

最尤法による隠れマルコフモデルの推定 — t 分布のケース —

比 佐 章¹・比 佐 優 子²

要旨

本論では、Lystig and Hughes (2012) をもとに、状態依存型の変数が t 分布に従う隠れマルコフモデル (Hidden Markov Model: 以下 HMM) に、最尤推定法を用いたケースについて論じる。Lystig and Hughes (2012) は、HMM の対数尤度関数の一階条件について論じており、これにより準ニュートン法を応用させることが可能となった。また二階条件についても論じており、情報行列の導出が容易になった。そこで本論では、これまで困難であった、分散・共分散行列に基づく仮説検定の推定と、それに基づく可能性についても論じた。

HMM に合致したデータを人工的に作成し、それに対して最尤法を適用すると、(1) 最尤法による推定は、自由度のパラメータを除くと、比較的少ないデータ数でも精度の高い推定量を求められる、(2) 情報行列や分散・共分散行列の推定には、より多くのデータ数が必要とされる、(3) 自由度のパラメータの推定量については、初期値をきちんと設定しないと、真のパラメータに収斂しにくいことなどが判明した。

またこのモデルは、新規公開市場のアンダープライシングのような、循環的な要因と、個別企業の直面するガバナンス要因という、異なる要因が同時に影響するような変数の分析を行うのに、効果的な手法であることを明らかにした。

¹ 横浜市立大学 国際総合科学部 経営科学コース 特任准教授

² 帝京大学 経済学部 講師

1. はじめに

隠れマルコフモデル（以下、HMM）とは、マルコフ過程（Markov Process）に従って変化する状態変数に依存した観測値を、確率的にとらえるモデルであり、分野を問わず広く応用されてきた。たとえば、自然科学の分野では、音声認識や遺伝子解析などの分析に、また医学の分野では、人々の病状の変化や、症状の進行の速さや、死亡するまでの平均余命などの分析などに用いられている。

またHMMは、レジーム・スイッチング＝モデルともいわれ、時系列に変化するデータの分析への応用がしやすいという特徴があり、経済学においては、景気循環や株価の収益率の変化などの分析に使われている。たとえば経済状態が好景気か不景気かという二つの状態が考えられるときに、どちらの状態なのかという判断をしたり、株価の変化が市場の状況（例えば通常の状態と金融危機の場合というように）で異なってくる場合、その違いを考慮して分析を行えるという特徴がある。

HMMによる分析では、尤度関数を定義した上で、マルコフ過程に従う推移確率と観測値の確率パラメータについて推定を行う。しかしこのモデルでは、状態の変化するパターンの組み合わせが、時系列の経過とともに増大していく。そのため通常の最尤法では、一階条件の定義が難しいという問題点があった。そこで各変数の状態確率の期待の推定手順（Eステップ）と、パラメータの最尤推定値を求める手順（Mステップ）を繰り返すことで、パラメータを推定するEMアルゴリズムや、ベイズ推定にもとづき乱数を発生させることでパラメータの信用区間を求めるMarkov Chain Monte Carlo Simulation法（以下、MCMC法）などによる分析が行われることが多かった。

しかしながらEMアルゴリズムでは、準ニュートン法などの手法を応用させることができないことから、収束計算の効率性の点で問題があると考えられる。またMCMC法は、計算手法が簡便になるというメリットはあるものの、基本的に乱数発生による推定のため、通常の最尤法で

用いられる情報行列 (Information Matrix) にもとづく Wald 検定が適用できないといった制約が存在することや、また十分な数の乱数を発生させると計算時間が必然的に長くなるといった問題点がある。

Lystig and Hughes (2012) は、こうした点を考慮し、Baum のフォワード・アルゴリズムにもとづく尤度関数の一階微分および二階微分の計算方法を提唱した。彼らは各時点で、初期時点からその直近までの観測値から求められる条件付確率を計算し、それを逐次計算していくことで尤度関数および一階・二階微分を求めることができることを明らかにした。そしてこの方法により、HMM の尤度関数の計算がより容易になるとしている。そこで本論では、彼らの論文をもとに、HMM における準ニュートン法の適用について議論する。

また Lystig and Hughes (2012) による分析方法では、尤度関数の二階条件から情報行列を導出することができる点に特徴がある。HMM において、尤度関数の二階条件を求めることは、一階条件を求める以上に難しいため、EM アルゴリズムの収束計算のプロセスで求められた変化の度合いから、近似的に求める手法が用いられてきた³。Lystig and Hughes (2012) による手法は、尤度関数の二階条件に関する計算がより精緻になるというメリットがあることから、本論文では二階条件による情報行列および分散・共分散行列の推定が、どの程度有意義かについて検証する。

また本論文では、各時点の確率変数が t 分布に従う場合について、検証を行い、t 分布の HMM の推定結果について論じる。近年、株価をはじめとする市場においてその価格の変動率が、正規分布よりもより裾野の厚い分布に従うことが指摘されている。経済物理の分野では、株価がべき乗分布に従う可能性があることが指摘されている。ただこれらの分析では、状態が複数存在するケースについては考えることは少ない。今回の分析では、状態が複数存在する場合において、各状態が t 分布に従

³ 詳しくは Meng and Rubin (1991) を参照。

うケースについて検証を行う。t分布は正規分布よりも裾野が厚い分布であり、自由度が十分大きい場合には正規分布に近づくことから、自由度のパラメータが推定できれば、その分布が正規分布に近いのかどうかを検証することが容易であること、また分布の形状から最尤法によるパラメータの推定がおこないやすいなどの理由から、t分布について推定を行うこととした。

最尤法によるt分布の推定は、正規分布の場合と違い、パラメータの推定、特に自由度の推定が難しいことが知られている。本論文では、HMMにおけるt分布の推定について議論する。

本論文の構成として、第2節でt分布によるHMMの構造と、最尤法の推定方法について説明を行う。続く第3節で、実際にHMMを設定したうえで、それに合致した数値を人工的に作成して、それに対して、Lystig and Hughes (2012) の一階条件・二階条件をもとにした最尤推定量が、どの程度の正確なのかについて議論する。第4節では、新規上場市場 (Initial Public Offering: 以下、IPO) におけるアンダープライシングについて、t分布によるHMM推定を行う。

2. 隠れマルコフモデル (HMM) とは

隠れマルコフモデル (HMM) とは、マルコフ過程 (Markov Process) に従って変化する状態変数に依存した観測値を、確率的にとらえるモデルである。すなわち複数の状態の間を移動する確率変数が、各時点での状態にあるかを推定することで、観測値の変化のパターンなどをとらえることが可能となる。

HMMの構造は以下のとおりである。 Y_t を観測値、 S_t を状態変数とする。下添えの $t(t=1,2,\dots,T)$ は、時系列データが観測される時点であり、 T は最終観測時点である。そして観測値 Y_t は確率変数であり、状態変数に依存するものとする。

以下のHMMでは、状態変数がマルコフ過程に従っている状態を、す

なわち状態変数 S_t が直前の状態変数 S_{t-1} の状態にのみ、確率的に依存している状況を想定する。これは $\Pr(S_t | S_1, S_2, \dots, S_{t-1}) = \Pr(S_t | S_{t-1})$ となる状況を意味する。次に確率変数である観測値 Y_t が、観測時点の状態 S_t にのみ依存し、そのもとで確率的に決定される状況 $\Pr(Y_t | Y_1, Y_2, \dots, Y_{t-1}, S_1, S_2, \dots, S_{t-1}, S_t) = \Pr(Y_t | S_t)$ を想定する。

次に、状態変数が、時間の経過とともに、以下のように確率的に変化する状況を考える。状態変数 S_t が i 、次期に変数 S_{t+1} が j になる確率を $\gamma_{ij} = \Pr(S_{t+1} = j | S_t = i)$ とする。状態が全部で m 通りだけ存在すると考えると ($i = 1, 2, \dots, m, j = 1, 2, \dots, m$)、状態変数の移行確率は、以下のような行列形式 Γ で表現される。

$$\Gamma = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{pmatrix}$$

m は状態変数の数である。状態 S_t が i となる確率を $u_{it} = \Pr(S_t = i)$ と表現すると観測値が $Y_t = y_t$ となる確率は次のように表現できる。

$$\Pr(Y_t = y_t) = \sum_{i=1}^m \Pr(S_t = i) \Pr(Y_t = y_t | S_t = i) = \sum_{i=1}^m u_{it} p_i(y_t)$$

$p_i(y_t) = \Pr(Y_t = y_t | S_t = i)$ である。これを行列表現すると $\Pr(Y_t = y_t) = u_t P(y_t) \mathbf{1}'$,

$$u_t = (u_{t1}, \dots, u_{tm}), P(y_t) = \begin{pmatrix} p_1(y_t) & & 0 \\ & \ddots & \\ 0 & & p_m(y_t) \end{pmatrix}, \mathbf{1}' = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

となる。以上の条件から、観測値が $(Y_1 = y_1, \dots, Y_T = y_T)$ となる尤度関数 L_T は次のように定義することができる。

$$L_T = \Pr(Y^{(T)} = y^{(T)}) = \delta P(y_1) \Gamma P(y_2) \cdots \Gamma P(y_T) \mathbf{1}' \quad (1)$$

$(Y^{(T)} = y^{(T)}) = (Y_1 = y_1, Y_2 = y_2, \dots, Y_T = y_T)$ であり、 $\delta = (u_{11}, \dots, u_{m1})$ は、初期時点における状態変数 S_1 の確率分布を表している（詳しくは

Zucchini and MacDonald (2009) を参照)。

次に各状態変数のもとで、観測値がとる確率分布について、以下のような t 分布をとるケース、すなわち各状態のもとで観測値 y_i がとる確率 $p_j(y_i | X_i, \theta_j)$ が、各状態 j で自由度 v_j の t 分布に従う分布を考える。さらに観測値 y_i が、説明変数 X_i によって決まる、いわゆる線形回帰モデルを想定する。すると観測値は、説明変数と確率分布のパラメータに依存したモデルとなり、以下のように表現できる。

$$p_j(y_i | X_i, \theta_j) = \frac{\Gamma\left(\frac{v_j+1}{2}\right)}{\Gamma\left(\frac{v_j}{2}\right)\sqrt{v_j\pi\sigma_j^2}} \left(1 + \frac{(y_i - \mu_j - \beta_j X_i)^2}{\sigma_j^2}\right)^{-\frac{v_j+1}{2}}, \theta_j = (\mu_j, \beta_j, \sigma_j^2, v_j)$$

なお $\Gamma(\cdot)$ はガンマ関数である。 θ_j は、状態 j における平均 $\mu_j + \beta_j X_i$ 、分散 σ_j^2 、自由度 v_j のパラメータを表している。

HMM では、以上の尤度関数をもとに、パラメータの推定量を求めることになるが、その方法としては、最尤推定法や、EM アルゴリズム、ベイズ推定による MCMC 法などが考えられてきた。HMM では、尤度関数の一階条件を直接推定すると、計算が複雑になることから、これまで EM アルゴリズムや MCMC などの手法が用いられることが多かった。

EM アルゴリズムとは、任意の初期時点から E ステップと M ステップを、繰り返し計算を行って収束計算を行うものである。また MCMC ではパラメータに任意の事前確率分布を仮定した後、観測値による事後確率分布を求めて、それをもとに確率標本を発生させることで、パラメータの確率分布を求める手法である (詳しくは、大森裕浩・和合肇 (2005) を参照)。しかし EM アルゴリズムは、一般に計算プロセスに時間がかかること、また MCMC に関しても乱数を発生させるのに十分な量の計算をさせる必要があり、同様の問題を抱えている。Lystig and Hughes (2012) による手法では、対数尤度関数の一階条件を直接導出できることから、収束計算が比較的短いとされる準ニュートン法が応用できる。

本論文では、Lystig and Hughes (2012) による手法を検証してみる。

Lystig and Hughes (2012) は、Baum アルゴリズムをもとに、尤度関数の一階微分および二階微分の計算方法を明らかにしている。彼らはまず、初期時点において観測値 y_1 のもとで、状態 S_1 が j をとる ($S_1 = j$) 確率を

$$\lambda_1(j) = \Pr(y_1, X_1, S_1 = j) = \delta(j) p_j(y_1 | X_1, \theta_j)$$

また観測値 y_t 、初期時点 (1 時点) から $t-1$ 時点までの状態変数が (S_1, \dots, S_{t-1}) という状況の下で、状態 S_t が j をとる ($S_t = j$) 確率を $\lambda_t(j)$ とすると、

$$\lambda_t(j) = \Pr(y_t, S_t = j | S_1, \dots, S_{t-1}) = \sum_{i=1}^m \lambda_t(i) \gamma_{ij} p_j(y_t | X_t, \theta_j) / \Lambda_{t-1}, \quad \Lambda_{t-1} = \sum_{j=1}^m \lambda_{t-1}(j)$$

として計算することが可能であるとしている。なお $\Lambda_{t-1} = \Pr(S_{t-1} | S_1, \dots,$

$S_{t-2})$ である。そして対数尤度関数の値 l_T を、 $l_T = \sum_{t=1}^T \log(\Lambda_t)$ として計算で

きることを明らかにしている。

次にこの対数尤度関数の一階条件を求める。この対数尤度に含まれるパラメータは、 $\mathbf{B} = (\delta_1, \dots, \delta_m, \gamma_{11}, \dots, \gamma_{mm}, \theta_1, \dots, \theta_m)$ である。まず $t = 1$ の時の、 $\psi_1(j; \mathbf{B}) = \partial \Pr(Y_1, S_1 = j) / \partial \mathbf{B}$ と定義すると、

$$\psi_1(j; \mathbf{B}) = \partial \Pr(Y_1, H_1 = j) / \partial \mathbf{B} = \left[\frac{\partial p_j(y_1)}{\partial \mathbf{B}} \right] \delta(j) + p_j(y_1) \left[\frac{\partial \delta_j}{\partial \mathbf{B}} \right]$$

となる。次に $t \geq 2$ について同様の計算を続けると、

$$\begin{aligned} \psi_t(j; \mathbf{B}) &= \frac{\partial \Pr(Y_2, \dots, Y_t, H_t = j) / \partial \mathbf{B}}{\Pr(Y_2, \dots, Y_{t-1})} \\ &= \sum_{i=1}^m \left\{ \psi_{t-1}(i; \mathbf{B}) \gamma_{ij} p_j(y_t) + \lambda_{t-1}(i) \gamma_{ij} \left[\frac{\partial p_j(y_t)}{\partial \mathbf{B}} \right] + \lambda_{t-1}(i) \left[\frac{\partial \gamma_{ij}}{\partial \mathbf{B}} \right] p_j(y_t) \right\} \times \Lambda_{t-1}^{-1} \end{aligned}$$

となる。これは $\lambda_2(j) = p_j(y_2) \delta_j$ 、 $\lambda_t(j) = \sum_{i=1}^m [\lambda_{t-1}(i) p_j(y_t) \gamma_{ij}] \times \Lambda_{t-1}^{-1}$ 、 $\Lambda_t = \sum_{j=1}^m \lambda_t(j)$

という条件から導かれる。後はこれを繰り返し計算していくことで、最終的に

$\Psi_T(B) = \sum_{j=1}^m \psi_T(j; B) / \Lambda_T$, $\partial l_T / \partial B = \Psi_T(B) / \Lambda_T$, となる条件, すなわち HMM の対数尤度関数の一階条件が求められる (詳しくは Lystig and Hughes (2012) を参照)。よって最終的に $\partial l_T / \partial B = \Psi_T(B) / \Lambda_T = 0$ となる条件を満たすパラメータの集合 B を求めればよいことになる。

あとは HMM の最尤推定値を求めるには, $\delta_j \geq 0$, $\gamma_{ij} \geq 0$, $\sigma_j^2 > 0$, $\nu_j > 0$ となる不等式制約と, $\sum_{j=1}^m \delta_j = 1$, $\sum_{j=1}^m \gamma_{ij} = 1$ となる等式制約のもとで, l_T を最大にするパラメータを求めればよいということになる。

これは不等式制約の下での非線形最適化問題であり, 通常, 任意の初期値を定め, たうて, 一階条件をもとに, 繰り返し計算を行うことで値を求めていくことになる。本論では, 不等式制約の下での最適化問題を解くのに準ニュートン法で, 逐次二次計画法を解くことにする。なおこの解法として, Goldfarb and Idnani (1983) の手法を用いた⁴。

t 分布の推定を行う際には, EM アルゴリズムで t 分布の自由度 ν_j をどのように推定するかという問題がある。確率密度関数が t 分布の時の自由度の推定は, 結果がかなりセンシティブであることが知られている⁵。

Lystig and Hughes (2012) は, HMM の対数尤度関数の二階条件についても議論している。各パラメータの集合を B_1, B_2 , すなわち $B_1 = B_2 = (\delta_1, \dots, \delta_m, \gamma_{11}, \dots, \gamma_{m1}, \dots, \gamma_{1m}, \dots, \gamma_{mm}, \theta_1, \dots, \theta_m)$ とする。まず $t = 1$ の時の, $\omega_1(j; B_1, B_2) = \partial^2 \Pr(Y_1, S_1 = j) / \partial B_1 \partial B_2$ と定義すると,

⁴ Goldfarb and Idnani (1983) による, 不等式制約の下での非線形最適化問題の計算の応用法については, 茨木・福島 (1991), 矢部・八巻 (1999) を参照。

⁵ EM アルゴリズムによる t 分布の推定については, Liu and Rubin (1995) や Liu (1997), 中村 (1995) を参照。

$$\begin{aligned}\omega_0(j; \mathbf{B}_1, \mathbf{B}_2) &= \partial^2 \Pr(Y_1, S_1 = j) / \partial \mathbf{B}_1 \partial \mathbf{B}_2 \\ &= \left[\frac{\partial^2 p_j(y_1)}{\partial \mathbf{B}_1 \partial \mathbf{B}_2} \right] \delta(j) + p_j(y_1) \left[\frac{\partial^2 \delta_j}{\partial \mathbf{B}_1 \partial \mathbf{B}_2} \right] + \left[\frac{\partial p_j(y_1)}{\partial \mathbf{B}_1} \right] \left[\frac{\partial \delta_j}{\partial \mathbf{B}_2} \right] + \left[\frac{\partial \delta_j}{\partial \mathbf{B}_1} \right] \left[\frac{\partial p_j(y_1)}{\partial \mathbf{B}_2} \right]\end{aligned}$$

となる。次に $t \geq 2$ について同様の計算を続けると、

$$\begin{aligned}\omega_t(j; \mathbf{B}_1, \mathbf{B}_2) &= \frac{\partial^2 \Pr(Y_t, S_t = j) / \partial \mathbf{B}_1 \partial \mathbf{B}_2}{\Pr(Y_2, \dots, Y_{t-1})} \\ &= \sum_{i=1}^m \left\{ \omega_{t-1}(i; \mathbf{B}_1, \mathbf{B}_2) \gamma_{ij} p_j(y_t) + \psi_{t-1}(i; \mathbf{B}_1) \gamma_{ij} \left[\frac{\partial p_j(y_t)}{\partial \mathbf{B}_2} \right] + \psi_{t-1}(i; \mathbf{B}_1) p_j(y_t) \left[\frac{\partial \gamma_{ij}}{\partial \mathbf{B}_2} \right] \right\} \times \Lambda_{t-1}^{-1} \\ &+ \sum_{i=1}^m \left\{ \psi_{t-1}(i; \mathbf{B}_2) \gamma_{ij} \left[\frac{\partial p_j(y_t)}{\partial \mathbf{B}_1} \right] + \psi_{t-1}(i; \mathbf{B}_2) p_j(y_t) \left[\frac{\partial \gamma_{ij}}{\partial \mathbf{B}_1} \right] \right\} \times \Lambda_{t-1}^{-1} \\ &+ \sum_{i=1}^m \left\{ \lambda_{t-1}(i) \gamma_{ij} \left[\frac{\partial^2 p_j(y_t)}{\partial \mathbf{B}_1 \partial \mathbf{B}_2} \right] + \lambda_{t-1}(i) \left[\frac{\partial^2 \gamma_{ij}}{\partial \mathbf{B}_1 \partial \mathbf{B}_2} \right] p_j(y_t) \right\} \times \Lambda_{t-1}^{-1} \\ &+ \sum_{i=1}^m \left\{ \lambda_{t-1}(i) \left[\frac{\partial p_j(y_t)}{\partial \mathbf{B}_1} \right] \left[\frac{\partial \gamma_{ij}}{\partial \mathbf{B}_2} \right] + \lambda_{t-1}(i) \left[\frac{\partial p_j(y_t)}{\partial \mathbf{B}_2} \right] \left[\frac{\partial \gamma_{ij}}{\partial \mathbf{B}_1} \right] \right\} \times \Lambda_{t-1}^{-1}\end{aligned}$$

となる。ここで $\Omega_t(\mathbf{B}_1, \mathbf{B}_2) = \sum_{j=1}^m \omega_t(j; \mathbf{B}_1, \mathbf{B}_2)$ と定義すると、対数尤度関数の

二階条件は、

$$-\frac{\partial^2 l_T}{\partial \mathbf{B}_1 \partial \mathbf{B}_2} = -\frac{\Omega_T(\mathbf{B}_1, \mathbf{B}_2)}{\Lambda_T} + \frac{\Psi_T(\mathbf{B}_1) \Psi_T(\mathbf{B}_2)}{(\Lambda_T)^2}$$

として計算できる。

以上のように、HMMの一階条件および二階条件は、初期時点から逐次計算で導出できることが分かった。ではこの方法によって最尤推定法によるパラメータの推定はうまくいくのであろうか。以下では、t分布によるHMMについて、一階条件とそれにもとづく準ニュートン法による推定と、二階条件からの情報行列の推定が、うまくできるのかをみてみる。そのために、モデルに合致したデータを、人工的に作成して、それを元に、パラメータの推定量を求めていくことができるかどうかをみ

ていくことにする。

3. 最尤法による t 分布の HMM の推定の検証

Lystig and Hughes (2012) による, HMM の一階条件から, 最尤法によるパラメータの推定を行うことが果たして可能なのだろうか。HMM でない最尤法でも, t 分布のパラメータの推定は難しいことがこれまでの研究でも知られている。中村 (1995) は, 多変量 t 分布の自由度のパラメータの推定そのものが難しいことを指摘している。また Liu and Rubin (1995) や Liu (1997) などは, 多変量 t 分布の推定にニュートン法などの手法を適用することは, 計算に時間がかかることを指摘しており, より効率的な形で推定するには, EM アルゴリズムをもとに, 自由度のパラメータを, ほかのパラメータとは別に推計する方法を提唱している。本論では, HMM において, 準ニュートン法でどの程度の精度でパラメータが推計できるかを検証してみる。

以下では, 次のようなモデルを想定する。状態が 2 つ存在し (状態 1 と状態 2), その状態に応じて, 観測値が次のような t 分布に従うとする。

$$\text{状態 1 の時} \quad y_t = \mu_1 + \beta_{1,1}X_{1,t} + \beta_{1,2}X_{2,t} + \varepsilon_{1,t} \quad \varepsilon_{1,t} \sim t(0, \sigma_1^2, v_1)$$

$$\text{状態 2 の時} \quad y_t = \mu_2 + \beta_{2,1}X_{1,t} + \beta_{2,2}X_{2,t} + \varepsilon_{2,t} \quad \varepsilon_{2,t} \sim t(0, \sigma_2^2, v_2)$$

変数 $\varepsilon_{j,t}$ は攪乱項である。そして $\varepsilon_{j,t} \sim t(0, \sigma_j^2, v_j)$ は, 変数 $\varepsilon_{j,t}$ が, 平均 0, 分散 σ_j^2 , 自由度 v_j に従うことを意味する。以下では, 状態 1 の時のパラメータを, $\mu_1 = 2$, $\beta_{1,1} = 0$, $\beta_{1,2} = 1$, $\sigma_1^2 = 0.2$, $v_1 = 70$, 状態 2 の時のパラメータを $\mu_2 = 1$, $\beta_{2,1} = 0.5$, $\beta_{2,2} = -2$, $\sigma_2^2 = 0.1$, $v_2 = 20$, と設定する。また状態変数の移行確率を, 以下のように設定する。

$$\Gamma = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}$$

これらの値は, 統計学でいうところのいわゆる「真のパラメータ」と

呼んでいるものに対応する。本論では、これらのパラメータ $\mathbf{B}=(\delta_1, \dots, \delta_m, \gamma_{11}, \dots, \gamma_{mm}, \theta_1, \dots, \theta_m)$ を、最尤法をもとに、その最尤推定量を $\hat{\mathbf{B}}=(\hat{\delta}_1, \dots, \hat{\delta}_m, \hat{\gamma}_{11}, \dots, \hat{\gamma}_{mm}, \hat{\theta}_1, \dots, \hat{\theta}_m)$ と求めることとする。

最尤推定量の漸近分布は、 $\hat{\mathbf{B}} \rightarrow N[\mathbf{B}, \{I(\mathbf{B})\}^{-1}]$ 、 $I(\mathbf{B})=-E[\partial^2 l_T / \partial B_1 \partial B_2]$ となる。しかしこのHMMでは、対数尤度関数のヘッシアン行列の期待値、すなわち $I(\mathbf{B})=-E[\partial^2 l_T / \partial B_1 \partial B_2]$ を計算することは難しい。そこで本論では $I(\hat{\mathbf{B}})=-[\partial^2 l_T / \partial \hat{B}_1 \partial \hat{B}_2]$ をつかって、パラメータの分散・共分散行列を計算する（詳しくは Greene (1997) を参照）。もしこの情報行列がうまく推定できると、Wald 検定による仮説検定が行えることになり、線形回帰モデルで重要なパラメータの仮説検定が容易になるというメリットがある。

このHMMの最尤法を計算するのに、非線形最適化問題を解くことになる。本論では、尤度関数の一階条件をもとに、不等式制約下における準ニュートン法を用いた、逐次二次計画法による推計をおこなう。準ニュートン法は、非線形最適化問題を解く際に、対数尤度関数の一階条件のみを使用するため、二階条件を計算する必要はない。これは繰り返し計算の途中で、ヘッシアンが最適条件を満たさない場合があるからである。準ニュートン法は、この問題を回避できる。さらに本論で採用した Lystig and Hughes (2012) でも、二階条件の計算はプログラムが若干複雑になるという問題がある。今回、対数尤度関数の二階条件は、最適条件を計算したのちに、パラメータの分散・共分散行列を計算する場合のみに使用した。

また非線形最適化問題を解くのに必要となる、逐次二次計画法の解法として、Goldfarb and Idnani (1983) の手法を用いた。この不等式制約の下での非線形最適化問題の計算については、茨木・福島 (1991) および矢部・八巻 (1999) を参考にした。なお本論の推定を実施するにあたり、Ox Console Version 7.00 64bit 版（詳しくは Doornik, 2012を参照）および The Arfima package Version 1.00（詳しくは Doornik and Ooms,

2003を参照) のソフトウェアを使用した。

次に初期値の設定について説明をする。非線形最適化問題を解く際に必要な、パラメータの初期値の設定については、 μ 、 β については、通常の最小二乗法の推定量を求めたうえで、0から1の範囲の一様分布の乱数を加えた。その上で、残差の二乗和を標本数で除した値を σ^2 の初期値とした。

δ および γ については、0から1の範囲の一様分布の乱数を発生させたうえで、それぞれのパラメータについて、合計が1になるよう値を調整した。具体的には、初期時点における状態変数の確率を表す δ については、乱数を発生させたうえで、その合計値を一度求め、再度その合計値で、 δ の各値を除することで、合計が1になるように基準化した。また γ についても、同様に、一様分布の乱数を発生させたうえで、各行の値を合計し、再度 γ の値を、各行の合計値で除することで、合計が1になるよう調整した。

またt分布の自由度を表す ν については、50という値を割り振った。この自由度のパラメータの推定は、非常にセンシティブであり、推定も難しい。そのため初期値をどのように設定するかは非常に重要な問題であるが、実際の推定では、自由度の真の値を知ることはできない。本論では、こうした問題点があることを理解したうえで、とりあえず初期値を、正規分布に近い分布になると思われる50に設定する。

以下、データ数 (T) を1000, 4000, 5000にして、それぞれのケースでパラメータおよび情報行列から求められる標準偏差についてみてみる。まずパラメータの推定量についてみてみると、自由度のパラメータ以外は、おおむね当初の設定したパラメータに近い値が求められていることがわかる。

他方、自由度のパラメータ ν の推定量は、初期値からほとんど変わっていないことがわかる。これはt分布の自由度が変化しても、尤度関数の値にさほど影響を与えないことを意味しているといえる。そのため真

の値に限りなく近い自由度の値を求めるためには、それなりの工夫が必要になってくる。準ニュートン法を用いた推定では、すくなくとも自由度の推定量を求めるためには、初期値をどのように設定するかが重要になってくることになる。

つぎに情報行列に関してしてみると、データ数が1000の場合、 δ_2 の推定量に関する分散の値が負になるという問題が発生した。すなわち分散の値が -9.83×10^{82} という極端に小さな、しかも取りうるはずのない負値をとっていることがわかる。しかしデータ数が5000（表1の第3列目の結果）になると正の値をとっている（ 6.89×10^{86} ）ことから、これはデータ数が小さいために真の情報行列に十分近づいていないためと思われる。

また情報行列による仮説検定についてみる。ここではそれぞれのパラメータについて0となる帰無仮説を、Wald検定をもとに行ってみる（Wald検定に関する詳しく説明はGreene（1997）などを参照）。この場合、統計量は帰無仮説を満たしているときに、自由度1のカイ二乗分布に従う。結果をみると、初期値の状態確率を表す δ 、 $\hat{\beta}_{1,1}$ （真のパラメータが0）および u_1 以外は、帰無仮説を棄却する結果となっている。また δ および u_1 については、データ数が多くなるにつれて検定量が大きくなっていることから、標本数が多くなるほど検定結果もより信用性が増しているといえる。

表 1 推計結果

	パラメータ の設定値	T=1000	T=4000	T=5000	T=5000 (初期値 : 混合分布)
δ_1	1	0.8444 [0.844]	0.99852 [0.999]	0.99887 [0.999]	1 [0.702]
δ_2	0	0.1556 N.A	0.0014787 [3.48×E+42]	0.0011301 [2.63×E+43]	3.6642E-23 N.A
γ_{11}	0.6	0.62673 *** [0.038]	0.58611 *** [0.02]	0.60668 *** [0.017]	0.60663 *** [0.017]
γ_{21}	0.3	0.30933 *** [0.024]	0.28008 *** [0.011]	0.30903 *** [0.011]	0.30902 *** [0.011]
γ_{12}	0.4	0.37327 *** [0.029]	0.41389 *** [0.016]	0.39332 *** [0.014]	0.39337 *** [0.014]
γ_{22}	0.7	0.69067 *** [0.036]	0.71992 *** [0.018]	0.69097 *** [0.016]	0.69098 *** [0.016]
μ_1	2	2.0131 *** [0.02]	2.0026 *** [0.012]	1.9994 *** [0.01]	1.9994 *** [0.01]
$\beta_{1,1}$	0	-0.0081049 [0.007]	-0.00076124 [0.004]	0.0026486 [0.003]	0.0027006 [0.003]
$\beta_{1,2}$	1	0.99686 *** [0.006]	1.0022 *** [0.004]	0.99662 *** [0.003]	0.99654 *** [0.003]
$\sigma_{1,2}$	0.2	0.1724 *** [0.015]	0.20577 *** [0.01]	0.21406 *** [0.009]	0.21969 *** [0.013]
ν_1	70	50.482 [75.651]	51.009 [43.023]	51.303 [35.082]	177.55 [846.921]
μ_2	1	0.99935 *** [0.014]	1.0056 *** [0.007]	1.0031 *** [0.006]	1.0033 *** [0.006]
$\beta_{2,1}$	0.5	0.49347 *** [0.005]	0.4989 *** [0.002]	0.50086 *** [0.002]	0.50094 *** [0.002]
$\beta_{2,2}$	-2	-1.999 *** [0.005]	-2.0005 *** [0.002]	-2.0004 *** [0.002]	-2.0005 *** [0.002]
$\sigma_{2,2}$	0.1	0.10065 *** [0.006]	0.10644 *** [0.003]	0.099475 *** [0.003]	0.093959 *** [0.003]
ν_2	20	50.41 *** [1.097]	51.432 *** [0.546]	51.44 *** [0.479]	20.533 *** [0.167]
	AIC	2065.9	8324	10571	10567
	BIC	2129.7	8405.8	10655	10652

***, **, *は、それぞれ 1%, 5%, 10% で有意
帰無仮説 : 各パラメータの値が 0

以上の結果から、(1)Lystig and Hughes (2012) をもとにした最尤法による推定は、自由度のパラメータを除くと、比較的少ないデータ数でも真のパラメータに近い推定量を推定できる、(2)Wald 検定を行うのに必要な情報行列の正確な推定は、比較的多いデータを必要とする、(3)自由度のパラメータの推定量については、推定が比較的難しいといえるであろう。

では自由度のパラメータを求めるにはどのようにすればよいのであろうか。もし真のパラメータがわかっているならば、真のパラメータに近い値に初期値を設定して、繰り返し計算をしてするのがよいであろう。しかし真のパラメータ自体を、分析者はわかっていないことが通常であり、この方法は非現実的であるといえるであろう。

これを解決する方法の一つとして、最初に混合分布に基づく最尤推定法をおこない、そこで各パラメータを推定したうえで、その値を初期値に設定して、準ニュートン法による HMM の推定を行うという方法である。EM アルゴリズムによる混合分布の推計は、繰り返し計算の回数を比較的多くとるの必要があり、計算時間が長いという欠点があるが、比較的安定的な解が得られることが知られている。

表1の最右の結果は、初期値を混合分布で求めたのちに、それを HMM の初期値として推計をした結果である。この結果、自由度のパラメータの一方 ν_2 は、20に近い値を推定できたことがわかった。しかし ν_1 については、170と70から外れた値をとっていることがわかる。またデータ数が1000の場合と同様、 δ_2 の推定量の分散値が負になるという問題が発生した。こうした点を考えると、初期値を混合分布で求める方法についても、いろいろと検討の余地があると思われる。

次節では、新規市場におけるアンダープライシングの分析をみながら、この方法について考察してみる。

4. 分析例

次に t 分布による HMM の分析例をみしてみる。以下の分析は、比佐・比佐 (2014) による新規公開市場におけるアンダープライシングに関する分析である。企業が株式市場に新規上場 (Initial Public Offering: 以下 IPO) する際、株式を公開して広く株主に販売をする。この時の価格を公示価格というが、その後、株式市場で取引されると、市場で改めて株価が決定される。その取引開始時点の最初の価格を初値と呼ぶが、上場前の公示価格と上場直後の初値との間に、アンダープライシングと呼ばれる価格の乖離が存在することが知られている。通常、株式を公示してから、上場して取引されるまでの間の期間はさほど長くないことから、本来ならばそれほど大きな価格差はつかないと考えられる。しかし実際の取引では、一般に、上場前の公示価格が上場直後の株価よりも過少になることが一般に知られている。しかもこの現象は、日本のみならず、世界中の IPO 市場でみられる一般的な現象である。

このような、一見すると非合理的な株価の動きがなぜ起こるのかという問題が、これまでの研究で行われていた。というのも、もし上場前の公示価格が上場直後の株価よりも平均的に低いのであれば未公開株を購入して、上場後に売却をすれば、利益を獲得することができる。そのためこの現象の存在がわかっているならば、この利益を求めて、多くの市場参加者がこぞって上場前の未公開株を購入し、上場直後に売却しようとするであろう。その結果、上場前の公示価格は上昇、上場直後の初値は下落するであろう。そしてこうした裁定行為によって、両者の価格の差は縮小し、最終的に乖離はなくなるはずである。しかし実際には、こうした乖離はつねに存在するのである。

一般に、上場後の株式市場のほうが、多くの投資家が市場に参加することから、より多くの情報が株価に反映されているとするならば、上場後の市場のほうがより効率的であると思われる。よって株式市場で効率的市場仮説が成立しているのであれば、公開後の初値のほうがより合理

的に価格が形成されていると考えられるであろう。以上の理由から、上場前の価格が、本来あるべき価格（すなわち初値）よりも低い、いわゆる過小な値段で公示されるという意味において、これらの現象はアンダープライシング（Underpricing）と呼ばれている⁶。

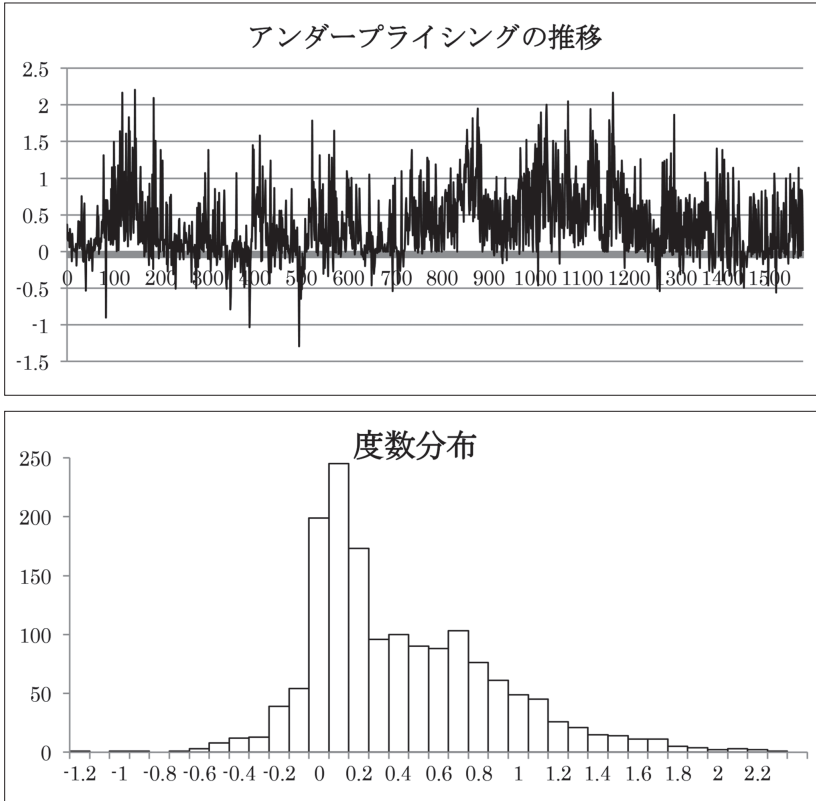
このアンダープライシングには、循環的な動きがみられることが知られている。1998年から2012年での14年間のIPO企業上場（1573社）⁷を上場した順番に時系列的に追いかけると⁸、次のような動きになっていることがわかる。

⁶ 上場前の公示価格については、最近はブックビルド方式が主流である。この方式はIPOの株式を引き受けた主幹事証券会社が、投資家への聞き取り調査などを通じて、詳細な調査をおこなうことで公示価格を決定している。そのため公示価格で非効率的な価格が行われているのかという点については議論の余地があると思われる。加えて株式市場では、初値が付いた後、急激に株価が下落し、その後数年かけてゆっくりと上昇する現象が存在する。こうした現象からも、実際の上場後の株式市場で本当に合理的な価格が形成されているかどうかは、検討の余地がある。

⁷ 東証一部上場および移行上場のIPOについては、すでに多くの投資家によってよく知られた企業が上場することから、情報の非対称性がさほど働かないといえる。またこうした上場は、IPO市場でも特殊な上場と位置付けられていることから、分析対象から除外している。

⁸ なお同日に上場した企業が複数ある場合は、資産規模の大きい企業のデータを先に配置した。その理由として、資産規模の大きな企業の株価の動きは、多くの投資家により注目されること、またその動向が他のIPOの株価にも影響を与えることと思われるからである。なおIPO企業の株に対しては、それを専門とする投資家もおり、特殊な株価がつく投資市場であるとの認識がある。またIPOに関しては、過去のIPOの株価の動向が、こうした投資の行動に影響を与える可能性があることも、こうした分析を行った理由である。つまり過去のIPOの株価の動向が、現在のIPOの株価の動きに影響を与えているのかどうかを検証することも、今回の分析の主旨である。

図1 アンダープライシングの推移と分布 (1998年-2012年)



出典：比佐・比佐 (2014)

表2 アンダープライシングの基本統計量

平均	標準偏差	25パーセンタイル	中央値	75パーセンタイル	最大値	最小値
0.366	0.465	0.025	0.235	0.676	2.207	-1.291

出典：比佐・比佐 (2014)

この初期収益率（アンダープライシング）は、対数の差分で評価している⁹。これをみると、アンダープライシングが上下に変動しながら推移していること（図1のアンダープライシングの推移）、また分布が複数存在している（図1の分布）ように思われる。分布がこのように複数になることは、Asquith et al. (1997) によって知られている。図1から、二つの確率分布の間を、推移している可能性があると思われる。これをHMMで分析を行うと、次のような結果となった（ $\beta = 0$ のケース）。

表3 パラメータおよび確率移行行列の推定（状態が2つの時）

状態 (S_i)	混合分布		HMM	
	状態 1	状態 2	状態 1	状態 2
$\hat{\delta}$	0.18736	0.81264	0.24685	0.75315
$\hat{\mu}_j$	0.45214	0.31994	0.53986	0.037929
$\hat{\sigma}_j^2$	0.23339	0.16539	0.23634	0.013139
\hat{v}_j	1892.8	2589.3	1892.8	2589.3
	AIC	BIC	AIC	BIC
	2080.3	2117.8	1537	1585.2

$$\hat{\Gamma} = \begin{pmatrix} 0.8274 & 0.1726 \\ 0.3247 & 0.6753 \end{pmatrix}$$

出典：比佐・比佐（2014）

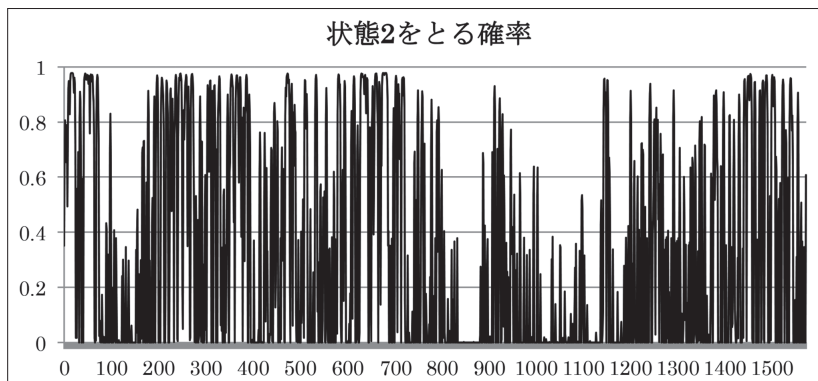
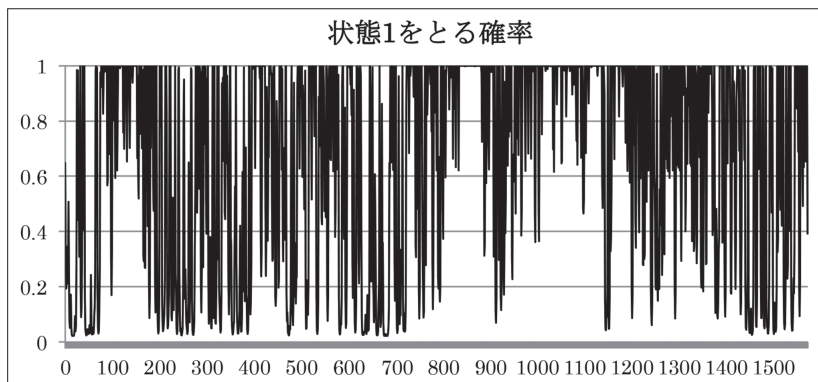
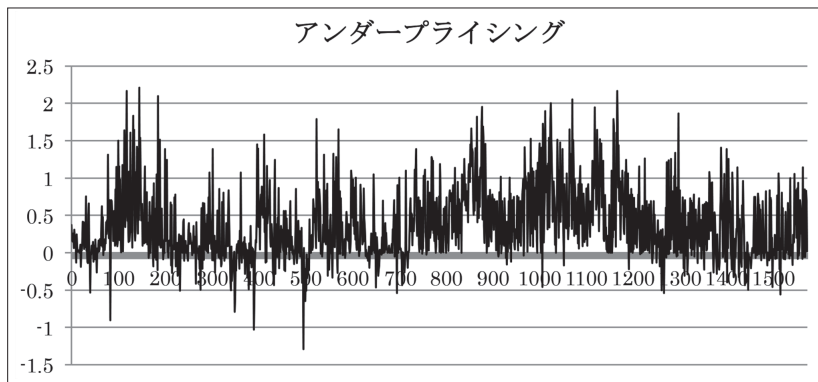
⁹ 今回、IPO企業の株価の公開価格を X_t 、初値を X_{t+1} としたときに、アンダープライシングを $\ln(X_{t+1}/X_t) = \ln(X_{t+1}) - \ln(X_t) = \Delta \ln(X_t)$ として、公開価格と初値をそれぞれ対数化したものの差分で評価した。アンダープライシングは、(初値 - 公開価格) / 公開価格、すなわち $(X_{t+1} - X_t) / X_t$ で評価する場合もあるが、そうすると下限が -1 になってしまうことで、分布にゆがみが生じてしまう。そのため今回は、t分布をフィットさせることもあり、下限が $-\infty$ である、対数の差分でアンダープライシングの収益率を評価した。

最尤推定法によると、アンダープライシングの度合いが高い状態を状態1、アンダープライシングの度合いが低い状態を状態2とすると、状態1の時は、平均0.53986、分散0.23634、自由度1892.8のt分布、状態2の時は、平均0.037929、分散0.013139、自由度2589.3のt分布になった。そして確率移行行列をみてみると、状態1にとどまる確率が0.8274、状態1から状態2に移行する確率が0.1726、状態2にとどまる確率が0.6753、状態2から状態1に移行する確率0.3247となる。またAICやBICを比較すると、混合分布を推計した場合よりも、HMMのほうが、その値が小さいことから、混合分布よりもHMMのほうがモデルとしてのフィットがよいといえる。

なお自由度の推定については、最初に混合分布に基づく最尤推定法をおこない、そこで各パラメータを推定したうえで、その値をもとに準ニュートン法によるHMMの推定を行った。中村(1995)の論文をもとに、EMアルゴリズムによる混合分布による推計を行った。混合分布による推計では、
$$\bar{L}_T = \sum_{t=1}^T \log \sum_{j=1}^m \pi_j p_j(y_t), \quad \sum_{j=1}^m \pi_j = 1$$
となる対数尤度関数を解くことになる。

なおEMアルゴリズムによる混合分布の推計は、繰り返し計算の回数を比較的多くとる必要があり、計算時間が長いという欠点があるが、比較的安定的な解が得られることが知られている。

図2 アンダープライシングと状態確率 (1998年—2012年)



出典：比佐・比佐 (2014)

上の図は、状態が2つの時の、 $\Pr(S_t = i | Y^{(T)} = y^{(T)})$ ($i = 1, 2$) の推移である。すなわち観測値($Y^{(T)} = y^{(T)}$)が与えられたもとで、それぞれの時点において、いずれの状態をとるのか(いわゆる local decode)を推定している。この隠れマルコフによる状態変数の確率をみевみると、アンダープライシングが増加するタイミングに合わせて、状態1をとる確率が上昇し、状態2をとる確率が下落していることがわかる。

こうしたアンダープライシングについて循環的な動きが存在する理由として、以下のようなものが考えらえる。すなわちIPOが盛んなときは、新規上場した企業も多く、また企業も株式市場で資金を調達することが容易であり、企業の投資活動も積極的に行われる。その一方で、IPOが低迷しているときは、企業の資金調達も難しくなるため、企業の投資活動も停滞するとされる。そのため、IPOの変動は、景気循環とも密接に関係しているのではないかという、見方がなされている。Yung et al. (2008)は、IPOを達成する企業数の総数は、投資家と株式公開企業との間の情報の非対称性によって変動し、その結果、アンダープライシングが拡大すると主張している。つまり、景気が良いときには、情報の非対称性が強まり、景気が悪いときには弱まるとした。

その一方で、アンダープライスの要因については、ミクロ的視点からの研究が中心に進められてきた。そして多くの研究では、新規上場における情報の非対称性の問題に着目している。そしてコーポレート・ガバナンスの観点から、銀行、証券会社、ベンチャーキャピタル(以下VCと記す)は情報の非対称性を緩和し、アンダープライシングを抑制するとみなされている。

アンダープライシングは、各国の証券市場でも確認され、その要因についてはアメリカを中心に様々な角度から分析がおこなわれている。Ljungqvist (2007)は、こうした研究の特徴を4つの要因に整理している。

一つ目は、投資家と株式公開企業との間に存在する情報の非対称性によって、IPOの変動が引き起こされるとする情報の非対称性モデルで

ある¹⁰。代表的な研究として、投資家間の情報の非対称性に着目した Rock (1986) の研究と、投資家と企業間の情報の非対称性に着目した Allen and Faulhaber (1989) の研究がある¹¹。二つ目は、Hughes and Thakor (1992) によるアンダープライシングの発生が、引き受け証券会社による引き受けリスクの調整によって引き起こされるとみなす仮説である¹²。三つ目は、企業の所有者の経営参加などの度合いなどに着目し、所有と経営の関係がアンダープライシングの発生要因とみなす見方である¹³。そして最後は、近年の行動経済学の視点から投資家の心理的要因に着目したプロスペクト理論である¹⁴。

いずれにしても、これらの要因は、IPO を実施する個別企業のガバナンス構造などに依存する形となっている¹⁵。すなわちガバナンス構造に

¹⁰ 情報の非対称性によって、引き起こされる説として、Rock (1986) の勝者の呪い、情報顕示理論、プリンシパル・エージェントモデル、シグナリング・モデルを挙げている。

¹¹ Allen and Faulhaber (1989) は、投資家は優良な企業とそうでない企業の見分けがつかないことに着目した。この場合、優良企業は、上場後に情報の非対称性が緩和されても問題がない。そのため、あえて IPO 時点で低い価格をつけることで、自らは質が高い企業であるというシグナルを市場に発信する。その結果としてアンダープライシングが発生するとした (Grinblatt and Hwang (1989), Welch (1989), Beatty and Ritter (1986))。

¹² Hughes and Thakor (1992) によれば IPO が盛んな状況では、売れ残り株の引き受け、公開直後の価格の下落などの取引リスクが高まるため、引き受け証券会社は引き受け手数料以上のコストを負うことを回避する。そのためアンダープライシングが生じるといふ。また証券会社は損出をこうむった投資家からの訴訟を回避する目的やエージェンシー関係から顧客である投資家をもうけさせるために、あえてアンダープライシングを導くという説も存在する (Tinic (1988), Baron (1982))。

¹³ Holmström and Tirole (1993) は、ストックオプションを通じて経営者に効率的な報酬を提示できるとしており、こうした制度の導入がシグナルとなっている可能性がある。

¹⁴ 投資家が他人の行動に横並び行動、また感情や過大な期待、投資家の一時的な過熱に振り回される結果、などがアンダープライシングを引き起こす要因としている。

¹⁵ コーポレートガバナンスの立場からみた、日本のアンダープライシングの研究に関しては、Hamao et al. (2000) や Kutsuna et al. (2002)、小西 (2004)、鈴木 (2004)、辰巳・桂山 (2005) を参照。

問題があって、情報の格差が大きくアンダープライスが過大になる可能性のある企業が IPO を実施したのちに、情報の格差が小さくアンダープライスの小さくなる可能性がある企業が IPO を実施すると、表面的には、IPO が変動したようにみえるであろう。しかしこれは循環的要因ではなく、各企業の個別要因によるものといえる。くわえてこれらの要因も、先ほどもみたように、IPO 市場が活性化しているときとそうでないときで、それらの効果が異なってくる可能性も考えられる。

このように IPO のアンダープライシングについては、景気変動による要素と、個別企業の情報の非対称性の度合いと、それによるコーポレート・ガバナンス構造による要素とが、それぞれ影響を与えているといえる。そしてこれらの要因を同時に考慮して、分析を行うには、HMM が適しているといえる。このモデルでは、景気変動による要因を推移確率行列 Γ や定数項 μ 、また個別企業によるコーポレート・ガバナンス要因を、 $p(y_i | X_i, \theta)$ に含まれる、説明変数 X_i の係数 (β) ごとに分けて、それぞれの効果を別々に推定することが可能である。

比佐・比佐 (2014) では、こうした点を考慮した分析を行っている。分析では、①負債比率が高い場合、金融機関などによるモニタリングが強まり、アンダープライシングが起りにくい¹⁶、②銀行の持株比率が高いと、アンダープライシングは起りやすい¹⁷、③ベンチャーキャピタル (以下、VC) の支援を受けている企業では、アンダープライシングは起りにくい¹⁸、④ストックオプションを導入している企業では、

¹⁶ 金融機関からの借り入れに依存が強い場合、金融機関によって負債を通じたモニターが強まるとする (Jensen (1986) か、企業は新規の設備投資案件に対して融資を受けることができず、企業価値を高める機会を逸する可能性 (Myers (1977)) があるという、異なる見解が存在する。

¹⁷ 日本では、銀行が中心となって企業の情報を集め、情報の非対称性の問題を解消してきたとされる。そして投資先企業は、上場後も銀行の顧客として長期的関係を構築する。そのため銀行によって株式を保有されているということは、その企業が優良企業であることを示すシグナルとなる。

¹⁸ VC は将来性の高い企業に投資を行い、投資先企業の経営監視・経営支援を行いな

アンダープライシングは起こりにくい¹⁹、⑤主幹事証券会社の株式の引き受けが大きいほど、アンダープライシングが起こりやすくなる²⁰、などの仮説を立てて検証を行っている。

比佐・比佐（2014）では、1998年から2012年に日本の株式市場でIPOを行った企業1573社についての、アンダープライシングに関する分析を行っている。表4は、IPO実施企業に関する基本統計量である。

また彼らは、状態を2つに分けた場合について分析を行っている（結果は表5）。回帰分析の結果をみると、状態1では定数項が平均的に高く（アンダープライスが高い）、また状態2では、定数項が平均的に低い（アンダープライスが低い）状況であるといえる。そしてIPOがさかんな状態1では、資産収益率（ROA）が高いほど、アンダープライスが大きいことがわかる。また資産規模が大きいほど、アンダープライスが小さくなること、規模の大きな企業が上場するときは、アンダープライスが起こりにくいことが判明した。また自己資本比率が大きいほどアンダープライスが小さくなること、また主幹事証券会社の株式引き受け比率が大きいほど、アンダープライスが起こりにくくなることも明らかにした。

他方、アンダープライスが起こりにくい状態2の時は、上場廃止ダミーが有意な結果となった。これは上場後、業績悪化などにより上場を維持することが難しかった企業に関して、アンダープライスが高めになる

がら、IPOを達成させることを目的としている。Megginson and Weiss (1991) は、VCの支援を受けている企業はアンダープライシングが低いことを示している。そのほか、VCとIPOの関係に関する論文としては、Barry et al. (1990), Lerner (1994), Gompers (1996), Brander et al (2002), Lockett and Wright (2003), Lee and Wahal (2004), Li and Masulis (2005) などがある。

¹⁹ ストックオプション制度は役員インセンティブを高めるだけでなく、投資家に対して企業の経営方針に関する肯定的な情報を発信していると考えられる。

²⁰ 証券会社については、アンダープライシングの発生を引き受け証券会社による引き受けリスクの調整とみなす仮説がある。IPOが盛んな状況では、売れ残り株の引き受け、上場直後の価格の下落などの取引リスクが高まるため、アンダープライシングに導くという。

ことを意味しており、こうした企業の株価については、公示価格の段階で、あらかじめ低い価格を設定されているといえよう。また資産収益率や資産規模が大きくなるほど、アンダープライスが大きくなる傾向があるといえる。また状態1の時と異なり、自己資本比率や主幹事証券会社への引き受け比率が大きいほど、アンダープライスが大きくなることがわかった。

また銀行の持株比率は、アンダープライスに何ら影響を与えておらず、銀行は未上場企業について、情報の非対称性を緩和するような働きをしていないと考えられる。つまり銀行は未上場段階では、企業の経営活動のモニタリング等について、あまり中心的役割を果たしていないからと思われる。

表4 新規上場企業に関する基本統計量

	平均	標準偏差	中央値	最大	最小
アンダープライシング (対数差分)	0.3659	0.4652	0.2351	2.2073	-1.2910
上場廃止ダミー	0.0445	0.2063	0	1	0
VC 持株比率 (%)	9.1536	13.0240	3.59	91.42	0
銀行持株比率 (%)	1.9208	3.1495	0	23.8	0
資産収益率 (ROA)	0.1173	0.1490	0.0957	1.5253	-1.5746
資産規模 (対数)	8.4874	1.4424	8.4194	16.8853	4.5109
ストックオプションダミー	0.5016	0.5002	1	1	0
自己資本比率 (%)	40.7765	23.3869	36.7914	100.0000	-20.2703
主幹事証券会社比率 (%)	66.0530	9.2989	65	90	21
CEO 年齢	52.0324	10.3978	53	79	25
ジャスダックダミー	0.1914	0.3935	0	1	0
新規店頭ダミー	0.3102	0.4627	0	1	0
ヘラクレス上場ダミー	0.0890	0.2848	0	1	0
マザーズダミー	0.1799	0.3842	0	1	0

出典：比佐・比佐 (2014)

表5 アンダープライスの決定要因

定数項	状態1	状態2
	係数 [標準誤差]	係数 [標準誤差]
定数項	1.2485 [0.2107] ***	-0.63904 [0.9785]
上場廃止ダミー	0.0073565 [0.0764]	0.67503 [0.3909] *
VC持株比率 (%)	-0.0030843 [0.0013] **	0.00010774 [0.0004]
銀行持株比率 (%)	0.0000435 [0.0059]	-0.0028782 [0.014]
資産収益率 (ROA)	0.14487 [0.0663] **	0.53799 [0.1123] ***
資産規模 (対数)	-0.095395 [0.0192] ***	0.17473 [0.0307] ***
ストックオプションダミー	0.060104 [0.0453]	-0.1471 [0.1307]
自己資本比率 (%)	-0.0020986 [0.0009] **	0.0040814 [0.0016] ***
主幹事証券会社比率 (%)	-0.0032841 [0.0019] *	0.014933 [0.0076] *
CEO年齢	0.00090409 [0.0039]	-0.023296 [0.0165]
ジャスダックダミー	0.21816 [0.1176] *	-0.78629 [0.5043]
新規店頭ダミー	0.025954 [0.1339]	-0.45768 [0.5493]
ヘラクレス上場ダミー	0.30592 [0.0845] ***	0.25833 [0.2955]
マザーズダミー	0.32341 [0.0747] ***	-0.69132 [0.1638] ***
σ^2	0.11185 [0.0177] ***	0.11827 [0.0029] ***
ρ	5.3775 [1.3285] ***	20.127 [3.8942] ***

$$\hat{\Gamma} = \begin{pmatrix} 0.802 & 0.198 \\ 0.472 & 0.528 \end{pmatrix}$$

***, **, * は、それぞれ1%, 5%, 10%で有意
出典：比佐・比佐 (2014)

HMMによるアンダープライスの分析を行うと、アンダープライシングには循環的な動きがみられることや、ガバナンス構造がアンダープライスに与える影響が、レジームによって異なる効果がある可能性があることが確認できるといえる。たとえばVCが投資を行った場合は、IPO市場が活発でアンダープライスが大きくなるような状況では、アンダー

プライスを小さくする働きをしていることがわかった。また状態とは関係なく、資産収益率が大きい企業ほど、アンダープライスが大きくなることがわかった。つまり VC が融資する場合、通常はアンダープライスを軽減するような効果はみられないが、IPO 市場が盛んになると、アンダープライスを縮小する役割を果たしていることが判明した（詳しい議論は、比佐・比佐（2014）を参照）。

このように循環的な動きがみられる現象が、個別の要因によるものなのかあるいは循環的要因なのか判別が難しい場合、上記のような HMM による分析が有益であることがいえる。

5. 結論

本論では、 t 分布における HMM の推定について分析を行った。Lystig and Hughes (2012) による対数尤度関数の一階微分条件をもとに、準ニュートン法による最尤推定法を、モデルにあわせて人工的に作成されたデータをつかって検証した。その結果、自由度以外のパラメータは、比較的少ない標本数でも、おおむね真のパラメータに近いものとなった。一方、自由度に関するパラメータに関しては、準ニュートン法による推定では、標本数を増やしていても、自由度の初期値の値からさほど変化しない結果となった。

この問題を解決するためには、初期値をあらかじめ真のパラメータに近い値に設定するという方法も考えられるが、そもそも真のパラメータがわからない状態で推定を行わなければならないわけで、この方法はあまりにも非現実的な解決法であるといえる。

これを解決する一つの方法として、EM アルゴリズムによる混合分布による推定を行い、それを HMM の最尤推定量を求める際の、準ニュートン法の初期値に設定するという方法が考えられる。今後の課題として、検討を行う余地があると思われる。

また情報行列にもとづく分散・共分散行列を推定すると、データ数が

十分に大きくなると、それなりに信憑性のあるものに収斂する可能性があることがわかった。それでも自由度のパラメータに関する分散値については、十分な値に収斂していない可能性がある。これは自由度のパラメータが真の値に十分近い値に収斂していないことが理由として考えられる。

またt分布によるHMMは、IPOのアンダープライシングのような、循環的に変動しながらも、同時に各企業の持つ属性にも影響を受けるような現象の分析を行うのに、適したモデルといえる。またガバナンス要因が、状態の違いによって効果が異なるかどうかの検証を行うことも可能である。分析では、景気循環との関係等を分析しながら、アンダープライシングがどのような要因で決定されるのかなどを分析することが可能となるであろう。

6. 参考文献

- 茨木俊秀・福島雅夫 (1991) 『Fortran77 最適化プログラミング』岩波書店。
- 大森裕浩・和合肇 (2005) 「マルコフ連鎖モンテカルロ法とその応用」和合肇編『ベイズ計量経済分析』第2章, 東洋経済新報社。
- 小西大 (2004) 「銀行系ベンチャーキャピタルの役割」『郵貯資金研究』第13巻, 45-59。
- 鈴木健嗣 (2004) 「引受シンジケートの構成が新規公開費用に与える影響」『一橋論叢』第132巻第5号, 87-106。
- 辰巳憲一・桂山靖代 (2005) 「IPOリターン・リバーサル: 初取引日前後のIPOパフォーマンスのアノマリー分析」『学習院大学経済論集』第42巻第3号, 197-208。
- 中村永友 (1995) 「クラスター化法の統計的評価とその応用」『総合研究大学院大学平成6年度学位論文』。
- 比佐優子・比佐章一 (2014) 「新規公開市場のアンダープライシングの

決定要因—隠れマルコフモデルによるアプローチ—」『帝京経済学研究』第74巻第2号, 81-100.

矢部博・八巻直一 (1999) 『非線形計画法』朝倉書店.

Asquith, D., Jones, J. D. and Kieschnick, R. (1997) “Evidence on Price Stabilization and Underpricing in Early IPO returns,” *The Journal of Finance*, 53, No.5, 1759-1773.

Allen, F. and Faulhaber, G. R. (1989) “Signaling by Underpricing in the IPO Market,” *Journal of Financial Economics*, 23, 303-323.

Baron, D. P. (1982) “A Model of the Demand for Investment Banking Advising and Distribution Services for New Issues,” *The Journal of Finance*, 37, No.4, 955-976.

Barry, C., Muscarella, C., Peavy, J. W. and Vetsuypens, M. (1990) “The Role of Venture Capital in the Creation of Public Companies: Evidence from the Going-Public Process,” *Journal of Financial Economics*, 27, 447-471.

Beatty, Randolph P., and Ritter, Jay R. (1986) “Investment Banking, Reputation, and the Underpricing of Initial Public Offerings,” *Journal of Financial Economics*, 15, 213-232.

Brander, J. A., Amit, R. and Antweiler, W. (2002) “Venture Capital Syndication: Improved Venture Selection Versus the Value-added Hypothesis,” *Journal of Economics and Management Strategy*, 11, 423-452.

Doornik, J. A. (2012) *Object-Oriented Matrix Programming Using Ox 3rd ed.*, London: Timberlake Consultants Press and Oxford: www.doornik.com..

Doornik, J. A. and Ooms, M. (2003) “Computational Aspects of Maximum Likelihood Estimation of Autoregressive Fractionally Integrated Moving Average Models,” *Computational Statistics and Data*

- Analysis*, 41, 333-348. Also see www.doornik.com..
- Goldfarb, D. and Idnani, A. (1983) "A Numerically Stable Dual Method for Solving Strictly Convex Quadratic Programs," *Mathematical Programming*, 27, 1-33.
- Greene, William H. (1997) *Econometric Analysis Third Edition*, Prentice-Hall, Inc..
- Grinblatt, M. and Hwang, C. Y. (1989) "Signalling and the Pricing of New Issues," *The Journal of Finance*, 44, No.2, 393-420.
- Gompers, P. A. (1996) "Grandstanding in the Venture Capital Industry," *Journal of Financial Economics*, 42, 133-156.
- Hamao, Y., Packer, F. and Ritter, J. (2000) "Institutional Affiliation and the Role of Venture Capital: Evidence from Initial Public Offerings in Japan," *Pacific Basin Finance Journal*, 8, 529-558.
- Holmström, B. and Tirole, J. (1993) "Market Liquidity and Performance Monitoring," *Journal of Political Economy*, 101, No.4, 678-709.
- Hughes, Patricia J. and Thakor, Anjan V. (1992) "Litigation Risk, Intermediation, and the Underpricing of Initial Public Offerings," *Review of Financial Studies*, 5, No.4, 709-742.
- Jensen, M. C. (1986) "Agency Costs of Free Cash Flow, Corporate Finance and Takeovers," *The American Economic Review*, 76, No.2, 323-329.
- Jensen, M. C. and Meckling, W. H. (1976) "Theory of the Firm Managerial Behavior, Agency Costs, and Capital Structure," *Journal of Financial Economics*, 3, 305-360.
- Kutsuna, K., Okamura, H., and Cowling, M. (2002) "Ownership Structure Pre- and Post-IPOs and the Operating Performance of JASDAQ Companies," *Pacific-Basin Finance Journal*, 10, 163-181.
- Lee, P. M. and Wahal, S. (2004) "Grandstanding, Certification and the

- Underpricing of Venture Capital Backed IPOs,” *Journal of Financial Economics*, 73, 375-407.
- Lerner, J. (1994) “The Syndication of Venture Capital Investments,” *Financial Management*, 23, 16-27.
- Li, X., and Masulis, R. W. (2005) “Venture Capital Investments by IPO Underwriters: Certification or Conflict of Interest?,” Working Paper, Vanderbilt University.
- Ljungqvist, A. (2007) “IPO Underpricing,” Eckbo, B. E. ed., *Handbook of Corporate Finance*, Elsevier, 375-422.
- Lockett, A. and Wright, M. (2003) “The Structure and Management of Alliances: Syndication in the Venture Capital Industry,” *Journal of Management Studies*, 40, 2073-2102.
- Liu, C. (1997) “ML Estimation of the Multivariate t Distribution and the EM Algorithm,” *Journal of Multivariate Analysis*, 63, 296-312.
- Liu, C. and Rubin, D. B. (1995) “ML Estimation of the t Distribution Using EM and its Extensions, ECM and ECME,” *Statistica Sinica*, 5, 19-39.
- Lystig, T. C. and Hughes, J. P. (2012) “Exact Computation of the Observed Information Matrix for Hidden Markov Models,” *Journal of Computational and Graphical Statistics*, 11, 678-689.
- Meggison, W. L. and Weiss, K. A. (1991) “Venture Capitalist Certification in Initial Public Offerings,” *Journal of Finance*, 46, 879-903.
- Meng, X. L., and Rubin, D. B. (1991) “Using EM to Obtain Asymptotic Variance-Covariance Matrices: The SEM Algorithm,” *Journal of the American Statistical Association*, 86, 899-909
- Myers, S. C. (1977) “Determinants of Corporate Borrowing,” *Journal of Financial Economics*, 5, 147-175.

- Rock, K. (1986) “Why New Issues Are Underpriced,” *Journal of Financial Economics*, 15, 187-212.
- Yung, C. and Gönül, Ç. and Wang, W. (2008) “Cycles in the IPO Market,” *Journal of Financial Economics*, 89, 112-208.
- Tinic, Seha M. (1988) “Anatomy of Initial Public Offerings of Common Stock,” *The Journal of Finance*, 43, No.4, 789-822.
- Welch, I. (1989) “Seasoned Offerings, Imitation Costs, and the Underpricing of Initial Public Offerings,” *The Journal of Finance*, 44, No.2, 421-449.
- Zucchini, Walter and MacDonald, Iain L. (2009) *Hidden Markov Models for Time Series An Introduction Using R*, A Chapman and Hall Book.