

## 企業業績予測システムの構築 動的線形モデルによる業績予測の可能性に関する研究

深 澤 弘 美

### A study of the ability of the dynamic linear models to predict corporate performance

Hiromi Fukasawa

本研究では、ベイズ統計学の立場で企業評価という問題を考える。これまでの財務指標を用いた分析の多くは、非ベイズ統計学の手法を適用している。また、業績変化の過程について統計的手法を用いて分析することが不十分である。そこで本論文では、ベイズ理論に基づいて企業業績の時系列分析を考え、統計的モデルを具体的に構築し、そのモデルを活用した予測システムを提案する。このシステムは、「動的線形モデル」を用いて時系列財務データの動きを捉えて将来の業績を予測し、業績優良もしくは不良企業に判別されるであろう確率を提示するものである。多面的かつ的確な予測が可能になれば、そのような企業評価は様々な場面での意思決定の大きな助けとなると考える。適切な統計的手法を用いた定量的な分析は、外部から企業を評価する1つの方法として活用され得るであろう。

This study concerns bayesian forecasting and predicting corporate performance. A new predicting system based on bayesian forecasting is proposed. The past analyses of financial ratio which take non-bayesian approach have not shown the detailed evaluation of corporate performance process on stochastic approach. In order to find a solution to this kind of problem, a system has been developed to forecast the financial time series data appropriately by the dynamic linear models. The dynamic linear models based on bayesian forecasting are easy to interpret and has significant flexibility. If we can predict corporate performance accurately by using financial data, it will support every place of decision making.

ベイズ統計学、時系列分析、財務指標分析、業績予測、倒産予測

Bayes, Time series, Financial ratio, Forecasting, Bankruptcy

本論文は、多摩大学大学院博士課程論文「企業業績予測システムの構築」を要約したものである。

(原稿受領日 1999. 10. 19)

### はじめに

本研究では、ベイズ統計学の立場で企業評価という問題を考える。ベイズ統計学の理論（以下単にベイズ理論という）に基づいて企業業績の時系列分析を考え、統計的モデルを具体的に設定し、そのモデルを活用した予測システムを構築する。経営情報を分析する際にデータを時系列に分析することが必要であることはいうま

でもないが、経営情報、特に会計の領域において時系列理論を適用した研究はまだ少ない。従来の企業評価は、単年度のデータを用いた分析が主である。変化の傾向を考慮するために、売上高増加率や営業利益増加率などの指標を用いた分析も行われてはいるが、対前年比もしくは直近3年程度の平均変化率の分析が多い。業績変化の過程を、統計的手法を用いて分析することが不十分である。また、各指標の変化率を含

む財務指標の分析に際して、統計的手法が用いられてはいるが、そのほとんどが非ベイズ統計学に基づくものである。非ベイズ統計学的手法を用いて、ある企業が「倒産する」か「倒産しない」かを判別するモデルが開発されてはいるが、このモデルには非現実的な条件が存在する。判別関数が説明変数の線形関数であることや、「倒産する」群と「倒産しない」群との分散が等しいことなどの前提条件は、現実のデータでは満たされないことから不十分な分析といわなければならない。これまでの研究の問題点である1)時系列分析が不十分であること、2)現実のデータに適合しない非ベイズ統計学による分析手法が適用されていること、以上2点を解決するためには、企業評価に適した新しいモデルを構築し、ベイズ理論に基づいた時系列分析を考えることが必要である。動的線形モデル[ West and Harrison (1997)] は、ベイズ理論に基づいた時系列分析の手法であり、企業の振る舞いを多面的かつ動的に分析・予測することが可能である。財務分析や倒産分析などの企業評価にも適用できるであろう。本研究では、この動的線形モデルを用いて業績を予測し、企業を評価する可能性を検討する。

ベイズ理論に基づく統計的手法により、年次、四半期、月次などの等間隔に得られる時系列財務データの動きを、できる限り忠実に表わすモデルを構築することができるならば、観測されたデータを作り出す母集団の把握およびより精度の高い予測が可能になると考える。景気の低迷が続き業績を悪化させる企業が少なくないどころか、金融機関までが破綻するという現在の経済状況において、企業を適切に分析・評価・予測する必要性は高い。容易に入手できる財務データを使って、多面的に企業の業績を予測することができれば、そのような企業評価はさまざまな場面での意思決定の大きな助けとなるで

あろう。もちろん定量的な情報である財務データの分析だけでは、十分な分析が可能とはいえない面もある。その限界と問題点を十分に考慮しなければならないが、適切な統計的手法を用いた定量的な分析は、外部から企業を評価する1つの方法として活用され得るであろう。

## 企業業績予測システムの概要

企業業績予測システムは、過去の業績優良企業および業績不良企業の情報をもとに現在あるいは将来の業績を評価・予測することを目的としている。次章に示す基準にもとづき、本研究で業績評価に有効と認定した5つの財務指標を用いて、動的線形モデル[ West and Harrison (1997)] を構築し、各指標の推移およびその予測を提示するシステムである。また最終的な評価値として、1994年の業績優良企業および業績不良企業の情報をもとに、1年後(図1の実行例では1995年)、3年後(図1の実行例では1998年)、6年後(図1の実行例では2001年)に当該企業が業績優良企業群に判別される確率を提示する。図1に示すグラフの\*印は観測値、実線は予測値を示している。またこのシステムは、日経財務データNEEDSに収録されている上場企業のうち製造業のみを対象にしている。1995年までのデータとしては1017社、1998年までのデータとしては1006社が実行可能である。製造業のみを対象とした理由は、非製造業の中でも特に小売業や建設業、農水産業では財務構造が製造業のそれと大きく異なるためである。

評価の対象とする財務指標は以下の5指標である。

売上高(対数)	Sales(Log)
総資本経常利益率	Return on Total Asset
自己資本比率	Equity Ratio
一人当たり人件費	Labor Costs per a head

キャッシュフロー対総負債比率

Cash Flow on Total Debt

プログラムはMATLAB( Version5.1 )で開発さ

れている。MATLABは高度な数値計算、視覚化、そしてプログラミングが扱える科学技術計算用の言語である。以下に実行例を示す。

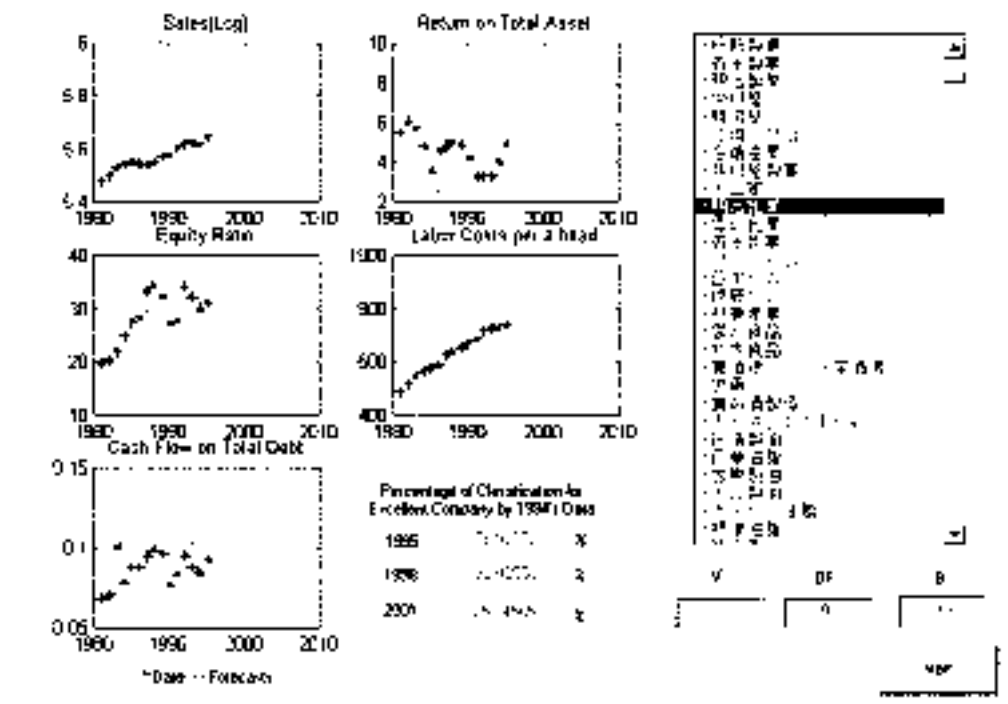


図1 企業業績予測システムの実行例

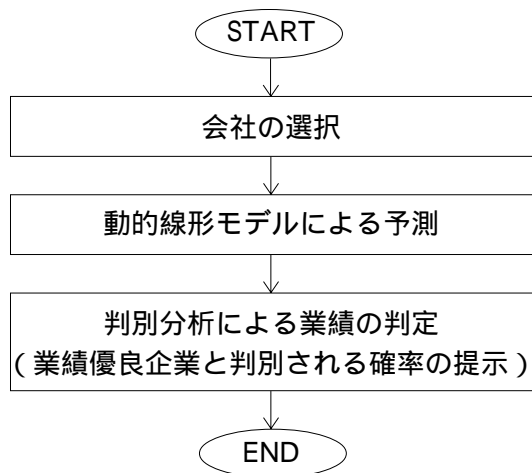


図2 システムの流れ

## 2.1 指標の選定基準

本システムでは、大矢知、金川、深澤( 1996 )で標準的財務指標として定めた10指標と、最近の倒産分析に関する研究においてその重要性が主張されているキャッシュフロー対総負債比率の計11指標の中から、もっとも判別効率の良い指標5つを選定した。表1に本システムで選定の候補とした11指標の業績優良企業群( 製造業上位300社 )と業績不良企業群( 製造業下位300社 )の平均値を示す。

表1 候補とした11指標の業績優良企業群(製造業上位300社)と業績不良企業群(製造業下位300社)の平均値

	業績優良企業群 (製造業上位300 社の平均値)	業績不良企業群 (製造業下位300 社の平均値)	
	1994	1994	
① 売上高(対数)	4.64	4.40	log(売上高)
② 売上高営業利益率(償却前)	14.01	-0.87	(営業利益+減価償却実施額)/売上高
③ 総資本経常利益率(利払前)	8.25	-3.88	(経常利益-償却不足額+支払利息+割引料+社債発行費+差金償却-受取利息)/《総資本》
④ 売上高増加率(3年平均)	0.84	-6.77	3年間各年度の売上高増加率の単純平均
⑤ 営業利益増加率(3年平均)	-20.31	-173.90	3年間各年度の営業利益増加率の単純平均
⑥ 経常収支比率	117.01	102.55	経常収入/経常支出
⑦ 自己資本比率	53.49	32.45	自己資本/総資本
⑧ 1人当り営業利益	380.20	-109.37	営業利益《従業員数》
⑨ 1人当り人件費	713.97	667.34	人件費《従業員数》
⑩ 1株当り純資産	1186.92	479.24	《自己資本》/《発行済株式数》
⑪ キャッシュフロー対総負債比率	17.47	-0.24	当期利益+減価償却実施額-配当金-役員賞与/総負債

金額の単位は百万円、《》は期中平均値

## 【選定基準】

1994年度売上高営業利益率(償却前)の上位300社および下位300社(製造業に限る)をもっとも効率よく判別する指標であること  
業績優良企業群および業績不良企業群のどちらかに偏ることなく平均的に効率よく判別できること

第4章に示す判別分析を適用した結果を比較することによって、効率の良い変数の組み合わせを選定する。ベイズ理論を適用することにより、「業績優良企業」「業績不良企業」の各群に判別される確率を算出できる。売上高営業利益率の上位300社(製造業に限る)が、各々業績優良企業群に判別される確率を算出し、50%を超えれば正しく判別されたと判断する。同様に売上高営業利益率の下位300社(製造業に限る)が業績不良企業群に判別される確率を算出し、50%を超えれば正しく判別されたと判断する。「業績優良企業を業績不良企業と誤って判別する確率」と「業績不良企業を業績優良企業と誤って判別する確率」(誤判別率)の値が低い方が判別効率が良いことを示す。また、両群の誤判別率に大きな差がある指標は採用しないものとする。

## 2.2 動的線形モデル[West and Harrison(1989, 1997)]の概要

「企業業績予測システム」では、柔軟で解釈しやすいモデルとしてベイズ理論を基礎とした動的線型モデル[West and Harrison(1997)]をもとに時系列データの予測を考える。

## 2.2 (1) 一次多項式モデル

動的線型モデルの最も基本的なモデルである1次多項式モデルは、観測値 $Y_t$ および状態変数 $\mu_t$ を次のように定義する。

## 【一次多項式モデル】

observation equation:

$$Y_t = \mu_t + v_t \quad v_t \sim N[0, V_t] \quad (2.1)$$

system equation:

$$\mu_t = \mu_{t-1} + \omega_t \quad \omega_t \sim N[0, W_t] \quad (2.2)$$

initial information:

$$(\mu_0 | D_0) \sim N[m_0, C_0] \quad (2.3)$$

 $Y_t$  : 観測値 ( $Y_t \in \mathbf{R}$ ) $\mu_t$  : 時刻  $t$  におけるシステムの状態 ( $\mu_t \in \mathbf{R}$ ) $v_t$  : 観測誤差 ( $v_t \in \mathbf{R}, V_t \in \mathbf{R}$ ) $\omega_t$  : システム誤差 ( $\omega_t \in \mathbf{R}, W_t \in \mathbf{R}$ ) $D_t$  : 時刻  $t$  までの情報  $\{D_0, Y_1, Y_2, \dots, Y_t\}$

$m_0$  : 状態変数の初期情報の平均

$C_0$  : 状態変数の初期情報の分散 ( $C_0 \in \mathbf{R}$ )

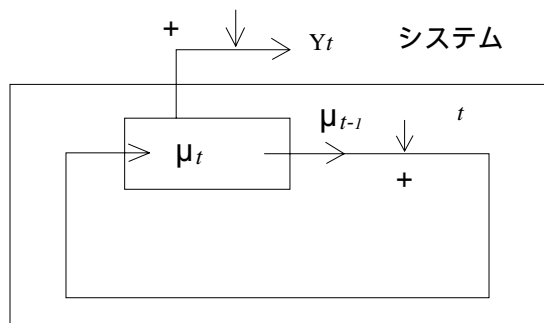


図3 動的線形モデルの流れ

ここでは、観測誤差  $\{v_1, v_2, \dots\}$  およびシステム誤差  $\{\omega_1, \omega_2, \dots\}$  はすべての時点において正規分布にしたがい、時点および互いに独立である。また  $(\mu_0 | D_0)$  から独立である。さらに誤差分散は各時点において既知であると仮定する。

information:

時刻  $t-1$  における  $\mu_{t-1}$  の事後分布

$$(\mu_{t-1} | D_{t-1}) \sim N[m_{t-1}, C_{t-1}] \quad (2.4)$$

時刻  $t-1$  における  $\mu_t$  の事前分布 ( $Y_t$  観測前)

$$(\mu_t | D_{t-1}) \sim N[m_{t-1}, R_t] \quad (2.5)$$

$$R_t = C_{t-1} + W_t$$

forecast:

1 時点先の予測 (時刻  $t-1$ 、 $Y_t$  観測前)

$$(Y_t | D_{t-1}) \sim N[f_t, Q_t] \quad (2.6)$$

$$f_t = m_{t-1} \quad Q_t = R_t + V_t$$

updating recurrence relationships:

時刻  $t$  における  $\mu_t$  の事後分布

$$(\mu_t | D_t) \sim N[m_t, C_t] \quad (2.7)$$

$$m_t = m_{t-1} + A_t e_t \quad C_t = A_t V_t$$

$$A_t = \frac{R_t}{Q_t} = \frac{R_t}{R_t + V_t} \quad e_t = Y_t - f_t$$

ここで、 $e_t$  は 1 時点先の予測の誤差を示し、観測値と予測値の差である。 $A_t$  はモデルの当てはまり具合を示す値 (適応係数) で、 $0 \sim 1$  の値をとる ( $0 < A_t < 1$ )。適応係数が 0 に近い値を

とれば事前分布のもつ情報は観測値のもつ情報に比べて大きいことを示し ( $R_t \ll V_t$ )、1 に近い場合は事前分布のもつ情報は観測値のもつ情報に比べて極めて小さいことを示す ( $R_t \gg V_t$ )。

さらに実際に予測を行う際には、 $k$  時点先の予測が必要となる。時刻  $t$  で得られる情報から時刻  $t+k$  の  $Y_{t+k}$  予測は次のように表される。

forecast distributions:  $k$  時点先の予測

$$(Y_{t+k} | D_t) \sim N[m_t, Q_t(k)] \quad (2.8)$$

$$Q_t(k) = C_t + \sum_{j=1}^k W_{t+j} + V_{t+k}$$

動的線形モデルは、インターベンション (分散の値を逐次修正) することができる。事前に入手できる情報を最大限活用してより正確な事後の予測を実現する、これがベイズ・アプローチ最大の利点である。また、分散が未知の場合にはディスカウントファクター ( $\delta$ ) を用いて観測値の分散  $W_t = C_{t-1}(1-\delta)/\delta$  を導き出すことが可能である。

## 【動的回帰モデル】

observation equation:

$$Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + v_t \quad v_t \sim N[0, V] \quad (2.9)$$

system equation:

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \omega_t \quad \omega_t \sim N[0, W_t]$$

$\mathbf{F}_t$  : 回帰行列 ( $n \times r$ )

$\mathbf{G}_t$  : システム行列 ( $n \times n$ )

$\mathbf{W}_t$  : システムの分散行列 ( $n \times n$ )

$\boldsymbol{\theta}_t$  : 状態ベクトル ( $n \times 1$ )

## 2.2 (2) 季節モデル

周期的なパターンを表現するために、さまざまな周期的な関数の線形結合が用いられるが、ここでは三角関数を用いて季節性を表現することを考える。サイン・コサインカーブを用いることにより、経済的で自然な表現が可能となる。季節効果モデルでは月次の変動を表現するには 12 の要素を考える必要があるのに対して、サイン・

コサインカーブで季節的な変動を表現することができれば、位相と振幅の2つの変数を決定すればよい。また、不規則な季節パターンにも対応が可能である。

周期的関数  $g(t)$ 、季節要素  $\psi_k = (\psi_0, \dots, \psi_{p-1})$  を三角関数の線形結合で表すと、周期が奇数の場合と偶数の場合で以下のように表せる。ここで、 $\alpha = 2\pi/p$  とする。

(1) 周期  $p$  が奇数の場合

$$\begin{aligned} & \text{整数 } q > 0, p = 2q - 1 \text{ に対して } p \text{ 実数個の } a_0, a_1, \dots, \\ & a_{q-1}; b_1, \dots, b_{q-1} \\ & \psi_j = a_0 + \sum_{r=1}^{q-1} [a_r \cos(\alpha r j) + b_r \sin(\alpha r j)] \quad (2.10) \\ & j = 0, \dots, p - 1 \end{aligned}$$

(2) 周期  $p$  が偶数の場合

$$\begin{aligned} & \text{整数 } q > 0, p = 2q \text{ に対して } p \text{ 実数個の } a_0, a_1, \dots, \\ & a_{q-1}; b_1, \dots, b_{q-1} \\ & \psi_j = a_0 + \sum_{r=1}^{q-1} [a_r \cos(\alpha r j) + b_r \sin(\alpha r j)] + a_q \cos(\pi j) \\ & j = 0, \dots, p - 1 \quad (2.11) \end{aligned}$$

(3) フーリエ係数  $a_r, b_r$

各周波数 ( $r = 1, \dots, q$ ) のフーリエ係数  $a_r, b_r$  は次式から求まる。

$$\begin{aligned} a_0 &= p^{-1} \sum_{j=0}^{p-1} \psi_j \\ a_p &= p^{-1} \sum_{j=0}^{p-1} \psi_j \cos(\pi j) \end{aligned} \quad (2.12)$$

周波数  $r = 1, \dots, q - 1$  のとき、

$$\begin{aligned} a_r &= 2p^{-1} \sum_{j=0}^{p-1} \psi_j \cos(\alpha r j) \\ b_r &= p^{-1} \sum_{j=0}^{p-1} \psi_j \sin(\alpha r j) \end{aligned} \quad (2.13)$$

(4) 各周波数の周期関数  $S_{rj}$

各周波数の周期関数の値  $S_{rj}$  は次式で表される。

$$\begin{aligned} S_{rj} &= a_r \cos(\alpha r j) + b_r \sin(\alpha r j), \quad r = 1, \dots, q - 1 \\ \text{and} \end{aligned} \quad (2.14)$$

$$S_{qj} = \begin{cases} a_q \cos(\pi r j) = a_q (-1)^j, & p = 2q \\ 0, & p = 2q - 1 \end{cases}$$

【全季節効果のフーリエ形式 DLM】

回帰行列  $F$  およびシステム行列  $G$  が、

$$F' = \begin{cases} (E'_2, \dots, E'_2), & p = 2q - 1, \text{ odd}; \\ (E'_2, \dots, E'_2, 1), & p = 2q, \text{ even}. \end{cases} \quad (2.15)$$

$$G' = \begin{cases} \text{block diag}[G_1, \dots, G_{q-1}], & p = 2q - 1, \text{ odd}; \\ \text{block diag}[G_1, \dots, G_{q-1}, -1], & p = 2q, \text{ even}; \end{cases}$$

$$G_r = J_2(1, \omega r) = \begin{bmatrix} \cos(\omega r) & \sin(\omega r) \\ -\sin(\omega r) & \cos(\omega r) \end{bmatrix}$$

$$r = 1, \dots, q - 1, \quad \omega = 2\pi/p$$

となるモデルを全季節効果のフーリエ形式 DLM と定義する。

ここで、予測関数は周期関数の和となる。

$$f_t(k) = \sum_{r=1}^q S_{rk} = \sum_{r=1}^q [a_{tr} \cos(\omega r k) + b_{tr} \sin(\omega r k)] + a_{tq} (-1)^k \quad (2.16)$$

ただし、周期が奇数の場合にはすべての時間でゼロとなる。状態ベクトルの事後分布の平均は、

$$E[\theta_t | D_t] = m_t = \begin{cases} (a_{t1}, b_{t1}; \dots; a_{t,q-1}, b_{t,q-1})', & p = 2q - 1; \\ (a_{t1}, b_{t1}; \dots; a_{t,q-1}, b_{t,q-1}; a_{t,q})', & p = 2q. \end{cases} \quad (2.17)$$

このとき  $\theta_t$  はフーリエ係数のベクトルで、その時点で算出された値である。

また、ディスカウントファクターはハーモニクごとに決めることができる。状態変数の事前分布の分散の第1項を、 $P_t = G C_{t-1} G'$ 、ハーモニク番号を、

$$v = \begin{cases} q - 1, & p = 2q - 1; \\ q, & p = 2q. \end{cases}$$

とおくと、各ハーモニクごとに  $P_{1t}, \dots, P_{vt}$  を定義できる。このとき観測値の分散は、

$$\begin{aligned} W_t &= \text{block diag}[P_{1t}(1 - \delta_1/\delta_1), \dots, P_{vt}(1 - \delta_v/\delta_v)] \\ 0 &< \delta_r \leq 1, \quad r = 1, \dots, v \end{aligned}$$

から求まる。

2.2 (3) 行列正規動的線形モデル

2.2(1)・2.2(2)節に示したモデルを検討した結果、本システムでは季節性を考慮しない多変量

動的線形モデル（行列正規動的線形モデル）を用いて5つの財務指標（年次）の予測を行うことが適当と判断した。動的線形モデルは、季節モデルを用いて月次データ等季節性のあるデータを分析することもできる。また、非線型モデルへの拡張も可能であり、あらゆるデータに対応できるモデルとして大変有効である。動的線形モデルに関連する研究としては、Kitagawa (1998)があげられる。この研究に示されているように、動的線形モデルを拡張し、非線型・非正規状態のための自己組織化状態空間モデルを構築することができる。

#### 【行列正規動的線形モデルの概要】

observation equation:

$$\mathbf{Y}_t = \mathbf{F}_t' \Theta_t + v_t \quad v_t \sim N[0, \mathbf{V}_t, \Sigma] \quad (2.18)$$

system equation:

$$\Theta_t = \mathbf{G}_t \Theta_{t-1} + \Omega_t \quad \Omega_t \sim N[0, \mathbf{W}_t, \Sigma]$$

$\mathbf{Y}_t$  : 観測値 ( $1 \times p$ )

$\Theta_t$  : 状態行列 ( $n \times q$ )

$v_t$  : 観測誤差 ( $1 \times q$ )

$\Omega_t$  : システム誤差 ( $n \times q$ )

$\mathbf{F}_t$  : 回帰行列 ( $n \times 1$ )

$\mathbf{G}_t$  : システム行列 ( $n \times n$ )

$\mathbf{V}_t$  : 観測誤差の分散行列 ( $q \times q$ )

$\Sigma$  : 誤差の分散行列 ( $q \times q$ )

$\mathbf{W}_t$  : システム誤差の分散行列 ( $n \times n$ )

$q$  : 変数の数

$n$  : モデルの次元

誤差はすべて時間そして系列間でも独立である。また、システム誤差のすべての要素の集合は、多変量正規分布にしたがっている。このような分布を行列正規分布と呼び、この分布の平均はゼロ行列、左分散行列  $\mathbf{W}_t$ 、右分散行列  $\Sigma$  である。密度関数は、

$$p(\Omega_t) = k(\mathbf{W}_t, \Sigma) \exp \left\{ -\frac{1}{2} \text{trace}[\Omega_t' \mathbf{W}_t^{-1} \Omega_t \Sigma^{-1}] \right\} \quad (2.19)$$

$$k(\mathbf{W}_t, \Sigma) = (2\pi)^{-qn/2} |\mathbf{W}_t|^{-q/2} |\Sigma|^{-n/2}$$

となり、 $\Sigma$  は逆ワイシャート分布に従う。

$$\Sigma \sim \mathbf{W}_n^{-1} [\mathbf{S}]$$

$$E[\Sigma^{-1}] = \mathbf{S}^{-1} \quad (2.20)$$

ここで、行列  $\mathbf{S}$  は  $\Sigma$  の推定値であり、このことから、

$$(\Theta_t | \Sigma) \sim N[\mathbf{m}, \mathbf{C}, \Sigma]$$

$$(\Theta_t, \Sigma) \sim N \mathbf{W}_n^{-1} [\mathbf{m}, \mathbf{C}, \mathbf{S}] \quad (2.21)$$

が導き出される。さらに、 $\Theta$  の周辺分布は行列  $T$  分布に従う。

$$\Theta \sim T_n [\mathbf{m}, \mathbf{C}, \mathbf{S}] \quad (2.22)$$

時刻 0 における事前分布

$$(\Theta_0, \Sigma | D_0) \sim N \mathbf{W}_{n_0}^{-1} [\mathbf{m}_0, \mathbf{C}_0, \mathbf{S}_0]$$

時刻 t-1 における事後分布

$$(\Theta_{t-1}, \Sigma | D_{t-1}) \sim N \mathbf{W}_{n_{t-1}}^{-1} [\mathbf{m}_{t-1}, \mathbf{C}_{t-1}, \mathbf{S}_{t-1}]$$

時刻 t における事前分布

$$(\Theta_t, \Sigma | D_{t-1}) \sim N \mathbf{W}_{n_{t-1}}^{-1} [\mathbf{a}_t, \mathbf{R}_t, \mathbf{S}_{t-1}]$$

$$\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1} \quad \mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t' + \mathbf{W}_t$$

1 時点先の予測  $(Y_t | \Sigma, D_{t-1}) \sim N[\mathbf{f}_t, Q_t, \Sigma]$

$$(Y_t | D_{t-1}) \sim T_{n_{t-1}}[\mathbf{f}_t, Q_t, \mathbf{S}_{t-1}]$$

$$\mathbf{f}_t = \mathbf{F}_t' \mathbf{a}_t \quad Q_t = \mathbf{F}_t' \mathbf{R}_t \mathbf{F}_t + V_t$$

時刻 t における事後分布

$$(\Theta_t, \Sigma | D_t) \sim N \mathbf{W}_{n_t}^{-1} [\mathbf{m}_t, \mathbf{C}_t, \mathbf{S}_{t-1}]$$

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t \mathbf{e}_t' \quad \mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t \mathbf{A}_t' Q_t$$

$$n_t = n_{t-1} + 1 \quad \mathbf{S}_t = n_t^{-1} [n_{t-1} \mathbf{S}_{t-1} + \mathbf{e}_t \mathbf{e}_t' / Q_t]$$

$$\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t Q_t^{-1} \quad \mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$$

k 時点先の予測  $(Y_{t+k} | D_t) \sim T_{n_t}[\mathbf{f}_t(k), Q_t(k), \mathbf{S}_t]$

$$\mathbf{f}_t(k) = [f_{t1}(k), \dots, f_{tq}(k)]'$$

$\mathbf{F}_t, \mathbf{G}_t, \mathbf{W}_t, \mathbf{R}_t, \mathbf{A}_t, Q_t$  は  $q$  変量に対して共通である。

さらに、観測値の分散に割引率  $B$  を導入することにより、

$$n_t = B n_{t-1} + 1 \quad (2.23)$$

$$\mathbf{S}_t = n_t^{-1} (B n_{t-1} \mathbf{S}_{t-1} + \mathbf{e}_t \mathbf{e}_t' / Q_t)$$

が時間で変化するモデルに拡張する。この際、事後分布は  $t$  の増加により退化することなく、

$n_t$  は有限の値  $(1-B)^{-1}$  に収束する。また、 $t$  が大きくなると事後分布の分散の予測値  $S_t$  は

$$S_t (1-\delta)^{-1} \sum_{r=0}^{t-1} B^r e_{t-r} e'_{t-r} / Q_{t-r} \quad (2.24)$$

となる。分散の予測値と時系列の共分散は過去の単なる分散と共分散の指数加重移動平均である。  $k$  時点までフィルタリングした分散は

$$S_t(-k)^{-1} = S_t^{-1} + B_t [S_t(-k+1)^{-1} - S_t^{-1}] \quad (2.25)$$

であり、さらには  $B=1$  のモデルを静的  $\Sigma$  モデルという。

「企業業績予測システム」は、割引率  $B$  をユーザが各データに適する値を指定してシミュレーション可能である。また、割引率  $B$  のほかに観測値の分散  $V$ 、システム誤差の分散の割引率  $DF$  も適宜指定可能である。

### 2.3 ベイズ理論による判別分析の概要

鈴木 (1987) に示されるように、ベイズ理論による判別分析とは、ある個体についての情報 (観測結果) に基づいて、その個体の帰属すべき母集団を判別する問題である。ある企業が「倒産する」か「倒産しない」かどちらかに判別・分類する場合にもベイズ理論に基づく判別分析を用いることができる。この手法により「倒産する」か「倒産しない」かの二者択一で業績を判断するだけではなく、それぞれに判別される確率を求めることができる。

対象となる 2 つのグループの母集団  $\pi_1, \pi_2$  について、問題とする任意の企業が  $\pi_i$  に属する事前確率を  $\xi(\pi_i)$  ( $i=1,2$ ) とする。

$$\xi(\pi_1) + \xi(\pi_2) = 1 \quad (3.1)$$

企業の財務データ  $x$  に対応する確率変数  $x$  は、その企業が母集団  $\pi_i$  に属するとき密度関数 (あるいは確率関数  $f(x|\pi_i)$ ) をもつとする。その企業が母集団  $\pi_i$  に属する事後確率は、財務データ  $x=x$  およびベイズの定理により以下の式で求めることができる。

$$\begin{aligned} \xi(\pi_i|x) &= \frac{\xi(\pi_i) f(x|\pi_i)}{\xi(\pi_1) f(x|\pi_1) + \xi(\pi_2) f(x|\pi_2)} \\ &\propto \xi(\pi_i) f(x|\pi_i) \quad (i=1,2) \end{aligned} \quad (3.2)$$

また、

$$\frac{\xi(\pi_1|x)}{\xi(\pi_2|x)} = \frac{\xi(\pi_1)}{\xi(\pi_2)} \cdot \frac{f(x|\pi_1)}{f(x|\pi_2)} \quad (3.3)$$

が成立つ。

各グループの母集団  $\pi_i$  が正規分布に従う場合、任意の企業が母集団  $\pi_i$  に属するならば、

$$\begin{aligned} \tilde{x}/\pi_i &\sim N(\mu_i, \Sigma_i) \\ \Sigma_i &: p \text{ 次正定符号} \quad (i=1,2) \end{aligned} \quad (3.4)$$

すなわち密度関数は、

$$\begin{aligned} f(x|\pi_i) &\equiv f(x|\mu_i, \Sigma_i) \\ &= (2\pi)^{-\frac{p}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp[-\frac{1}{2} (x-\mu_i) \Sigma_i^{-1} (x-\mu_i)'] \end{aligned} \quad (3.5)$$

で表せる。事後確率は、

$$\xi(\pi_i|x) = \frac{\xi(\pi_i) |\Sigma_i|^{-\frac{1}{2}} \exp(-\frac{1}{2} Q_i)}{\sum_{j=1}^2 \xi(\pi_j) |\Sigma_j|^{-\frac{1}{2}} \exp(-\frac{1}{2} Q_j)} \quad (i=1,2) \quad (3.6)$$

$$\text{ただし } Q_i = (x - \mu_i) \Sigma_i^{-1} (x - \mu_i)'$$

である。ここで、2 つのグループへの判別を考えると、

$$\frac{\xi(\pi_1|x)}{\xi(\pi_2|x)} = \frac{\xi(\pi_1)}{\xi(\pi_2)} \cdot \left( \frac{|\Sigma_2|}{|\Sigma_1|} \right)^{\frac{1}{2}} \exp(-\frac{1}{2} (Q_1 - Q_2)) \quad (3.7)$$

となる。

なお本研究では、この判別分析の手法をシステム構築の際と業績判定の際の 2 通りに活用している。

システム構築時には、本章第 2 節に示すように、判別分析を用いて、もっとも効率よく企業の業績を予測する変数 (財務指標) を選定した。業績優良企業群に判別される事前確率および業績不良企業群に判別される事前確率  $\xi(\pi_i)$  ( $i=1,2$ ) は、各 50% に設定し、業績優良企業群と業績不良企業群の分散行列  $\Sigma_i$  ( $i=1,2$ ) は、サンプル (1994 年度売上高営業利益率上位・下位) 各 300 社のデータより算出した。式 (3.7) に判別すべき企



業の観測値および事前確率、平均、共分散行列を代入することにより、事後確率（業績優良企業群に判別される確率）が求められる。

さらに業績予測システム実行時にも、各企業の業績を判定するために判別分析の手法を用いた。ここでは、各企業の観測値および予測値が、業績優良企業群に判別される確率を算出している。

## 結 語

本研究では、一貫してベイズ統計学の理論に基づいた手法を用い、企業評価という問題をモデル化し、実用的な予測システムを構築した。この研究により、企業が合併・買収を検討する場合や、投資家が投資価値判断を行う場合等のさまざまな意思決定の場面で企業を評価する際に、動的線形モデルおよびベイズ理論が有用であることが実証された。また本研究では、ベイズ統計学の理論を適用した動的線形モデルの有効性を実証すると同時に、このモデルと判別モデルを統合した予測システムが、企業評価に有益であることも検証できた。さらには、評価のもととなる財務指標の再検討や、動的線形モデルの動的非線形モデルへの拡張などの課題にとりくみ、本研究を発展させるという目標も得ることができた。本研究での成果を活かし、我々のシステムをより実用的で信頼性の高いシステムに改良するために、さらに研究を進めなければならない。今後の課題としては、以下の3点があげられる。

- 1) 財務指標の再検討
- 2) 業績優良企業と業績不良企業の判別のもととなるサンプルの検討
- 3) 動的非線形モデルへの拡張

財務指標については、「わが国上場企業の経営財務指標（採用ランク別）」「経営者が考えてい

る経営財務指標（重要度ランク順）」（ダイヤモンド・ハーバードビジネス：1997年8～9月号）に示される指標について検討する予定である。また判別のサンプルについては、業績優良と業績不良のサンプルの選択方法として、格付け結果や株価などの客観的な値を採用したモデルの検討が必要と考える。さらには動的線形モデルを拡張し、自己組織化状態空間モデル[Kitagawa (1998)]に示されるような非線型モデルについても研究を進める予定である。

論述を締めくくるにあたり、本研究を進めるにあたってご教示いただいた、多摩大学 鈴木雪夫教授、尾高敏樹教授、松谷泰行教授、青山学院大学 大矢知浩司教授、東北大学 河野大機教授に深く感謝申し上げます。研究の細部にわたり多大なるご指導ならびに大変有益なご示唆をいただきましたこと、厚く御礼申し上げます。また、本研究のいたるところでアイデアをいただいた多摩大学大学院の教職員ならびに学生の方々に記して感謝の意を表します。

## 参考文献

- Kitagawa, G.(1998), A self-organizing state-space model, *Journal of the American Statistical Association* 93(443), 1203-1215.
- West, M. and Harrison, J. (1989), *Bayesian Forecasting and Dynamics Models*. Springer.
- West, M. and Harrison, J.(1997), *Bayesian Forecasting and Dynamics Models* Second Edition, Springer.
- 大矢知浩司、金川一夫、深澤弘美(1996) 財務分析ツール・アンド・データ 白桃書房
- 鈴木雪夫(1987) 統計学 朝倉書店

## 著者プロフィール

1986年 青山学院理工学部経営工学科卒業  
1996年 多摩大学大学院修士課程修了  
1999年 同博士課程修了（博士：経営情報学）  
青山学院大学附属情報科学研究センター助手、同大学非常勤講師、多摩大学非常勤講師を経て、1997年青葉学園短期大学講師、現在に至る。