダムウォークのみでは、こうした急峻なトレンドの変化を検出することは困難である。つまり、移動平均は、過去の複数のデータについての平均であるので、直近のデータが大きな変動を示したとしても、母集団に比べて直近のデータ数が

少なければ、ランダムウォークを適用しても急峻なトレンド変化を捉えることができない。

[0005] 上述の問題を解決するため、発明者らは、時系列データの急峻な変化を検出するPUC K (Potentials of Unbalanced Complex Kinetics) モデルを提案した(非特許文献1)。PUC Kモデルでは、金融等の分野における時系列データの変動を、時刻(t)での真の市場価格(Noise-Reduced Price) $P(t)$ 、時刻(t)での市場価格の観測誤差(Observation Noise) $f_Y(t)$ 、時刻(t)で観測した市場価格(best-bid, best-ask, Mid-quote Data等) $Y(t)$ の3要素により、以下の式(1)で表される。

[数1]

$$Y(t) = P(t) + f_Y(t) \quad \dots (1)$$

[0006] ティックとは、取引市場において、約定成立による取引対象物の価格の変動が生じるイベントのことをいう。よって、1ティック間とは、ある時点での約定タイミングと、それよりも1つ前の約定タイミングとの間の時間間隔を意味する。また、ティック数とは、ある期間内における約定成立回数、すなわち取引対象物の価格の変動が生じるイベントの発生回数を意味する。以降(t)は、ティック数で数えた時刻または実時間を表す。PUC Kモデルでは、数ティックスケールの変位の逆相関性を除去した後の真の市場価格 $P(t)$ の変動を取り扱う。1ティックの間での真の市場価格の変動は、以下の式(2)で表すことができる。

[数2]

$$P(t+1) - P(t) = -\frac{b(t)}{M-1} (P(t) - P_M(t)) + f_P(t) \quad \dots (2)$$

但し、式(2)の右辺第1項は、1ティックの間での真の市場価格の変動に作用するポテンシャルの寄与分を表し、右辺第2項の $f_P(t)$ は、時刻(t)での真の市場価格を変動させる変動誤差を表す。 $b(t)$ は時刻(t)でのポテンシャル係数、Mは価格変動の中心を見積もるために必要な直近のデ

一タ数、 $P_M(t)$ は直近 M 個の真の市場価格を用いて見積もった価格変動の中心価格である。また、式(2)における時刻 (t) でのポラシナル係数 $b(t)$ は、直近のデータ数 M に対して依存性を有し、 $\{1/(M-1)\}$ に比例して減少することが経験的に知られている。よって、式(2)では、時刻 (t) でのポラシナル係数 $b(t)$ の直近のデータ数 M への依存性を避けるため、時刻 (t) でのポラシナル係数 $b(t)$ に $\{1/(M-1)\}$ を乗じている。

[0007] 中心価格 $P_M(t)$ は、以下の式(3)で表される。

[数3]

$$P_M(t) = \frac{1}{M-1} \sum_{h=0}^{M-1} P(t-h) \quad \dots (3)$$

[0008] さらに、式(2)は、時々刻々変化する線形な中心力が働く場の中でのラ

ンダムウォークとみなすと、時刻 (t) でのポラシナル係数 $b(t)$ を2次以上のポラシナルの係数と捉え、式(2)の右辺第1項が表すポラシナルの寄与分にかかるポラシナルを、ポラシナル $W(q, t)$ として、以下の式(4)で表わすことができる。但し、 $!$ はポラシナルの次数を示す。

[数4]

$$W(q, t) = \sum_{i=2}^K b_i(t) q^i \quad (K \geq 2) \quad \dots (4)$$

価格変動の平均速度の絶対値は、ポラシナル関数の接線勾配の大きさに比例する。価格変動の期待される方向は、勾配を下る方向である。したがって、ポラシナル関数のどこに今の市場価格があるのかを理解するためには、中心価格と真の市場価格の乖離を計算し、ポラシナル関数のどこに市場価格があるのかを理解する必要がある。

さらに、式(4)を用いると、式(2)を以下の式(5)に書き換えることができる。

[数5]

$$P(t+1) - P(t) = -\frac{d}{dq} W(q, t) \Big|_{q = \frac{P(t) - P_M(t)}{M-1}} + f_P(t) \quad \dots (5)$$

[0009] 上述のように、式（４）及び（５）において、時刻（ t ）でのポテンシャル係数 $b_i(t)$ の値により、ポテンシャル $W(q, t)$ の寄与度が決定される。例えば、時刻（ t ）で各次のポテンシャル係数 $b_i(t) = 0$ であれば、 $P(t+1) - P(t) = f_P(t)$ となり、変動誤差の成分のみが価格差に寄与する。一方、時刻（ t ）での各次のポテンシャル係数 $b_i(t)$ の絶対値が大きくなるにつれて、ポテンシャル $W(q, t)$ の寄与が大きくなる。その結果、ランダムウォークモデルで計算される拡散係数よりも速い拡散、遅い拡散が時刻（ t ）での各次のポテンシャル係数 $b_i(t)$ の関数として定量的に推定できることが知られている（以下、異常拡散と称する）。特に、２次ポテンシャルの係数 $b_2(t) \leq -2$ の場合には、市場価格の変動は、もはや過去のデータ移動平均から逸脱し、発散してしまうことが知られている。

[0010] このように、PUCモデルでは、２次以上のポテンシャル項を考える事により、上述のような短期間の異常拡散を表現することが可能となる。これにより、これまでのランダムウォークのみを用いた手法では実現できない、異常拡散の検出を行うことができる。

先行技術文献

非特許文献

[0011] 非特許文献1：Misako Takayasu et. al., "Econophysics Approaches to Large-Scale Business Data and Financial Crisis", Springer Japan, June 2010, pp.79-98.

非特許文献2：Kenta Yamada, Hideki Takayasu, Takatoshi Ito and Misako Takayasu, "Solvable Stochastic Dealer Model for Financial Markets", Physical Review E, 79, Issue 5, 051120 (2009).

非特許文献3：中村和幸、樋口知之、「最近のベイズ理論の進展と応用」[1]

「一逐次ベイズとデータ同化」、2009年、電子情報通信学会誌、Vol. 92、No. 12、p1062～1067

発明の概要

発明が解決しようとする課題

[0012] ところが、発明者らは、上述のPUCKモデルを用いる手法では、以下のような問題点があることを見出した。今日の金融市場などでは、コンピューターシステムを用いた高速の自動取引などが導入されており、価格を示す時系列データの変動を迅速かつ正確に検出することが求められる。しかしながら、通常のPUCKモデルでは、時刻(t)での各次のポテンシャル係数 $b_i(t)$ を見積もるには、1000ティック程度のデータを用い、式(2)の $\{P(t+1) - P(t)\}$ と $\{P(t) - P_M(t)\}$ の値をプロットし、最小二乗法などでその散布図の傾きを求めて時刻(t)でのポテンシャル係数 $b(t)$ (以下、特に断らない限り、各次のポテンシャル係数 $b_i(t)$ ($2 \leq i \leq K$)の集合を、単にポテンシャル係数 $b_i(t)$ と称する)を計測する必要があった。そのため、データのサンプリング開始から1000個程度の時系列データが収集されるまでは、過去情報が入れ替わらず、急峻なトレンド変化への追従性については、実用上、未だ不足していた。

[0013] また、モデル上、時刻(t)でのポテンシャル係数 $b(t)$ の値は、直近のデータ数Mに依存しないはずであるが、観測から式(2)を満たす適切な直近のデータ数Mの範囲が時間ごとに変わる効果が無視できず、時刻(t)でのポテンシャル係数 $b(t)$ の推定誤差が大きくなるという問題点があった。そこで、PUCKモデルに合う最適な直近のデータ数Mを逐次探すが重要な課題であった。

[0014] さらに、時刻(t)における最適な真の価格 $P(t)$ を推定するために、数ティック程度の最適移動平均の重みを、1日程度ごとの全データを用いて別途計算を更新しておく必要があった。

[0015] 本発明は上記の事情に鑑みて成されたものであり、本発明の目的は、時系列データの変化点検出方法及びプログラム、未来の時系列データ値の確率密

度分布予測方法及びプログラムを提供することである。

課題を解決するための手段

[0016] 本発明の第1の態様である時系列データの変化点検出方法は、時刻($t-1$)での真の市場価格及び時刻($t-1$)までの M (M は、正の整数)個の真の市場価格の移動平均である中心価格によって定義されるポテンシャル項と変動誤差との和によって、時刻(t)での真の市場価格 $P(t)$ を算出するPUCKモデルに、粒子フィルタ法を適用する時系列データの変化点検出方法であって、それぞれ異なる値を有するPUCKモデルの状態を表すパラメータを持つ粒子群より次の時刻のパラメータの確率密度関数を得る第1のステップと、前記複数の粒子のそれぞれについて、時刻(t)で観測した市場価格に対する前記時刻(t)での真の市場価格の適合度を評価する第2のステップと、

前記適合度に応じて、前記複数の粒子の中から粒子を再サンプリングする第3のステップと、を備え、前記第3のステップでは、乱数を生成し、第1の所定の値と前記乱数とを比較し、前記乱数が前記第1の所定の値よりも大きい場合には、時刻(t)でのPUCKモデルの状態を表すパラメータを平均値とする確率密度分布を成分とする確率密度関数を粒子として生成し、前記乱数が前記第1の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成するものである。これにより、時系列データである市場価格に急峻なトレンド変化が生じて、良好な追随性を与えるポテンシャル係数を算出し、ポテンシャル係数の変動を検出することにより変化点検出ができる。

[0017] 本発明の第2の態様である時系列データの変化点検出方法は、上述の時系列データの変化点検出方法において、時刻(t)でのPUCKモデルの状態を表すパラメータを平均値とする確率密度分布は、時刻(t)でのPUCKモデルの状態を表すパラメータを平均値とする正規分布であるものである。これにより、正規分布を用いて、具体的に確率密度分布を生成することができる。

[0018] 本発明の第3の態様である時系列データの変化点検出方法は、上述の時系列データの変化点検出方法において、前記時刻（ $t + 1$ ）での真の市場価格は、前記PUCモデルに基づいたポテンシャル項と、時刻（ t ）での真の市場価格の変動誤差と、を加算することにより算出されることを特徴とするものである。これにより、粒子フィルタ法を用いた再サンプリングを好適に行うことができる。

[0019] 本発明の第4の態様である時系列データの変化点検出方法は、上述の時系列データの変化点検出方法において、上述の時系列データの変化点検出方法において、前記第3のステップでは、それぞれの粒子に関して乱数を生成し、前記第1の所定の値と前記乱数とを比較し、前記乱数が前記第1の所定の値よりも大きい場合には、時刻（ t ）におけるMを前提とする条件付き確率密度関数を粒子として生成し、前記乱数が前記第1の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成するものである。これにより、ポテンシャル係数のみならず、時刻（ t ）から見て直近の時系列データ数を示すMについても同時に時間発展を適用することができる。

[0020] 本発明の第5の態様である時系列データの変化点検出方法は、上述の時系列データの変化点検出方法において、前記時刻（ $t + 1$ ）での真の市場価格は、前記ポテンシャル項と、時刻（ t ）での真の市場価格の変動誤差と、を加算することにより算出されることを特徴とするものである。これにより、時刻（ t ）での真の市場価格を精度よく見積もることが可能となる。

[0021] 本発明の第6の態様である時系列データの変化点検出方法は、上述の時系列データの変化点検出方法において、前記時刻（ t ）で観測した市場価格は、前記時刻（ t ）での真の市場価格と時刻（ t ）での市場価格の観測誤差との和であることを特徴とするものである。これにより、前記時刻（ t ）で観測した市場価格に対する時刻（ t ）での真の市場価格の適合度を、精度よく見積もることができる。

[0022] 本発明の第7の態様である時系列データの変化点検出方法は、上述の時系

列データの変化点検出方法において、前記時刻（ t ）での真の市場価格の変動誤差及び前記時刻（ t ）での市場価格の観測誤差は、正規分布等に従う確率密度関数で与えられることを特徴とするものである。これにより、時刻（ t ）での真の市場価格及び前記時刻（ t ）で観測した市場価格に対する時刻（ t ）での真の市場価格の適合度を、精度よく見積もることができる。

[0023] 本発明の第8の態様である未来の時系列データ値の確率密度分布の予測方法は、上述の時系列データの変化点検出方法により算出した時刻（ t ）における前記ポテンシャル項のポテンシャル係数、前記Mの値、及び時刻（ t ）での真の市場価格の変動誤差と、に基づいて、前記時刻（ t ）での真の市場価格 $P(t)$ を時間発展させることにより、時刻（ $t+N$ ）での前記時刻（ t ）での真の市場価格 $P(t)$ を算出するものである。これにより、PUCCKモデルを用いて、未来の時系列データ値の確率密度分布を予測することができる。

[0024] 本発明の第9の態様である未来の時系列データ値の確率密度分布の予測方法は、上述の未来の時系列データ値の確率密度分布の予測方法において、前記粒子の総数を N_p 、前記粒子を特定する番号を j （ただし、 j は $1 \leq j \leq N_p$ を満たす整数）、番号 j の時刻（ t ）での粒子のポテンシャル係数を $b^{(j)}_i(t)$ （ただし、 i はポテンシャルの次数を示す2以上の整数）、前記番号 j の粒子の時刻（ t ）での真の市場価格の変動誤差を $f^{(j)}_p(t)$ として、前記番号 j の粒子の時刻（ t ）での真の市場価格を $P^{(j)}(t)$ 、前記番号 j の粒子の時刻（ $t+N$ ）での真の市場価格を $P^{(j)}(t+N)$ 、前記番号 j の粒子の時刻（ t ）での市場価格の中心価格を $P^{(j)}_M(t)$ 、時刻（ t ）におけるMの値を $M(t)$ とした場合に、前記番号 j の粒子の前記時刻（ $t+N$ ）での真の市場価格 $P^{(j)}(t+N)$ は、 K を2以上の整数として、

[数6]

$$P^{(j)}(t+N) = P^{(j)}(t) - \sum_{n=0}^{N-1} \left\{ \sum_{i=2}^K b^{(j)}_i(t) \left(\frac{P^{(j)}(t+n) - P^{(j)}_M(t+n)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_p(t)$$

で表されるものである。これにより、PUCKモデルを用いて、未来の時系列データ値の確率密度分布を具体的に予測することができる。

[0025] 本発明の第10の態様である未来の時系列データ値の確率密度分布の予測方法は、上述の未来の時系列データ値の確率密度分布の予測方法において、番号 j の粒子の時刻 $(t+1)$ でのポテンシャル係数 $b^{(j)}_{t+1}$ は、前記番号 j の粒子の時刻 (t) でのポテンシャル係数 $b^{(j)}_t$ を時間発展させることにより算出されるものである。これにより、PUCKモデルを用いて、未来の時系列データ値の確率密度分布を具体的に予測することができる。

[0026] 本発明の第11の態様である時系列データの変化点検出プログラムは時刻 $(t-1)$ での真の市場価格及び時刻 $(t-1)$ までの M (M は、正の整数)個の真の市場価格の移動平均である中心価格によって定義されるポテンシャル項と変動誤差との和によって時刻 (t) での真の市場価格 $P(t)$ を算出するPUCKモデルに粒子フィルタ法を適用する時系列データの変化点検出をコンピュータに実行させる時系列データの変化点検出プログラムであって、それぞれ異なる値を有するPUCKモデルの状態を表すパラメータを持つ粒子群を生成することによりパラメータの確率密度関数を得る第1の処理と、適合度評価部に、前記複数の粒子のそれぞれについて、時刻 (t) で観測した市場価格に対する前記時刻 (t) での真の市場価格の適合度を評価する第2の処理と、サンプリング部に、前記適合度に応じて、前記複数の粒子の中から粒子を再サンプリングさせる第3の処理と、をコンピュータに実行させ、前記乱数が前記第1の所定の値よりも大きい場合には、時刻 (t) でのPUCKモデルの状態を表すパラメータを平均値とする確率密度分布を成分とする確率密度関数を粒子として再生成する処理と、前記乱数が前記第1の所定の値よりも小さい場合には、一様分布にしたがって粒子を生成する処理と、を実行させるものである。これにより、時系列データである市場価格に急峻なトレンド変化が生じて、良好な追随性を与えるポテンシャル係数を算出し、ポテンシャル係数の変動を検出することにより変化点検出ができ

る。

[0027] 本発明の第12の態様である時系列データの変化点検出プログラムは、上述の変化点検出プログラムにおいて、時刻(t)でのPUCKモデルの状態を表すパラメータを平均値とする確率密度分布は、時刻(t)でのPUCKモデルの状態を表すパラメータを平均値とする正規分布であるものである。これにより、正規分布を用いて、具体的に確率密度関数を生成することができる。

[0028] 本発明の第13の態様である時系列データの変化点検出プログラムは、上述の変化点検出プログラムにおいて、前記第3のステップでは、前記適合度が大きい粒子ほど、多くサンプリングされるだけでなく、一定の割合で、実現可能性のある適合度が低い粒子を生成すること、を特徴とするものである。これにより、粒子フィルタ法を用いた再サンプリングを好適に行うことができる。

[0029] 本発明の第14の態様である時系列データの変化点検出プログラムは、上述の変化点検出プログラムにおいて、前記第1のステップでは、乱数を生成する処理と、時刻(t)におけるPUCKモデルのパラメータを前提とする条件付き確率密度関数を粒子として生成する処理と、前記乱数が前記第3の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成する処理と、を実行させることを特徴とするものである。これにより、ポテンシャル係数のみならず、時刻(t)から見て直近の時系列データ数を示すMについても同時に時間発展を適用することができる。

[0030] 本発明の第15の態様である時系列データの変化点検出プログラムは、上述の変化点検出プログラムにおいて、前記時刻(t)での真の市場価格は、前記ポテンシャル項と、時刻(t)での真の市場価格の変動誤差と、を加算することにより算出されることを特徴とするものである。これにより、前記時刻(t)で観測した市場価格に対する時刻(t)での真の市場価格の適合度を、精度よく見積もることができる。

[0031] 本発明の第16の態様である時系列データの変化点検出プログラムは、上

述の変化点検出プログラムにおいて、前記時刻（ t ）で観測した市場価格は、前記時刻（ t ）での真の市場価格と時刻（ t ）での市場価格の観測誤差との和であることを特徴とするものである。これにより、前記時刻（ t ）で観測した市場価格に対する時刻（ t ）での真の市場価格の適合度を、精度よく見積もることができる。

[0032] 本発明の第17の態様である時系列データの変化点検出プログラムは、上述の変化点検出プログラムにおいて、前記時刻（ t ）での真の市場価格の変動誤差及び前記時刻（ t ）での市場価格の観測誤差は、正規分布等に従う確率密度関数で与えられることを特徴とするものである。これにより、時刻（ t ）での真の市場価格及び前記時刻（ t ）で観測した市場価格に対する時刻（ t ）での真の市場価格の適合度を、精度よく見積もることができる。

[0033] 本発明の第18の態様である未来の時系列データ値の確率密度分布の予測プログラムは、上述の時系列データの変化点検出プログラムにより算出した時刻（ t ）における前記ポテンシャル項のポテンシャル係数、前記Mの値、及び時刻（ t ）での真の市場価格の変動誤差と、に基づいて、前記時刻（ t ）での真の市場価格 $P(t)$ を時間発展させることにより、時刻（ $t+N$ ）での前記時刻（ t ）での真の市場価格 $P(t)$ を算出するものである。これにより、PUCKモデルを用いて、未来の時系列データ値の確率密度分布を予測することができる。

[0034] 本発明の第19の態様である未来の時系列データ値の確率密度分布の予測プログラムは、上述の未来の時系列データ値の確率密度分布の予測プログラムにおいて、前記粒子の総数を N_p 、前記粒子を特定する番号を j （ただし、 j は $1 \leq j \leq N_p$ を満たす整数）、番号 j の粒子の時刻（ t ）でのポテンシャル係数を $b^{(j)}_i(t)$ （ただし、 i はポテンシャルの次数を示す2以上の整数）、前記番号 j の粒子の時刻（ t ）での真の市場価格の変動誤差を $f^{(j)}_p(t)$ として、前記番号 j の粒子の時刻（ t ）での真の市場価格を $P^{(j)}(t)$ 、前記番号 j の粒子の時刻（ $t+N$ ）での真の市場価格を $P^{(j)}(t+N)$ 、前記番号 j の粒子の時刻（ t ）での市場価格の中心価格を $P^{(j)}_M(t)$

）、時刻（ t ）における M の値を $M(t)$ とした場合に、前記番号 j の粒子の前記時刻（ $t+N$ ）での真の市場価格 $P^{(j)}(t+N)$ は、 K を2以上の整数として、

[数7]

$$P^{(j)}(t+N) = P^{(j)}(t) - \sum_{n=0}^{N-1} \left\{ \sum_{i=2}^K \left\{ b^{(j)}_i(t) \left(\frac{P^{(j)}(t+n) - P_M^{(j)}(t+n)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_P(t) \right\}$$

で表されるものである。これにより、PUCKモデルを用いて、具体的に未来の時系列データ値の確率密度分布を予測することができる。

[0035] 本発明の第20の態様である未来の時系列データ値の確率密度分布の予測プログラムは、上述の未来の時系列データ値の確率密度分布の予測プログラムにおいて、番号 j の粒子の時刻（ $t+1$ ）でのポテンシャル係数 $b^{(j)}_i(t+1)$ は、前記番号 j の粒子の時刻（ t ）でのポテンシャル係数 $b^{(j)}_i(t)$ を時間発展させることにより算出されるものである。

発明の効果

[0036] 本発明によれば、時系列データの急峻な変化を精度よく検出できる時系列データの変化点検出方法及びプログラム、未来の時系列データ値の確率密度分布予測方法及びプログラムすることができる。

図面の簡単な説明

[0037] [図1]実施の形態1で用いる粒子フィルタ法の手順を示すフローチャートである。

[図2]実施の形態1にかかる式（24）で与えられる確率分布を示すグラフである。

[図3]実施の形態1にかかる式（25）で与えられる確率分布を示すグラフである。

[図4]実施の形態1における時刻（ t ）でのポテンシャル係数 $b(t)$ の決定方法を示すフローチャートである。

[図5]実施の形態1における時刻（ t ）での直近のデータ数 $M(t)$ の決定方

法を示すフローチャートである。

[図6]実施の形態2にかかる時系列データの変化点検出プログラムが実行される時系列データの変化点検出装置200の構成を模式的に示すブロック図である。

[図7]実施の形態3にかかる粒子フィルタ法を適用したPUCKモデルの時間発展方法を示すフローチャートである。

[図8]実施の形態3にかかる未来の時系列データがとる値の確率密度分布の予測方法により時刻($t + N$)における市場価格の確率密度分布を求めた場合の確率密度分布の推移を示すグラフである。

[図9]実施の形態3にかかる未来の時系列データがとる値の確率密度分布の予測方法により時刻($t + N$)における市場価格がとる値の確率密度分布を示すグラフである。

[図10]実施の形態4にかかる未来の時系列データがとる値の確率密度分布の予測プログラムが実行される未来の時系列データがとる値の確率密度分布の予測装置400の構成を模式的に示すブロック図である。

[図11]実施の形態5にかかる粒子フィルタ法を適用したPUCKモデルの時間発展方法を示すフローチャートである。

[図12]現在の市場価格を数理モデルに代入して、粒子番号 j 毎に時間発展させ、その経過を粒子毎に表示したグラフである。

[図13]実施の形態6にかかる未来の時系列データがとる値の確率密度分布の予測プログラムが実行される未来の時系列データがとる値の確率密度分布の予測装置600の構成を模式的に示すブロック図である。

発明を実施するための形態

[0038] 以下、図面を参照して本発明の実施の形態について説明する。各図面においては、同一要素には同一の符号が付されており、必要に応じて重複説明は省略される。

[0039] 実施の形態1

本発明の実施の形態1にかかる時系列データの変化点検出方法について説

明する。本実施の形態における時系列データの変化点検出方法は、前述のPUCモデルに、いわゆる粒子フィルタ法（非特許文献3）を適用したものである。粒子フィルタ法を適用することにより、時刻（ t ）でのポテンシャル係数 $b(t)$ 及び直近のデータ数 $M(t)$ の値を、逐次予測することができる。

[0040] まず、粒子フィルタ法について、概要を説明する。図1は、実施の形態1で用いる粒子フィルタ法の手順を示すフローチャートである。逐次モンテカルロ法的一种である粒子フィルタ法では、状態ベクトル x_t が、以下の式（6）で与えられる。つまり、時刻（ t ）における状態ベクトル x_t は、1ティック前の状態ベクトル x_{t-1} にパラメータ θ を与えることにより得られる。

[数8]

$$x_t \sim Q_t(\cdot | x_{t-1}, \theta) \quad \dots (6)$$

ここで、 $\sim$ は、 x_{t-1} と θ をパラメータとし、時刻（ t ）に依存する確率密度関数、 $Q_t(\cdot)$ に従う確率変数 x_t を生成する操作である。

[0041] また、観測ベクトル y_t は、以下の式（7）で与えられる。ここで、 $R_t(\cdot)$ は x_{t-1} と θ をパラメータとし、時刻（ t ）に依存する確率密度関数である。

[数9]

$$y_t \sim R_t(\cdot | x_t, \theta) \quad \dots (7)$$

[0042] 粒子フィルタ法においては、状態ベクトル x_t の確率密度関数は、観測ベクトル $\{y_1, y_2, \dots, y_t\}$ により決定される条件付き確率密度関数CPDFとして、以下の式（8）で表される。

[数10]

$$CPDF = p(x_t | y_1, y_2, \dots, y_t) \quad \dots (8)$$

[0043] ここで、粒子フィルタ法で想定する粒子（状態ベクトル）が N （ N は正の整数）個であるとする、 N 個の状態ベクトルの条件付き確率密度関数（ P

D F) は、以下の式 (9) で表される。

[数11]

$$(N \text{ 個の状態ベクトルの条件付きPDF}) = \sum_{j=1}^N \delta(x_l - x_{l|k}^{(j)}) \quad \dots (9)$$

ここで、 x は、 j ($1 \leq j \leq N$) 番目の粒子において、時刻 k の観測ベクトルから推定された時刻 l の状態ベクトルである。

[0044] 粒子フィルタ法では、 N 個の粒子の条件付き確率密度関数 (PDF) について、まず、 $t = 0$ における初期値を与える (図 1 のステップ S 1)。なお、初期値は、以下の式 (10) で表される。

[数12]

$$(N \text{ 個の状態ベクトルの条件付きPDF}) = \sum_{j=1}^N \delta(x_0 - x_{0|0}^{(j)}) \quad \dots (10)$$

そして、粒子フィルタ法による計算を開始するにあたり、時刻 (t) を $t = 1$ に設定する。

[0045] 次いで、 N 個の粒子について、式 (6) を用いて、時刻 ($t - 1$) における確率密度関数を基に時刻 (t) における予測粒子の確率密度関数 (PDF) を生成する (図 1 のステップ S 2)。なお、 j 番目の粒子の状態ベクトルは、以下の式 (11) で表される。

[数13]

$$x_{t|t-1}^{(j)} \sim q_t(\cdot | x_{t-1|t-1}^{(j)}, \theta) \quad (i = j, \dots, N) \quad \dots (11)$$

ここで、 $q_t(\cdot)$ は、式 (9) によって定義される確率密度関数である。

[0046] 次いで、各粒子の重み係数 W_j (以下の式 (12)) を、尤度 L_j (以下の式 (13)) を用いて計算する。なお、尤度 L_j は、式 (7) より得られる。換言すると、ここでは、実際の時系列データから得られる観測ベクトル y_t と、計算により得られた状態ベクトル x_t と、の適合度を評価する (図 1 のステップ S 3)。この適合度は、尤度を用いることで評価することができる。

[数14]

$$w_j = \frac{L_j}{\sum_{j=1}^N L_j} \quad \dots (12)$$

[数15]

$$L_j = p(y_t | x_{t|t-1}^{(j)}) \quad \dots (13)$$

[0047] その後、各粒子の重み係数に従って、粒子の再サンプリングを行う（図1のステップS4）。つまり、式（11）で示される各粒子を、重み係数に応じた個数だけサンプリングし、合計N個の粒子をサンプリングする。この際、適合度が小さければ重み係数は小さくなるので、その粒子は残らない可能性が高くなる。一方、適合度が大きければ重み係数は大きくなり、その粒子は複製され増加することとなる。以上が粒子フィルタ法の概要である。

[0048] 次に、上記の手順を、時刻（t）を1ステップずつ進めて、最後のステップ（t = t_{end}）に至るまで繰り返し行う（図1のステップS5及びS6）。これにより、時間の進展に応じて、逐次粒子フィルタ法を適用することが可能である。

[0049] 次に、本実施の形態において、PUCKモデルに粒子フィルタ法を用いる手順について説明する。本実施の形態では、図1のステップS2における、確率密度関数の生成時のPUCKモデルの取り扱いに特徴を有する。以下、ステップS2における具体例について説明する。以下では、状態ベクトルを用いて、時刻（t）での各次のポテンシャル係数を $b_i(t)$ 、直近のデータ数を $M(t)$ 等で表す。時刻（t-1）での各次のポテンシャル係数 $b_i(t-1)$ は以下の式（14）、直近のデータ数 $M(t-1)$ は以下の式（15）で表される。但し、 θ_b 及び θ_m はノイズを生成するためのパラメータである。

[数16]

$$b_i(t-1) \sim Q_b(\cdot | b_i(t-2), \theta_b) \quad \dots (14)$$

[数17]

$$M(t-1) \sim \mathcal{Q}_M(\cdot | M(t-2), \theta_M) \quad \dots (15)$$

[0050] 式(15)を前述の式(3)に適用することにより、以下の式(16)を得ることができる。

[数18]

$$P_M(t-1) = \frac{1}{M(t-1)} \sum_{j=0}^{M(t-1)-1} P(t-1-j) \quad \dots (16)$$

[0051] また、式(14)及び(15)を前述の式(2)に適用することにより、以下の式(17)を得ることができる。

[数19]

$$P(t) = P(t-1) - \sum_{i=1}^K \left(\frac{b_i(t-1)(P(t-1) - P_M(t-1))}{M(t-1)-1} \right)^{i-1} + f_P(t) \quad \dots (17)$$

[0052] この場合においても、前述の式(1)と同様に、式(18)が成立する。

[数20]

$$Y(t) = P(t) + f_Y(t) \quad \dots (18)$$

[0053] なお、時刻(t)での真の市場価格の観測誤差 $f_P(t)$ 及び時刻(t)での市場価格の観測誤差 $f_Y(t)$ は、以下の式(19)のように、正規分布等に従うものとする。

[数21]

$$f_P(t) \sim N(0, \sigma_P^2), \quad f_Y(t) \sim N(0, \sigma_Y^2) \quad \dots (19)$$

以下、一般に、 $N(\cdot)$ の部分は、正規分布のみならず、さらに高次のモーメント等で特徴付けられる確率分布でもよい。また、 $f_P(t)$ 、 $f_Y(t)$ は、異なる確率分布でもよい。

[0054] 時刻（ t ）での真の市場価格の観測誤差 $f_P(t)$ の標準偏差 σ_P は、状態ベクトルを用いて、以下の式（20）で表すことができる。

[数22]

$$\sigma_{P,t} \sim Q_P(\cdot | \sigma_{P,t-1}, \theta_P) \quad \dots (20)$$

[0055] 時刻（ t ）での市場価格の観測誤差 $f_Y(t)$ の標準偏差 σ_Y は、状態ベクトルを用いて、以下の式（21）で表すことができる。

[数23]

$$\sigma_{Y,t} \sim Q_Y(\cdot | \sigma_{Y,t-1}, \theta_Y) \quad \dots (21)$$

[0056] 式（20）及び（21）より、式（19）を以下の式（22）に書き換えることができる。

[数24]

$$f_P(t) \sim N(0, \sigma_{P,t}^2), \quad f_Y(t) \sim N(0, \sigma_{Y,t}^2) \quad \dots (22)$$

[0057] さて、上述のモデルの時間発展について考える。以下、変数 $A(t)$ は、平均値 $A(t-1)$ 、分散 σ_A^2 の正規分布等与えられるものとする。つまり、変数 $A(t)$ は、以下の式（23）で表される。

[数25]

$$A(t) \sim N(A(t-1), \sigma_A^2) \quad \dots (23)$$

[0058] 変数 $A(t)$ は、時刻（ t ）よりも前の観測値 $\{y(1), \dots, y(t-1)\}$ により決定される。このときの $A(t)$ を $A_{t|t-1}$ と表記すると、式（23）は式（24）に書き換えることができる。

[数26]

$$A_{t|t-1} \sim N(A_{t-1|t-1}, \sigma_A^2) \quad \dots (24)$$

図2は、式（24）で与えられる確率分布を示すグラフである。図2に示す

ように、 $A_{t|t-1}$ は、 A_{min} から A_{max} の範囲で、平均値 $A_{t-1|t-1}$ 、標準偏差 σ_A の正規分布等となる。

[0059] しかし、式(24)及び図2に示す $A_{t|t-1}$ は、2つの問題点が存在する。1つは、例えば実際の金融市場などでは、市場価格の急峻な変動が存在するが、ランダムウォークに基づく上述のモデルでは、こうした急峻な変化に追従することができない。また、現実系では、パラメータがとりうる値の範囲が有限である。よって、本実施の形態では、変数 $A(t)$ を、切断正規分布等と一様分布との混合分布として、以下の式(25)で表すこととする。

[数27]

$$A(t) \sim (1 - A_{mut}) \text{truncN}(A(t-1), \sigma_A^2; A_{min}, A_{max}) + A_{mut} U(A_{min}, A_{max}) \quad \dots (25)$$

式(25)の $\{(1 - A_{mut}) \text{truncN}(A(t-1), \sigma_A^2; A_{min}, A_{max})\}$ は、平均値 $A(t-1)$ 、分散 σ_A^2 、区間 (A_{min}, A_{max}) の規格化された切断された正規分布等である。 truncN は、平均や標準偏差等で特徴づけられる一般の分布を上限と下限で切断したものに拡張することができる。また、 A_{mut} は、市場価格の急峻な変化などの突発現象に対応するために、式(25)で新たに導入した係数である。つまり、 $A(t)$ は、正規分布等と一様分布 $U(A_{min}, A_{max})$ の2つの成分が混合した形式で表される。 A_{mut} の値により、一様分布 $U(A_{min}, A_{max})$ の混合率が決定される。

[0060] 図3は、式(25)で与えられる確率分布を示すグラフである。図3に示すように、 $A_{t|t-1}$ は、 A_{min} から A_{max} の範囲で、平均値 $A_{t-1|t-1}$ 、標準偏差 σ_A の正規分布等と一定値 A_{mut} とが混合された分布として表現される。

[0061] 続いて、PUCKモデルの各パラメータに、式(25)を適用する具体例を説明する。PUCKモデルの各パラメータに式(25)を適用する処理については、上述のステップS1において行われる。図4は、実施の形態1における時刻 (t) でのポテンシャル係数 $b(t)$ の決定方法を示すフローチャートである。まず、時刻 (t) でのポテンシャル係数 $b(t)$ について説

明する。このとき、式（25）を式（26）に書き換えることができる。

[数28]

$$b_i(t) \sim (1 - b_{i,mut})N(b_i(t-1), \sigma_b^2) + b_{i,mut}U(b_{i,min}, b_{i,max}) \quad \dots (26)$$

[0062] 式（26）を用いて、粒子フィルタ法を適用することにより、時刻（t）での各次のポテンシャル係数 $b_i(t)$ を見積もることが可能である。ここで、時刻（t）での各次のポテンシャル係数 $b_i(t)$ における正規分布等と一様分布との混合比を決定する必要がある。そのために、ここでは、乱数 u を生成する（図4のステップS11）。乱数 u は、 $0 \leq u \leq 1$ を満たす任意の実数である。式（25）における b_{mut} は、時刻（t）での各次のポテンシャル係数 $b_i(t)$ の時間発展において、突発現象が起きる割合を示す係数である。本実施の形態では、 $b_{i,mut}$ と乱数 u を比較する（図4のステップS12）ことにより、正規分布等と一様分布との混合比を決定する。 $b_{i,mut}$ は、無作為に生成される乱数 u と比較されるので、結局のところ、混合比は $(1 - b_{i,mut}) : b_{i,mut}$ に収束することとなる。

[0063] すなわち、 $u \geq b_{i,mut}$ である場合には、時刻（t）でのポテンシャル係数 $b_i(t)$ は、正規分布等のみから成り立つものとする（図4のステップS13）。よって、この場合の時刻（t）での各次のポテンシャル係数 $b_i(t)$ は、以下の式（27）で表される。

[数29]

$$b_{i,l|l-1}^{(j)} \sim \text{trunc}N\left(b_{i,l-1|l-1}^{(j)}, \sigma_b^2, b_{i,min}, b_{i,max}\right) \quad \dots (27)$$

[0064] 一方、 $u < b_{i,mut}$ である場合には、時刻（t）での各次のポテンシャル係数 $b_i(t)$ は、一様分布のみから成り立つものとする（図4のステップS14）。よって、この場合の時刻（t）での各次のポテンシャル係数 $b_i(t)$ は、以下の式（28）で表される。

[数30]

$$b_{i,t|t-1}^{(j)} \sim U(b_{i,min}, b_{i,max}) \quad \dots (28)$$

[0065] 本実施の形態においては、上述のように、時系列データの急峻に変化する突発現象を考慮するために一様分布を導入している。換言すれば、一様分布が無い場合には、粒子の世代が進むごとに、適合度の低い粒子については、淘汰されて消滅してしまう。このような状態で突発現象が生じた場合には、突発状態に適合する粒子が既に消滅しており、突発現象に追従できない。ところが、本実施の形態におけるように、突発現象を考慮した一様分布を導入することにより、適合度の低い粒子の消滅を防ぐことができる。これにより、突発現象に対して適合度の高い粒子を残存させることが可能となり、通常の粒子フィルタ法と比べて、突発現象に対する追従能力を高めることが可能となる。このように、一様分布を導入することにより、通常の正規分布やローレンツ分布を用いることでは実現できない、特有の効果を実現することが可能となる。

[0066] 次いで、直近のデータ数 $M(t)$ について説明する。直近のデータ数 $M(t)$ についても、時刻 (t) でのポテンシャル係数 $b(t)$ と同様に、突発現象を考慮するための一様分布を導入する。本実施の形態では、直近のデータ数 $M(t)$ を、時刻 $(t-1)$ における確率密関数と一様分布との混合分布として表す。この直近のデータ数 $M(t)$ についての処理は、上述のステップ S1 において行われる。図5は、実施の形態1における時刻 (t) での直近のデータ数 $M(t)$ の決定方法を示すフローチャートである。この場合、直近のデータ数 $M(t)$ は、式(29)で表される。

[数31]

$$M(t) \sim (1 - M_{mut})G(M(t) | M(t-1), \sigma_M, M_{max}, M_{min}) + M_{mut}U(M_{min}, M_{max}) \quad \dots (29)$$

[0067] ここで、時刻 (t) でのポテンシャル係数 $b(t)$ の場合と同様に、乱数

u を生成し（図5のステップS21）、 M_{mut} と乱数 u を比較する（図5のステップS22）ことにより、時刻（ $t-1$ ）における確率密度関数と一様分布との混合比を決定する。

[0068] すなわち、 $u \geq M_{mut}$ である場合には、直近のデータ数 $M(t)$ は、時刻（ $t-1$ ）における確率密度関数のみから成り立つものとする（図5のステップS23）。よって、この場合の直近のデータ数 $M(t)$ は、以下の式（30）で表される。

[数32]

$$\begin{cases} M_{t|t-1}^{(j)}(t) \sim G(\cdot | M_{t-1|t-1}^{(j)}, r, M_{\max}, M_{\min}) \\ M_{t|t-1}^{(j)}(t) \sim U(M_{\min}, M_{\max}) \end{cases} \quad \dots (30)$$

ただし、ここでの、 $U(M_{\min}, M_{\max})$ は、下限を $M_{\min}$ 、上限を $M_{\max}$ とする整数値をとるような一様な確率密度関数である。

[0069] 一方、 $u < M_{mut}$ である場合には、直近のデータ数 $M(t)$ は、一様分布のみから成り立つものとする（図5のステップS24）。よって、この場合の直近のデータ数 $M(t)$ は、式（31）で表される。

[0070] 以下に想定される確率密度関数 $G(\cdot)$ の具体例を示す。 $M_{\min} < M(t-1) < M_{\max}$ の場合には、確率密度関数 $G(\cdot)$ は、以下の式（31）で表される。

[数33]

$$\begin{cases} G(M(t-1)+1 | M(t-1), r, M_{\min}, M_{\max}) = r \\ G(M(t-1) | M(t-1), r, M_{\min}, M_{\max}) = 1-2r \\ G(M(t-1)-1 | M(t-1), r, M_{\min}, M_{\max}) = r \end{cases} \quad \dots (31)$$

ただし、 $0 < r < 1$ とする。

また、境界条件においては、確率密度関数 $G(\cdot)$ は、以下の式（32）で表される。

[数34]

$$\begin{cases} G(M_{\max} - 1 | M_{\max}, r, M_{\min}, M_{\max}) = r \\ G(M_{\max} | M_{\max}, r, M_{\min}, M_{\max}) = 1 - r \\ G(M_{\min} + 1 | M_{\min}, r, M_{\min}, M_{\max}) = r \\ G(M_{\min} | M_{\min}, r, M_{\min}, M_{\max}) = 1 - r \end{cases} \quad \dots (32)$$

上記の範囲以外での $M(t-1)$ に関しては、確率密度関数 $G(\cdot)$ は、以下の式 (33) で表される。

[数35]

$$G(\cdot | M(t-1), r, M_{\min}, M_{\max}) = 0 \quad \dots (33)$$

[0071] 式 (22) の変動誤差 $f_p(t)$ の標準偏差 $\sigma_p(t)$ 、(観測誤差 $f_y(t)$ の標準偏差 $\sigma_y(t)$ も以下同様の手順)、は、次式で与えられる。

[数36]

$$\sigma_p(t) \sim (1 - \sigma_{pmut}) \text{trunc}N(\sigma_p(t-1), \gamma_p^2; \sigma_{pmin}, \sigma_{pmax}) + \sigma_{pmut} U(\sigma_{pmin}, \sigma_{pmax}) \quad \dots (34)$$

ここで、 σ_{pmut} は、0 と 1 の間の数値であり、 σ_{pmin} と σ_{pmax} はそれぞれ $\sigma_p(t)$ のとりうる範囲の下限と上限を表す。また、 γ_p^2 は分散を表す数値である。

[数37]

$$\sigma_y(t) \sim (1 - \sigma_{ymut}) \text{trunc}N(\sigma_y(t-1), \gamma_y^2; \sigma_{ymmin}, \sigma_{ymmax}) + \sigma_{ymut} U(\sigma_{ymmin}, \sigma_{ymmax}) \quad \dots (35)$$

ここで、 σ_{ymut} は、0 と 1 の間の数値であり、 σ_{ymmin} と σ_{ymmax} はそれぞれ $\sigma_y(t)$ のとりうる範囲の下限と上限を表す。また、 γ_y^2 は分散を表す数値である。

[0072] 以上、説明したように、本実施の形態にかかる時系列データの変化点検出方法によれば、時系列データが安定的なトレンドにある場合には、安定的なトレンドへの適合度が高いパラメータ (ポテンシャル係数 b 、直近のデータ

数 M 、標準偏差 σ_p 及び σ_v)を示す粒子を増加させ、トレンドに好適に追隨することが可能である。

[0073] さらに、本実施の形態にかかる時系列データの変化点検出方法によれば、安定的なトレンドから逸脱した、急峻な変動、すなわち突発現象が生じた場合でも、トレンドへの追隨が可能である。すなわち、通常の粒子フィルタ法であれば、安定的なトレンドが継続すると、安定的なトレンドに適合する粒子のみが増加し、安定的なトレンドに適合しないものの突発現象には適合する粒子は淘汰され、消滅してしまう。

[0074] ところが、本実施の形態にかかる時系列データの変化点検出方法では、最新の時刻におけるパラメータの確率密度関数を算出するにあたり、各パラメータに対して、式(28)及び式(30)に示すように、一定の割合で一樣分布を付与する。これにより、安定的なトレンドに対する適合度が悪いために消滅する粒子にも、図1のスツプS2で確率密度関数を生成する場合に、一定の割合で一樣分布が付与される。これにより、安定的なトレンドに適合しないものの突発現象には適合する粒子は、粒子フィルタ法による逐次計算が進んでも残存することができる。そして、ある時点で突発現象が発生した場合に、残存している粒子の内で突発現象への適合度の高い粒子を増加させることで、好適に突発現象のような動力学的現象に追隨することが可能である。

[0075] そして、ポテンシャル係数 $b_1(t)$ 及び直近のデータ数 $M(t)$ を逐次算出することにより、時系列データの變動の状態を検出することが可能である。2次のポテンシャル係数 $b_2(t)$ が正の値をとり、その他の $b_i(t)$ が0近傍であれば時系列の變動は安定的である。しかし、逐次解析の結果、2次のポテンシャル係数 $b_2(t)$ が負の値をとる粒子の適合度が増加する場合には、時系列の變動は不安定な状態となる。つまり、2次のポテンシャル係数 $b_2(t)$ の値の變動を監視することにより、時系列データの變動が安定的であるか、不安定であるか、の境界点を検出することが可能となる。また、3次のポテンシャル係数 $b_3(t)$ が正の値をとるときには時系列に急激な下

落トレンドが、 $b_3(t)$ が負の値をとるときには時系列に急激な上昇トレンドが、発生する可能性が予想される。

[0076] 加えて、粒子フィルタ法では、ポテンシャル係数 $b_i(t)$ 及び直近のデータ数 $M(t)$ について、複数の粒子として表現するが、時刻の進展に応じて、適合度が高い粒子は変化するので、適合度が高いポテンシャル係数 $b_i(t)$ 及び直近のデータ数 $M(t)$ の値も変化する。適合度の高い直近のデータ数は可変であるので、仮に時系列データの蓄積数が少ない場合でも、利用可能なデータを用いて、適合度の高い直近のデータ数 M を算出することが可能である。つまり、通常の PUCK モデルでは、ある程度のデータ数（1000 点程度）のデータ数が必要であるが、本方法では、より少ないデータ数でも利用可能である。さらに、既存手法では不可欠だった最適移動平均等のデータのノイズ除去処理の計算が不必要となった。従って、PUCK モデルのパラメータ推定の既存手法と比べて、データの蓄積が少ない段階でも PUCK モデルを適用することができるという利点を有する。

[0077] さらにまた、ポテンシャル係数 $b_i(t)$ の値と適合度、すなわち粒子数を参照することにより、時系列データの変動環境を解析することが可能である。具体的には、円／ドル取引を例とすると、2 次のポテンシャル係数 $b_2(t)$ が正の値をとる粒子についての適合度が高い傾向が有る場合には、価格の変化が安定であり、ランダムウォークと比較し、短い時間スケールでは逆方向に引き戻すように価格が変動しやすい。この場合には、市場では、価格の上昇あるいは下降のトレンドが反転しやすく、結果として、大きな価格変動は生じにくい。他方、2 次のポテンシャル係数 $b_2(t)$ が負の値をとる粒子についての適合度が高い傾向が有る場合には、価格の変化は不安定であり、ランダムウォークと比較し、短い時間スケールでは、同じ方向の価格変化が持続的に発生しやすく、上昇あるいは下降のトレンドは増幅されやすい。この場合には、トレンドを追従する戦略を持つトレーダーが市場において数的に支配的であると推測できる（非特許文献 2）。このように、本方法を、為替取引のような金融市場における価格などの時系列データに適用することに

より、市場におけるトレーダーの集団的な行動特性を定量的に評価することが可能となる。

[0078] 上述の市場におけるトレーダーの集団的な行動特性を定量的に評価できるという効果は、本実施の形態にかかる時系列データの変化点検出方法特有の効果である。本実施の形態にかかる時系列データの変化点検出方法では、複数の値の異なる時刻 (t) でのポテンシャル係数 $b_i(t)$ 及び直近のデータ数 M について、それぞれ適合度が異なる粒子として表現することができる。これにより、時刻 (t) において適合度が高く、時系列データの変動に支配的な影響を与えると考えられる、値の異なる時刻 (t) でのポテンシャル係数 $b_i(t)$ 及び直近のデータ数 $M(t)$ を、複数個同時に評価することができる。これにより、異なる行動特性を有するトレーダーの、行動の態様 (時刻 (t) でのポテンシャル係数 $b_i(t)$) や行動を決定するため参照すると考えられるデータ数 (直近のデータ数 $M(t)$) を、同時に評価することができる。これに対し、前述のような粒子フィルタ法を適用しない通常の P U C K モデルでは、直近のデータ数 M を仮定し、仮定した直近のデータ数 M に対する時刻 (t) でのポテンシャル係数 $b_i(t)$ を計算することしかできなかった。従って、原理的に、ポテンシャル係数 $b_i(t)$ 及び中心を見積もるために必要な直近のデータ数 $M(t)$ を同時に評価することはできなかった。つまり、時刻 (t) でのポテンシャル係数 $b_i(t)$ 及び直近のデータ数 $M(t)$ 等の P U C K モデルのパラメータの確率密度関数等を推定することはできなかった。つまり、P U C K モデルによる複数個の時刻 (t) でのポテンシャル係数 $b_i(t)$ 及び直近のデータ数 M の同時評価は、本実施の形態にかかる時系列データの変化点検出方法を用いることで初めて実現することが可能となる。

[0079] なお、本実施の形態にかかる時系列データの変化点検出方法は、上述のステップ S 1 ~ S 6 を表現するアルゴリズムが記述されたプログラムとして提供することが可能である。このプログラムをコンピュータなどのハードウェアに実行させることにより、上述の時系列データの変化点検出などの、本実

施の形態にかかる時系列データの変化点検出方法により得られるのと同様の効果をえることができる。例えば、当該プログラムを実行し、時刻 (t) のポテンシャル係数 $b_i(t)$ 等を表示装置に逐次表示し、ポテンシャル係数の経時的な変化をリアルタイムで観察することが可能である。

[0080] 実施の形態 2

次に、実施の形態 2 にかかる時系列データの変化点検出プログラムについて説明する。図 6 は、実施の形態 2 にかかる時系列データの変化点検出プログラムが実行される時系列データの変化点検出装置 200 の構成を模式的に示すブロック図である。時系列データの変化点検出装置 200 は、実施の形態 1 にかかる時系列データの変化点検出方法のステップ S1～S6 を表現するアルゴリズムが記述されたプログラムが実行される装置である。図 6 に示すように、時系列データの変化点検出装置 200 は、記憶部 1、演算部 2、表示部 3 及びバス 4 を有する。

[0081] 記憶部 1 は、ハードディスク、DRAM、SRAM、フラッシュメモリなどの記憶装置により構成され、過去の時系列データ、図 1 のステップ S1 で設定する初期値の情報、及び実施の形態 1 にかかる時系列データの変化点検出方法のステップ S1～S6 を表現するアルゴリズムが記述されたプログラム 200 が記憶されている。なお、記憶部に記憶されている過去の時系列データ、図 1 のステップ S1 で設定する初期値の情報及びプログラム 200 は、外部から適宜書き換えることが可能である。また、時系列データについては、時間の進展に応じて、新しいデータ要素を逐次追加することが可能である。

[0082] 演算部 2 は、バス 4 を介して、記憶部 1 から時系列データ、図 1 のステップ S1 で設定する初期値の情報、及びプログラム 200 を読み込む。演算部 2 は、読み込んだプログラム 200 を実行し、時系列データの変化点検出動作を行う。プログラム 200 は、初期設定部 21、PDF 生成部 22、適合度評価部 23、再サンプリング部 24 及びカウンタ部 25 を有する。

[0083] 具体的には、初期設定部 21 では、図 1 のステップ S1 に対応する処理を行う。すなわち、N 個の粒子の条件付き確率密度関数 (PDF) について、

まず、式（１０）で示される $t = 0$ における初期値を与える。そして、粒子フィルタ法による計算を開始するにあたり、時刻（ t ）を $t = 1$ に設定する。

[0084] PDF生成部２２では、図１のステップＳ２に対応する処理を行う。すなわち、 N 個の粒子について、式（６）を用いて、時刻（ $t - 1$ ）における確率密度関数（PDF）を基に、時刻（ t ）における予測粒子の確率密度関数（PDF）を生成する（式（１１））。

[0085] 適合度評価部２３では、図１のステップＳ３に対応する処理を行う。すなわち、各粒子の重み係数 W_i （式（１２））を、尤度 L_i （式（１３））を用いて計算する。なお、尤度 L_i は、式（７）より得られる。つまり、実際の時系列データから得られる観測ベクトル y_t と、計算により得られた状態ベクトル x_t と、の適合度を評価する。

[0086] 再サンプリング部２４では、図１のステップＳ４に対応する処理を行う。すなわち、各粒子の重み係数に従って、粒子の再サンプリングを行う。つまり、式（１１）で示される各粒子を、重み係数に応じた個数だけサンプリングし、合計 N 個の粒子をサンプリングする。この際、適合度が小さければ重み係数は小さくなるので、その粒子は残らない可能性が高くなる。一方、適合度が大きければ重み係数は大きくなり、その粒子は複製され増加することとなる。

[0087] 演算部２は、再サンプリング部２４によって得られた重み係数（すなわち粒子数）及び各粒子で表される時刻（ t ）でのポテンシャル係数 $b_i(t)$ や直近のデータ数 $M(t)$ の情報を、バス４を介して、表示部３に出力する。

[0088] カウント部２５では、図１のステップＳ５及びＳ６に対応する処理を行う。すなわち、最新の処理の時刻が最後のステップ（ $t = t_{end}$ ）であるかを判定する（図１のステップＳ５での処理）。そして、最新の処理の時刻が最後のステップ（ $t = t_{end}$ ）でなければ、時刻（ t ）を１ステップ進めて、初期設定部２１に帰還させる。これにより、初期設定部２１、PDF生成部２２、適合度評価部２３、再サンプリング部２４及びカウント部２５の間で、

ループ処理が実現される。一方、最新の処理の時刻が最後のステップ（ $t = t_{end}$ ）であれば処理を終了し、処理完了報告を、バス 4 を介して、表示部 3 に出力する。

[0089] 表示部 3 は、演算部 2 から出力される重み係数（すなわち粒子数）及び各粒子で表される時刻（ t ）でのポテンシャル係数 $b(t)$ や直近のデータ数 $M(t)$ の情報、処理完了報告を、視認可能な状態で、例えば液晶表示画面上に表示する。この際、時刻（ t ）でのポテンシャル係数 $b(t)$ 等を表示装置に逐次表示し、ポテンシャル係数の経時的な変化をリアルタイムで観察することが可能である。

[0090] 実施の形態 3

次に、実施の形態 3 にかかる未来の時系列データがとる値の確率密度分布の予測方法について説明する。実施の形態 1 にかかる時系列データの変化点検出方法で説明したように、PUCK モデル及び粒子フィルタ法を用いることにより、時刻（ t ）におけるポテンシャル係数を求めることができる。さらに、この PUCK モデルを時間発展させることで、未来の市場価格がとる値を確率密度分布として得ることが可能である。

[0091] j 番目の粒子の時刻（ $t+1$ ）における市場価格（ $t+1$ ）は、式（1）を用いて、以下の式（36）で表すことができる。

[数38]

$$P^{(j)}(t+1) = P^{(j)}(t) - \sum_{i=2}^K \left\{ b^{(j)}_i(t) \left(\frac{P^{(j)}(t) - P_M^{(j)}(t)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_P(t)$$

・・・(36)

[0092] 図 7 は、実施の形態 3 にかかる粒子フィルタ法を適用した PUCK モデルの時間発展方法を示すフローチャートである。まず、実施の形態 1 にかかる時系列データの変化点検出方法で計算された時刻（ t ）での粒子のパラメータ（ $b^{(j)}_i(t)$ 、 $M^{(j)}(t)$ 、 $f^{(j)}_P(t)$ ）を準備するステップ S 3 1）。

- [0093] 次いで、実施の形態 1 にかかる時系列データの変化点検出方法で計算された時刻 (t) での真の市場価格 P (t) を準備する (ステップ S 3 2)。
- [0094] 準備したパラメータ及び時刻 (t) での真の市場価格 P^(j) (t) を用いて、式 (3 6) を計算する。
- [0095] そして、最新の処理の時刻 (t) が、t = t + N に到達したかを判断する (ステップ S 3 4)。最新の処理の時刻が最後のステップ (t = t + N) でなければ、時刻 (t) を 1 ステップ進めて (ステップ S 3 5)、式 (3 6) での計算を再度実行する。最新の処理の時刻が最後のステップ (t = t + N) であれば、処理を終了する。
- [0096] これにより、時刻 (t) でのパラメータ及び真の市場価格を用いて、式 (3 6) を時刻 (t + N) まで時間発展させることができる。粒子フィルタ法を用いて推定した時刻 (t) での j 番目の粒子の P U C K モデルのパラメータセット (ポテンシャル係数 b^(j)_i (t)、価格変動の中心を見積もるために必要なデータ数 M^(j) (t)、変動誤差 f^(j)_p (t)) とすると、式 (3 6) を時刻 (t + N) まで時間発展させた場合の時刻 (t + N) での予想市場価格は、以下の式 (3 7) で表される。

[数 39]

$$P^{(j)}(t+N) = P^{(j)}(t) - \sum_{n=0}^{N-1} \left\{ \sum_{i=2}^K b^{(j)}_i(t) \left(\frac{P^{(j)}(t+n) - P_M^{(j)}(t+n)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_p(t) \Bigg\} \dots (37)$$

- [0097] 以上、各粒子について式 (3 7) を適用することにより、時刻 (t + N) における市場価格 P^(j) (t + N) の確率密度分布を得ることが可能となる。図 8 は、実施の形態 3 にかかる未来の時系列データがとる値の確率密度分布の予測方法により時刻 (t + N) における市場価格の確率密度分布を求めた場合の確率密度分布の推移を示すグラフである。なお、図 8 では、N = 1 0 0 として、米ドル／日本円の為替レートを例として用いた。図 8 では、式 (3 7) で K = 3 の場合 (すなわち、2 次及び 3 次のポテンシ

ャルが作用する) 場合の時刻 (t) から時刻 ($t + N$) までの確率密度分布について、第1四分位偏差 $QD1$ 、中心値 Vc 及び第3四分位偏差 $QD3$ を表示した。なお、図8では、 $N = 100$ として計算した。

[0098] 図9は、実施の形態3にかかる未来の時系列データがとる値の確率密度分布の予測方法により時刻 ($t + N$) における市場価格がとる値の確率密度分布を示すグラフである。なお、図9では、図8と同様に、 $N = 100$ として、米ドル/日本円の為替レートを例として用いた。図9に示すように、時刻 ($t + N$) での確率密度分布は、価格上昇側 (分布曲線の頂部よりも右側の部分) での裾が広く、歪んだ分布を有することが確認できる。つまり、図9の例では、時刻 ($t + N$) での市場価格は、上昇傾向の方がより強いと理解することができる。

[0099] 以上のように、本実施の形態にかかる未来の時系列データがとる値の確率密度分布の予測方法によれば、3次以上のポテンシャル導入することで、価格変動の非線形挙動の影響を反映させることが可能となる。これにより、未来の市場価格の予測において、図9に示すような、左右非対称の歪んだ確率密度分布を生成することができる。これは、一般的なりスク評価などで用いる正規分布などの左右対称分布を用いる手法では実現できない、本実施の形態によりはじめて実現できる効果である。これにより、現実の金融市場等において生起する左右非対称の歪んだ確率密度分布を好適に予測することができ、より現実に即した時系列データがとる値の確率密度分布の予測方法を提供することができる。

[0100] 以上、本実施の形態にかかる未来の時系列データがとる値の確率密度分布の予測方法によれば、粒子フィルタ法を用いて推定した時刻 (t) でのPUCモデルのパラメータセットを基礎として、未来の市場価格がとる値の確率密度分布を得ることができる。また、式(37)においては、 $K \geq 3$ の場合には3次以上のポテンシャルを導入できるため、市場価格変動の非線形な挙動を含めた予測が可能となる。これにより、より市場の実態に即した市場価格の予測が可能となる。

[0101] 本実施の形態にかかる未来の時系列データがとる値の確率密度分布の予測方法を為替や株式などの価格予測に適用すれば、為替や株式の価格変動のリスクを従来手法よりも高精度に見積もることができる。よって、従来手法と比べて、より適正にリスクを見積もったオプションなどの金融商品の設計・開発に活用することができる。

[0102] 実施の形態 4

次に、実施の形態 4 にかかる未来の時系列データがとる値の確率密度分布の予測プログラムについて説明する。図 10 は、実施の形態 4 にかかる未来の時系列データがとる値の確率密度分布の予測プログラムが実行される未来の時系列データがとる値の確率密度分布の予測装置 400 の構成を模式的に示すブロック図である。未来の時系列データがとる値の確率密度分布の予測装置 400 は、実施の形態 3 にかかる未来の時系列データがとる値の確率密度分布の予測方法のステップ S 31 ~ S 35 を表現するアルゴリズムが記述されたプログラムが実行される装置である。図 10 に示すように、未来の時系列データがとる値の確率密度分布の予測装置 400 は、記憶部 1、演算部 2a、表示部 3 及びバス 4 を有する。記憶部 1、表示部 3 及びバス 4 は、実施の形態 2 と同様であるので、説明を省略する。

[0103] 演算部 2a は、バス 4 を介して、記憶部 1 から、プログラム 20 により算出されたパラメータ及び時系列データ及びプログラム 40 を読み込む。演算部 2 は、読み込んだプログラム 40 を実行し、未来の時系列データがとる値の確率密度分布の予測を行う。プログラム 40 は、パラメータ読み込み部 41、データ読み込み部 42、算出部 43 及びカウント部 44 を有する。

[0104] 具体的には、パラメータ読み込み部 41 は、プログラム 20 で計算された時刻 (t) での j 番目の粒子のパラメータ ($b^{(j)}_i(t)$ 、 $M^{(j)}(t)$ 、 $f^{(j)}_p(t)$) を読み込む。つまり、パラメータ読み込み部 41 は、図 7 のステップ S 31 に対応する処理を行う。なお、プログラム 20 で計算された時刻 (t) での粒子のパラメータは、例えば記憶部 1 に格納され、必要に応じてパラメータ読み込み部 41 により読み出される。

[0105] データ読み込み部42は、プログラム20で計算された時刻(t)での真の市場価格 $P^{(j)}(t)$ を読み込む。つまり、データ読み込み部42は、図7のステップS32に対応する処理を行う。なお、プログラム20で計算された時刻(t)での真の市場価格 $P^{(j)}(t)$ は、例えば記憶部1に格納され、必要に応じてデータ読み込み部42により読み出される。

[0106] 算出部43は、図7のステップS33に対応する処理を行う。すなわち、式(36)を用いて、設定されている時刻よりも1ティック先の市場価格を算出する。

[0107] カウント部44は、図7のステップS34及びS35に対応する処理を行う。すなわち、最新の処理の時刻が最後のステップ($t = t + N$)であるかを判定する(図7のステップS34での処理)。そして、最新の処理の時刻が最後のステップ($t = t + N$)でなければ、時刻(t)を1ステップ進めて、算出部43に帰還させる。これにより、式(36)の計算が繰り返され、時刻($t + N$)に到達した場合には、式(37)に示す式と同等の計算件が得られる。最新の処理の時刻が最後のステップ($t = t + N$)であれば、処理を終了し、算出結果に基づいて、時刻($t + N$)における市場価格の確率密度分布を、表示部3に出力する。

[0108] 表示部3は、演算部2から出力される時刻($t + N$)における市場価格の確率密度分布を、視認可能な状態で、例えば液晶表示画面上に表示する。この際、時刻(t)から時刻($t + N$)までの間の市場価格の確率密度関数を表示装置に逐次表示することも可能である。

[0109] 以上より、本実施の形態によれば、実施の形態4にかかる未来の時系列データがとる値の確率密度分布の予測プログラム及び予測装置を具体的に実現することができる。

[0110] 実施の形態5

次に、実施の形態5にかかる未来の時系列データがとる値の確率密度分布の他の予測方法について説明する。実施の形態3では、PUCKモデルを時間発展させることで、未来の市場価格がとる値を確率密度分布として得るこ

とが可能であることを説明した。つまり、時間発展の開始時点である時刻（ t ）でのポテンシャル係数 $b(t)$ を用いて時間発展を行う。なお、本実施の形態では、ポテンシャルの次数 i 、粒子の番号 j については、表示を省略している。

[0111] 実施の形態3では、時刻（ t ）でのポテンシャル係数 $b(t)$ の値が、将来の時刻（ $t+1$ ）から（ $t+N$ ）の間でも同様だと仮定して、将来を予測している。従って、実施の形態3では、短期的な価格変動の時間変化しか予測することができない。これに対し、本実施の形態のように、ポテンシャル係数 $b(t)$ の時間発展の方程式が得られれば、将来の時刻（ $t+1$ ）から（ $t+N$ ）の間のポテンシャル係数 $b(t)$ の変化を予測することができる。これにより、より正確な将来の時刻 $t+N$ での価格分布が得られる。従って、長期的な価格変動の時間変化を予測する場合、ポテンシャル係数 $b(t)$ の時間発展の方程式を推定することは重要である。例えば、時刻（ $t+1$ ）でのポテンシャル係数 $b(t+1)$ と時刻（ t ）でのポテンシャル係数 $b(t)$ との関係を、以下の式（38）で定義する。

[数40]

$$b(t+1) - b(t) = -\frac{\partial}{\partial b} G(b) \Big|_{b=b(t)} + f_b(t) \quad \dots (38)$$

なお、右辺第1項（偏微分項）の関数 $G(b)$ は、時刻（ t ）までのポテンシャル係数の時間発展を記述する関数である。 $f_b(t)$ は、ノイズ項である。

[0112] マイナス符号を除く右辺第1項（偏微分項）を、ポテンシャル係数の時間変化に作用するポテンシャル λ とすると、式（38）は式（39）に変形できる。

[数41]

$$b(t+1) - b(t) = -\lambda b(t) + f_b(t) \quad \dots (39)$$

[0113] 式(39)より、時刻(t+1)でのポテンシャル係数 $b(t+1)$ は、以下の式(40)で表される。

[数42]

$$b(t+1) = (1-\lambda)b(t) + f_b(t) \quad \dots (40)$$

式(40)より、時刻(t+1)でのポテンシャル係数 $b(t+1)$ は、係数(1- λ)が作用し、ノイズ項 $f_b(t)$ が加えられる。

[0114] ここで、式(40)で示すポテンシャル係数を用いて、例えば2次のポテンシャルが作用する場合の市場価格分布の予測について説明する。2次のポテンシャルが作用する場合の1ティックの間での真の市場価格の変動は、式(2)、(4)及び(5)より、以下の式(41)で表すことができる。

[数43]

$$P(t+1) - P(t) = -\left(\frac{b(t)}{(M-1)}(P(t) - P_M(t))\right) + f_P(t) \quad \dots (41)$$

[0115] 図11は、実施の形態5にかかる粒子フィルタ法を適用したPUCKモデルの時間発展方法を示すフローチャートである。図11では、図7と比べて、ステップS33をステップS52に置換している。また、ステップS32及びS35とステップS52との間に、ステップS51が追加されている。

[0116] ステップS51は、式(40)を用いて、ポテンシャル係数 $b(t+1)$ を計算する。次いで、ステップS52では、準備したパラメータ及び時刻(t)での真の市場価格 $P^{(j)}(t)$ を用いて、式(41)を計算する。その他のステップについては、図7と同様であるので、説明を省略する。

[0117] 図12は、現在の市場価格を数理モデルに代入して、粒子番号j毎に時間発展させ、その経過を粒子毎に表示したグラフである。ここでは、例として、開始価格を100とし、1450ティック後までの市場価格の分布を予測した。図12に示すように、本実施の形態によれば、ポテンシャル係数を時間発展させることで、将来のポテンシャル係数の変動をも予測し、市場価格

の長期トレンドをより正確に予測することが可能である。

- [0118] 長期トレンドの予測は市場価格分布として与えられるので、ある将来の時点における価格変動についてのリスクの見積もりが可能である。これにより、本実施の形態によれば、将来の為替や株式の価格変動のリスクを実施の形態3よりも高精度に見積もることができる。よって、更に適正にリスクを見積もったオプションなどの金融商品の設計・開発に活用することができる。
- また、多数の金融機関が自己資産のリスクを計測するために用いているValue at Riskなどのリスク管理の指標を改善することができる。

- [0119] なお、本実施の形態で説明したポテンシャル係数の時間発展は例示に過ぎない。よって、ポテンシャル係数の時間発展は、式(40)で示す例に限られない。

[0120] 実施の形態6

次に、実施の形態6にかかる未来の時系列データがとる値の確率密度分布の予測プログラムについて説明する。図13は、実施の形態6にかかる未来の時系列データがとる値の確率密度分布の予測プログラムが実行される未来の時系列データがとる値の確率密度分布の予測装置600の構成を模式的に示すブロック図である。未来の時系列データがとる値の確率密度分布の予測装置600は、実施の形態5にかかる未来の時系列データがとる値の確率密度分布の予測方法のステップS31、32、S51、S52、S34、S35を表現するアルゴリズムが記述されたプログラムが実行される装置である。未来の時系列データがとる値の確率密度分布の予測装置600は、未来の時系列データがとる値の確率密度分布の予測装置400の変形例である。図13に示すように、未来の時系列データがとる値の確率密度分布の予測装置600は、図10のプログラム40をプログラム60に置換した構成を有する。

- [0121] プログラム60は、プログラム40の算出部43を算出部62に置換し、更にポテンシャル算出部61を追加した構成を有する。ポテンシャル算出部61は、図11のステップS51に対応する処理を行う。すなわち、式(4

0) を用いて、ポテンシャル係数を計算する。算出部 62 は、図 10 のステップ S52 に対応する処理を行う。すなわち、式 (41) を用いて、設定されている時刻よりも 1 ティック先の市場価格を算出する。未来の時系列データがとる値の確率密度分布の予測装置 600 のその他の構成は、未来の時系列データがとる値の確率密度分布の予測装置 400 と同様であるので、説明を省略する。

[0122] 以上より、本実施の形態によれば、実施の形態 5 にかかる未来の時系列データがとる値の確率密度分布の予測プログラム及び予測装置を具体的に実現することができる。

[0123] なお、本発明は上記実施の形態に限られたものではなく、趣旨を逸脱しない範囲で適宜変更することが可能である。例えば、半導体製造工程において装置から生じるゆらぎを伴う時系列データに対して本発明を適用することによって、工程の処理が安定的に進んでいるかどうかをリアルタイムで計測し、何らかの原因で不安定化した場合には速やかに異常を検出することが可能となる。この場合、例えば、市場価格を装置が出力する管理・計測データに置き換えることで、同様の解析を行うことが可能である。

[0124] 上述の実施の形態では、式 (4) に示すように、時刻 (t) でのポテンシャル係数 $b(t)$ を 2 次以上のポテンシャル $W(q, t)$ の係数として説明した。例えば、現実の金融市場などでは、2 次ポテンシャルでは表現できない、更なる急激なトレンドの変動が有ることが知られている。そのため、時刻 (t) でのポテンシャル $U(q, t)$ を 2 次に限定せず、高次ポテンシャル項を含む関数形を導入して高次ポテンシャル項の係数を推定することにより、急激なトレンド変動に追随することができる。特に、3 次ポテンシャルを導入することにより、金融市場における暴騰又は暴落について、2 次ポテンシャルと比べてより高速に追随することが可能である。また、暴騰（すなわち、トレンド上昇）又は暴落（すなわち、トレンド下降）の方向性についても、併せて解析することが可能となる。

[0125] また、2 次ポテンシャルモデルを用いて、時刻 (t) における $b(t)$ お

よびM (t) の同時確率分布等を推定することにより、複数の時間スケールでのダイナミクスの変化の異なる戦略が推定可能となる (マルチスケールPUCモジュール)。さらに、3 以上の高次ポテンシャル係数を推定することで、トレンドを追うダイナミクスの変化の異なる戦略、および、価格を一方的に変動させている集団的行動の可視化が可能となる (高次ポテンシャルを含むマルチスケールPUCモジュールの展望)。

[0126] この出願は、2011 年9 月8 日に提出された日本出願特願2011-196512 及び2012 年3 月7 日に提出された日本出願特願2012-50819 を基礎とする優先権を主張し、その開示の全てをここに取り込む。

産業上の利用可能性

[0127] 本発明は、為替市場や株式市場での市場価格の時間変動の分析又は予測、管理・計測データなどの時系列データの時間変動の分析又は予測に適用することができる。

符号の説明

- [0128] 1 記憶部
- 2 演算部
- 3 表示部
- 4 バス
- 20、40、60 プログラム
- 21 初期設定部
- 22 PDP 生成部
- 23 適合度評価部
- 24 再サンプリング部
- 25 カウント部
- 41 パラメータ読み込み部
- 42 データ読み込み部
- 43、62 算出部
- 44 カウント部

6 1 ポリシヤル算出部

2 0 0 時系列データの变化点検出装置

4 0 0、6 0 0 未来の時系列データが与る値の確率密度分布の予測装置
S 1 ~ S 6、S 1 1 ~ S 1 4、S 2 1 ~ S 2 4、S 4 1 ~ S 4 5、S 5 1、
S 5 2 ストップ

請求の範囲

[請求項1] 時刻 ($t - 1$) での真の市場価格及び時刻 ($t - 1$) までの M (M は、正の整数) 個の真の市場価格の移動平均である中心価格によって定義されるポテンシャル項と変動誤差との和によって、時刻 (t) での真の市場価格 $P(t)$ を算出する P U C K モデルに、粒子フィルタ法を適用する時系列データの変化点検出方法であって、

それぞれ異なる値を有する P U C K モデルの状態を表すパラメータを持つ粒子群より次の時刻のパラメータの確率密度関数を得る第 1 のステップと、

前記複数の粒子のそれぞれについて、時刻 (t) で観測した市場価格に対する前記時刻 (t) での真の市場価格の適合度を評価する第 2 のステップと、

前記適合度に応じて、前記複数の粒子の中から粒子を再サンプリングする第 3 のステップと、を備え、

前記第 3 のステップでは、

乱数を生成し、

第 1 の所定の値と前記乱数とを比較し、

前記乱数が前記第 1 の所定の値よりも大きい場合には、時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする確率密度分布を成分とする確率密度関数を粒子として生成し、

前記乱数が前記第 1 の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成する、

時系列データの変化点検出方法。

[請求項2] 時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする確率密度分布は、時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする正規分布である、

請求項 1 に記載の時系列データの変化点検出方法。

[請求項3] 前記第 1 のステップでは、P U C K モデルの状態を表すパラメータ

を持つ粒子群より次の時刻のPUCKモデルのパラメータが更新される、

請求項1又は2に記載の時系列データの変化点検出方法。

[請求項4]

前記第3のステップでは、

それぞれの粒子に関して乱数を生成し、

前記第1の所定の値と前記乱数とを比較し、

前記乱数が前記第1の所定の値よりも大きい場合には、時刻(t)におけるMを前提とする条件付き確率密度関数を粒子として生成し、

前記乱数が前記第1の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成する、

請求項1乃至3のいずれか一項に記載の時系列データの変化点検出方法。

[請求項5]

時刻($t+1$)での真の市場価格は、前記ポテンシャル項と、時刻(t)での真の市場価格の変動誤差と、を加算することにより算出されることを特徴とする、

請求項1乃至4のいずれか一項に記載の時系列データの変化点検出方法。

[請求項6]

前記時刻(t)で観測した市場価格は、前記時刻(t)での真の市場価格と時刻(t)での市場価格の観測誤差との和であることを特徴とする、

請求項1乃至5のいずれか一項に記載の時系列データの変化点検出方法。

[請求項7]

前記時刻(t)での真の市場価格の変動誤差及び前記時刻(t)での市場価格の観測誤差は、正規分布に従う確率密度関数で与えられることを特徴とする、

請求項5又は6に記載の時系列データの変化点検出方法。

[請求項8]

請求項1乃至7に記載した時系列データの変化点検出方法により算出した時刻(t)における前記ポテンシャル項のポテンシャル係数、

前記Mの値、及び時刻（t）での真の市場価格の変動誤差と、に基づいて、前記時刻（t）での真の市場価格P（t）を時間発展させることにより、時刻（t + N）での前記時刻（t）での真の市場価格P（t）を算出する、

未来の時系列データ値の確率密度分布の予測方法。

[請求項9]

前記粒子の総数をN_p、

前記粒子を特定する番号をj（ただし、jは1 ≤ j ≤ N_pを満たす整数）、

番号jの粒子の時刻（t）でのポテンシャル係数をb^(j)_i（t）

（ただし、iはポテンシャルの次数を示す2以上の整数）、

前記番号jの粒子の時刻（t）での真の市場価格の変動誤差をf^(j)_p（t）として、

前記番号jの粒子の時刻（t）での真の市場価格をP^(j)（t）、

前記番号jの粒子の時刻（t + N）での真の市場価格をP^(j)（t + N）、

前記番号jの粒子の時刻（t）での市場価格の中心価格をP^(j)_M（t）、

時刻（t）におけるMの値をM（t）とした場合に、

前記番号jの粒子の前記時刻（t + N）での真の市場価格P^(j)（t + N）は、Kを2以上の整数として、

[数6]

$$P^{(j)}(t+N) = P^{(j)}(t) - \sum_{n=0}^{N-1} \left\{ \sum_{i=2}^K \left\{ b^{(j)}_i(t) \left(\frac{P^{(j)}(t+n) - P^{(j)}_M(t+n)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_P(t) \right\}$$

で表される、

請求項8に記載の未来の時系列データ値の確率密度分布の予測方法。

[請求項10]

番号jの粒子の時刻（t + 1）でのポテンシャル係数b^(j)_i（t

+ 1) は、前記番号 j の粒子の時刻 (t) でのポテンシャル係数 $b(j)(t)$ を時間発展させることにより算出される、

請求項 9 に記載の未来の時系列データ値の確率密度分布の予測方法。

[請求項 11]

時刻 ($t - 1$) での真の市場価格及び時刻 ($t - 1$) までの M (M は、正の整数) 個の真の市場価格の移動平均である中心価格によって定義されるポテンシャル項と変動誤差との和によって時刻 (t) での真の市場価格 $P(t)$ を算出する P U C K モデルに粒子フィルタ法を適用する時系列データの変化点検出をコンピュータに実行させる時系列データの変化点検出プログラムであって、

それぞれ異なる値を有する P U C K モデルの状態を表すパラメータを持つ粒子群を生成することによりパラメータの確率密度関数を得る第 1 の処理と、

適合度評価部に、前記複数の粒子のそれぞれについて、時刻 (t) で観測した市場価格に対する前記時刻 (t) での真の市場価格の適合度を評価する第 2 の処理と、

再サンプリング部に、前記適合度に応じて、前記複数の粒子の中から粒子を再サンプリングさせる第 3 の処理と、をコンピュータに実行させ、

前記乱数が前記第 1 の所定の値よりも大きい場合には、時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする確率密度分布を成分とする確率密度関数を粒子として再生成する処理と、

前記乱数が前記第 1 の所定の値よりも小さい場合には、一様分布にしたがって粒子を生成する処理と、を実行させる、

時系列データの変化点検出プログラム。

[請求項 12]

時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする確率密度分布は、時刻 (t) での P U C K モデルの状態を表すパラメータを平均値とする正規分布である、

請求項 1 1 に記載の時系列データの変化点検出プログラム。

[請求項13] 前記第 3 のステップでは、前記適合度が大きい粒子ほど、多くサンプリングされるだけでなく、一定の割合で、実現可能性のある適合度が低い粒子を生成することを特徴とする、

請求項 1 1 又は 1 2 に記載の時系列データの変化点検出プログラム。

[請求項14] 前記第 1 のステップでは、

乱数を生成する処理と、

時刻 (t) における P U C K モデルのパラメータを前提とする条件付き確率密度関数を粒子として生成する処理と、

前記乱数が前記第 1 の所定の値よりも小さい場合には、一様分布を成分とする確率密度関数を粒子として生成する処理と、を実行させることを特徴とする、

請求項 1 1 乃至 1 3 のいずれか一項に記載の時系列データの変化点検出プログラム。

[請求項15] 前記時刻 (t) での真の市場価格は、前記ポテンシャル項と、時刻 (t) での真の市場価格の変動誤差と、を加算することにより算出されることを特徴とする、

請求項 1 1 乃至 1 4 のいずれか一項に記載の時系列データの変化点検出プログラム。

[請求項16] 前記時刻 (t) で観測した市場価格は、前記時刻 (t) での真の市場価格と時刻 (t) での市場価格の観測誤差との和であることを特徴とする、

請求項 1 1 乃至 1 5 のいずれか一項に記載の時系列データの変化点検出方法。

[請求項17] 前記時刻 (t) での真の市場価格の変動誤差及び前記時刻 (t) での市場価格の観測誤差は、正規分布に従う確率密度関数で与えられることを特徴とする、

請求項 15 又は 16 に記載の時系列データの変化点検出プログラム

。

[請求項18]

請求項 11 乃至 17 に記載した時系列データの変化点検出プログラムにより算出した時刻 (t) における前記ポテンシャル項のポテンシャル係数、前記 M の値、及び時刻 (t) での真の市場価格の変動誤差と、に基づいて、前記時刻 (t) での真の市場価格 P (t) を時間発展させることにより、時刻 (t + N) での前記時刻 (t) での真の市場価格 P (t) を算出することを特徴とする、

未来の時系列データ値の確率密度分布の予測プログラム。

[請求項19]

前記粒子の総数を N_p、

前記粒子を特定する番号を j (ただし、j は $1 \leq j \leq N_p$ を満たす整数)、

番号 j の粒子の時刻 (t) でのポテンシャル係数を $b^{(j)}_i(t)$

(ただし、i はポテンシャルの次数を示す 2 以上の整数)、

前記番号 j の時刻 (t) での真の市場価格の変動誤差を $f^{(j)}_P(t)$ として、

前記番号 j の粒子の時刻 (t) での真の市場価格を $P^{(j)}(t)$ 、

前記番号 j の粒子の時刻 (t + N) での真の市場価格を $P^{(j)}(t + N)$ 、

前記番号 j の粒子の時刻 (t) での市場価格の中心価格を $P^{(j)}_M(t)$ (t)、

時刻 (t) における M の値を M (t) とした場合に、

前記番号 j の粒子の前記時刻 (t + N) での真の市場価格 $P^{(j)}(t + N)$ は、K を 2 以上の整数として、

[数7]

$$P^{(j)}(t + N) = P^{(j)}(t) - \sum_{n=0}^{N-1} \left\{ \sum_{i=2}^K b^{(j)}_i(t) \left(\frac{P^{(j)}(t + n) - P^{(j)}_M(t + n)}{M^{(j)}(t) - 1} \right)^{i-1} \right\} + f^{(j)}_P(t)$$

で表される、

請求項 18 に記載の未来の時系列データ値の確率密度分布の予測プログラム。

[請求項20] 番号 j の粒子の時刻 $(t + 1)$ でのポテンシャル係数 $b^{(j)}_i(t + 1)$ は、前記番号 j の粒子の時刻 (t) でのポテンシャル係数 $b^{(j)}_i(t)$ を時間発展させることにより算出される、

請求項 19 に記載の未来の時系列データ値の確率密度分布の予測プログラム。