日本オペレーションズ・リサーチ学会和文論文誌
Transactions of the Operations Research Society of Japan
Vol. 57, 2014, pp. 92–111

企業格付判別のためのSVM手法の提案および逐次ロジットモデルとの比較による有効性検証

田中 克弘　　中川 秀敏
一橋大学大学院 国際企業戦略研究科

(受理 2012 年 11 月 27 日； 再受理 2014 年 9 月 16 日； 電子公開版 2014 年 11 月 21 日)

和文概要　本研究では，サポートベクターマシン (SVM) に基づく多群判別分析の手法を提案し，金融リスクマネジメントのための企業格付に同方法を応用する．信用格付による企業分類の判別的中率を改善することを期待して，マージン最大化と変数選択のために 0-1 整数変数を混合した最適化問題を導入して線形判別関数を推定する．提案した SVM 手法と最も広まっている統計モデルの一つである逐次ロジット・モデル手法との比較分析を通じて，提案手法がある程度有効性を有することを示す．

キーワード: リスク管理，信用リスク，多群判別分析，逐次ロジットモデル，サポートベクターマシン

1. 導入

本論文は，サポートベクターマシン（以下，「SVM」）を格付判別に対応できるよう拡張させ，逐次ロジットモデルと比較実証し，その有効性を有することを示したものである．格付は，格付毎に商品に応じて残高枠を設けること[1]，信用リスク計測のために内部格付という形で各企業の格付を付与すること，証券投資を行う上で企業の信用力を図る物差しとしても大いに判断材料として使われていることなどから金融リスクマネジメントの観点で必要なものとなっている．その一方で，格付を定量的に判別する最適なモデルが何かという点は未だ議論が続いている．

そういった中，例えば個別企業の倒産確率推計といった信用リスクを見る際に広く知られているモデルに，統計学の枠組みで構築し最尤推定法を用いて容易に解けることが知られているロジットモデルがある[2]．一方で，最近データマイニングの分野で注目を集めている，Vapnik[12] が提示した判別線と企業の信用力との距離に注目して二次（あるいは線形）計画問題で構築し，同じく容易に解けることが知られている SVM がある．ロジットモデルと SVM は，企業の信用力（以下，「信用スコア」）を線形で表現するという点で共通する一方，信用スコアに含まれるパラメータ推定を行う数理最適化モデルに違いがある．こうした背景から，倒産と非倒産といった二者択一のケースを対象に双方のモデルの判別精度を比較実証した例が Tony et al. [10]，Okada et al. [6] 等で報告されているが格付といった複数判別を対象にした実証例はあまり見かけない．このため，本論文では外部格付が 3 つあるケースを想定し，格付機関が付与した外部格付を各モデルがどの程度判別できているのかを比較実証することとした．

[1]例えば大和証券 [2] は「信用リスクが生じる取引については，事前に取引先の格付等に基づく与信枠を設定」としている．
[2]JCR[5] によると，JCR はロジットモデルを使用して個別企業のデフォルト確率を推定しているとある．

実証にあたってはモデルに次の二つの点を加えている．一つは変数選択で，信用スコアを説明する上で最適な財務指標の組み合わせの選択を，0-1 整数変数を用いることでモデルでまとめて行えるよう対応する．更に変数選択時に同程度の影響度を持つ指標が二つあるときに選択する指標を一つだけとするようにも対応した．これは信用スコアを説明する上で，違う性質の指標の組み合わせ同士で説明した方が，似た性質の指標の組み合わせ同士よりも，望ましいと考えられるからである[3]．もう一つは複数ある判別面を平行と考えないことである．これは，AAA と AA というどちらにしても安全圏という判別と，BBB と BB という投資適格となるか投資不適格となるかという瀬戸際の判別において，特定の財務指標が与える影響度が等しいと仮定することを不自然と考えたための対応である[4]．

モデルについては，SVM は Bugera[1] が逐次的に格付を判別する手法を提示し，更にそれを元に Saito and Konno[7] が日本市場のデータを用いて実証を行った逐次格付を判別していくモデルがあるため，それをベースとして扱う[5]．又比較対象のロジットモデルは安川 [15] が実証した逐次ロジットモデルをベースとしたモデルを扱う．このため双方とも逐次的に格付判別を行っていく枠組みをとっている．

本論文の構成は次の通りである．第二節では数式の定義，逐次ロジットモデルおよび SVM について，第三節では変数選択と扱うデータおよび実証方法について説明する．第四節では実証結果を報告し，最後に第五節では結論と今後の課題を述べる．

2. 定式化

2.1. 定義

この節では，逐次ロジットモデルおよび SVM の記述に共通する記号を定義する．

まず H 個の企業があり，それぞれが $J(\geq 3)$ 段階の外部格付を付与されている状況を考える．ここでは 1 が最も信用力が高い格付で，J が最も信用力が低い格付とする．次に，企業 $h \in \{1, \ldots, H\}$ は，K 種類の財務データがベクトル $\boldsymbol{x}^h = (x_1^h, \ldots, x_K^h)^{\mathrm{T}}$ で特徴付けられるとし，格付 $j \in \{1, \ldots, J\}$ に含まれる企業の集合を M_j と表すことにする．つまり $\{1, \ldots, H\} = M_1 \cup M_2 \cup \ldots \cup M_J$ かつ，$M_a \cap M_b = \emptyset \ (a \neq b)$ である．

さて企業ごとの信用力を，線形回帰式でモデル化した値を一般に信用スコアと呼び，ロジットモデルと SVM は共通してこれを財務指標を説明変数とした一次式でモデル化する．この信用スコアを表現するパラメータである係数ベクトルを $\boldsymbol{\beta}_i = (\beta_{i,1}, \ldots, \beta_{i,K})^{\mathrm{T}} \in \mathbf{R}^K$, $i \in \{1, \ldots, J-1\}$ として特徴づける．

2.2. 逐次ロジットモデル

ロジットモデルのフレームワークは，企業 h が実際に付与されている格付 M_j に属している確率をモデル化し，その確率を可能な限り高めようとする最尤推定法を用いて，パラメータ推定を行うものである．又，ロジットモデルを用いた複数判別には逐次ロジットモデルと順序ロジットモデルが知られているが，複数ある判別面を平行と仮定しない枠組みである逐次

[3]数理最適化の観点からも付け加えると，制約を強め実行可能な範囲を狭めることで計算時間の短縮も期待できる．

[4]安川 [15] では，判別面が全て平行という仮定について実証し，平行性を仮定するモデルは望ましくないと報告している．

[5]SVM については，Bugera[1] はペナルティ最小化のみを考慮しているが，Saito and Konno[7] は最小幾何マージン最大化およびペナルティ最小化を考慮している点に違いがある．

ロジットモデルを取り扱うこととした[6]．なお記述は木島・小守林 [4]，安川 [15] を参照にした．以下では，本研究で扱う逐次ロジットモデルの定式化を与え，判別方法をまとめておく．

まず判別に用いる閾値を $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{J-1}) \in \mathbf{R}^{J-1}$ とし，$\tau_{i-1} \geq \tau_i, i \in \{1, \ldots, J-1\}$ とする[7]．次に，h が M_j に属するか否かを判別するための信用スコアを $R_i(\boldsymbol{x}^h) = \boldsymbol{\beta}_i^{\mathrm{T}} \boldsymbol{x}^h$ と書く．ここで企業 h が格付 j である $(h \in M_j)$ 確率を p_{hj} と表すことにする．

最も代表的なロジットモデルである二項ロジットモデルを考えると，このモデルはある対象物が真か偽かという二者択一なケースに対して，どちらになるかを判別する際に用いられる．本論文では企業 h が格付 j 以下かそうでないかを判別するために用い，企業 h が格付 j 以下と判別する確率 P_j を次の通り書く．

$$P_j(x^h) = \frac{1}{1 + \exp(-(\boldsymbol{\beta}_i^{\mathrm{T}} \boldsymbol{x}^h + \tau_i))}, i \in \{1, \ldots, J-1\}, j \in \{1, \ldots, J\}. \tag{2.1}$$

逐次ロジットモデルは，この二項ロジットモデルによる判別を繰り返し用いる枠組みである．具体的には，まず全企業に対して二項ロジットモデルを用いて $R_1(\boldsymbol{x}^h)$ の係数パラメータ $\boldsymbol{\beta}_1$ と判別閾値 τ_1 を推定し，M_1 に属するか否かを判別する．次に M_1 に属しないと判別された企業に対して再び二項ロジットモデルを用いて $R_2(\boldsymbol{x}^h)$ の係数パラメータ $\boldsymbol{\beta}_2$ と判別閾値 τ_2 を推定し，M_2 に属するか否かを判別する．このようにして格付が3つ $(J = 3)$ のケースでの p_{hj} を表にまとめると，表1の通りである．

表 1: 外部格付が3つ存在するとしたときの各格付に判別される確率

	1回目	2回目	M_j に判別する確率 (p_{hj})
M_1	M_1 と判別する確率 (P_1)	—	P_1
M_2	M_1 と判別しない確率 $(1 - P_1)$	M_2 と判別する確率 (P_2)	$(1 - P_1)P_2$
M_3	M_1 と判別しない確率 $(1 - P_1)$	M_2 と判別しない確率 $(1 - P_2)$	$(1 - P_1)(1 - P_2)$

また一般的にはこの手順を $(J-1)$ 回繰り返すことで，p_{hj} は次のように書ける．

$$p_{hj} = \begin{cases} P_j, & \text{if } j = 1 \\ \left(\prod_{m=1}^{j-1} (1 - P_m) \right) P_j, & \text{if } j = 2, \ldots, J-1 \\ \prod_{m=1}^{j-1} (1 - P_m), & \text{if } j = J \end{cases}$$

なお，1回目に M_1 と M_2, M_3 を，2回目に M_2 と M_3 というように上から区別しているが，逆に1回目に M_1, M_2 と M_3 を，2回目に M_1 と M_2 というように下から区別するやり方も考えられる．それぞれの視点で構築したモデルの解は完全には一致しないが，格付判別に関し

[6]それ以外にも本論文での実証に関連して付け加えると，3. 1で述べる変数選択で0-1整数変数を含めるようモデルを拡張する際，逐次ロジットモデルは後述する近似方法を用いれば実行可能解を算出できるが，順序ロジットモデルはそれでも実行可能解が算出できない構造のモデルである．こういった実証上の取扱いやすさも考慮した．

[7]木島・小守林 [4] を見ると，逐次ロジットモデルでは閾値をいれない場合が一般的のようだが，安川 [15] が実証したモデルでは，閾値を入れている．閾値をいれない場合は全て0と設定したケースなので，双方のモデルの枠組みに違いがあるわけではない．

ては特にどちらがより適正かと言い切れるものでもないので，本論文では前者で述べたやり方で取り扱うこととした[8].

さてパラメータ $\boldsymbol{\beta}, \boldsymbol{\tau}$ の推定に最尤推定法を用いるため[9] 尤度関数 $L(\boldsymbol{\beta}_i, \boldsymbol{\tau}|\delta^{hj})$ を次のとおり書く.

$$L(\boldsymbol{\beta}_i, \boldsymbol{\tau}|\delta^{hj}) = \prod_{h=1}^{H}\prod_{j=1}^{J} p_{hj}^{\delta^{hj}} \tag{2.2}$$

なお δ^{hj} は次を示す.

$$\delta^{hj} = \begin{cases} 1, & \text{企業 } h \text{ が格付 } M_j \text{に属するとき} \\ 0, & \text{企業 } h \text{ が格付 } M_j \text{に属しないとき} \end{cases}$$

次に対数尤度関数 $lnL(\boldsymbol{\beta}_i, \boldsymbol{\tau}|\delta^{hj})$ を最大化する. 以後このモデルを $(Logit)$ と呼ぶ.

$$(Logit) \left| \begin{array}{ll} \max & lnL(\boldsymbol{\beta}_i, \boldsymbol{\tau}|\delta^{hj}) = \displaystyle\sum_{h=1}^{H}\sum_{j=1}^{J} \delta^{hj}\log(p_{hj}) \\ \text{s.t.} & \boldsymbol{\beta}_i, \boldsymbol{\tau}, i \in \{1, \ldots, J-1\}. \end{array} \right.$$

このモデルは (二項) ロジットモデルと同じ枠組みなので、標準的な数理最適化のソフトウェアを用いて解くことができる.

2.3. SVM

SVM のフレームワークは，企業の信用スコアと判別面との距離に注目して判別面を推定するものである.

まず，格付 $j \in \{1, \ldots, J\}$ に対して，格付数値が j 以下（信用スコアが格付 j と同等かより良い）か j より大きい（格付 j より信用スコアが劣る）かに注目して，集合 $M_U^i = \bigcup_{p=1}^{i} M_p$ および $M_L^i = \bigcup_{p=i+1}^{J} M_p, i \in \{1, \ldots, J-1\}$ を定義する．次に信用スコアの表記をするため $\beta_{i,0}$ を導入すると，信用スコアは $R_i(\boldsymbol{x}^h) = \beta_{i,0} + \boldsymbol{\beta}_i^{\mathrm{T}}\boldsymbol{x}^h$ と書ける．このとき企業 h の信用スコアと判別する超平面との絶対値換算した距離は $R_i(\boldsymbol{x}^h)/\|\boldsymbol{\beta}_i\|$ で表される[10].

まずここから二値判別によるモデルを整理するため，$i = 1$ としある企業 h が格付 1 より大きいか，あるいは 1 より小さいかというケースを考える．ここである企業の財務および格付データを用いれば，企業の格付が 1 より大きいかあるいは 1 以下かを完全に線形判別できる判別面を書けると仮定する．このとき企業 h の信用スコアと判別面との距離は $R_i(\boldsymbol{x}^h)/\|\boldsymbol{\beta}_i\|$ と書け，このとき最も短い距離を $1/\|\boldsymbol{\beta}_i\|$ と[11] 書く.

ここで完全に線形判別可能という前提から判別面は無数に書けるが，各企業と判別面の距離が最も離れた，いわば最も明確に分離できる判別面を書くのが望ましい．この観点から，企業の信用スコアと判別面との距離の最小値 $1/\|\boldsymbol{\beta}_i\|$(以下，最小幾何マージン) を可能な限

---

[8]ただしケースによってはいずれかのみを考慮したほうが良いと思われる．例えば，木島・小守林 [4] は，倒産状態，無配当状態，有配当状態を区別するケースを取り上げ，最初に倒産状態と存続状態に区別し，次に存続状態の企業を対象に無配当状態と有配当状態に区分けするといったモデル構築を行っている.

[9]逐次ロジットの尤度関数を，木島・小守林 [4] は言及していないが，安川 [15] は pp. 204 で言及している.

[10]$Q(Q = 1, 2, \ldots)$ ノルムの定義は次の通りである．$\|\boldsymbol{\beta}_i\|_Q = (\sum_{k=1}^{K} |\beta_k|^Q)^{1/Q}$

[11]分子の 1 の値はスケーリングの観点で任意に置いた値で，正の値であれば的中率に変化はない.

り大きくした判別面を推定する．これを踏まえて次の通り書いたモデルを，本論文ではハードマージンSVMと呼称する．

$$\left|\begin{array}{ll} \min & \|\boldsymbol{\beta}_1\|^2 \\ \text{s.t.} & \beta_{1,0}+\boldsymbol{\beta}_1^{\mathrm{T}}x^h \geq 1, h \in M_U^1, \\ & \beta_{1,0}+\boldsymbol{\beta}_1^{\mathrm{T}}x^h \leq -1, h \in M_L^1. \end{array}\right. \tag{2.3}$$

このモデルは凸型の二次関数の最小化問題なので，前提通り完全に線形判別可能なデータなら解ける．しかし，実務上データセットが常にその前提通りとは限らないため，正しく判別できない企業が含まれるデータを取り扱う前提でのモデル上の工夫が必要となる．

そこでスラック変数 $\xi_1^h \in \mathbf{R}^1, h \in \{1,\ldots,H\}$ を導入し，判別を誤った企業 h の信用スコアと判別面との距離を別途 $\xi_1^h/\|\boldsymbol{\beta}_1\|$ と書く．このときスラック変数に0以上と制約を付け，かつ最小幾何マージンの分子を1とおけば，正しい判別を行った企業のスラック変数は1以下で，誤った判別を行った企業のスラック変数は1より大きい値といえる．

これを踏まえ，最小幾何マージンの最大化と同時にスラック変数の合計値(以下，ペナルティ)の最小化も考慮した次のモデルを本論文ではソフトマージンSVMと呼称する．

$$\left|\begin{array}{ll} \min & \|\beta_1\|_2^2 + C\sum_{h=1}^{H}\xi_1^h \\ \text{s.t.} & \beta_{1,0}+\boldsymbol{\beta}_1^{\mathrm{T}}x^h + \xi^h \geq 1, h \in M_U^1, \\ & \beta_{1,0}+\boldsymbol{\beta}_1^{\mathrm{T}}x^h - \xi^h \leq -1, h \in M_L^1, \\ & \xi^h \geq 0, h \in \{1,\ldots,H\} \end{array}\right. \tag{2.4}$$

このモデルは二次計画問題の最小化問題なので，標準的な数理最適化のソフトウェアを用いて解くことができる．

ここで本論文ではこのソフトマージンSVMについて，先行研究を元に次二つの改良を行う．一つは第一項で2ノルムとしているところを1ノルムに変更することである．

$$\|\boldsymbol{\beta}_1\|_2^2 \Rightarrow \|\boldsymbol{\beta}_1\|_1^1. \tag{2.5}$$

これは計算速度の向上が目的で，二次計画問題を線形計画問題に書き直すことと同じ意味を持つ．また，的中率にこの項のノルム形式はそれほど大きく影響しないことが山下[14]で示唆されている[12]．

もう一つは第二項で各格付 M_j にあるサンプル数 m_j に応じ H/m_j を乗じることである．

$$\sum_{h=1}^{H}\xi_1^h \Rightarrow \frac{H}{m_U}\sum_{h\in M_U}\xi_1^h + \frac{H}{m_L}\sum_{h\in M_L}\xi_1^h. \tag{2.6}$$

これは格付毎のサンプル数の違いが推定するパラメータに影響を与える可能性を懸念したもので，サンプル数にばらつきがあると比較的サンプル数が多めの格付群の精度を優先することに伴い全体の的中率の悪化を防ぐことを目的としている．

次にここからは複数判別によるモデルを整理するため，$i \in \{1,\ldots,J-1\}$ とするケースを考える．具体的には，まず全企業に対してソフトマージンSVMを用いて $R_1(\boldsymbol{x}^h)$ の係数

[12] $Q=1$ (1ノルム) のとき線形計画問題に，$Q=2$ (2ノルム) のとき二次計画問題になる．

パラメータ $\boldsymbol{\beta}_1$ を推定し，M_1 に属するか否かを判別する．次に再びソフトマージンSVMを用いて $R_2(\boldsymbol{x}^h)$ の係数パラメータ $\boldsymbol{\beta}_2$ を推定し，M_2 以上に信用力が高い格付に属するか否かを判別する．この手順を $(J-1)$ 回繰り返すことでモデルを構築していく．なお格付が3つ $(J=3)$ のケースで図に表すと，図1の通りである．

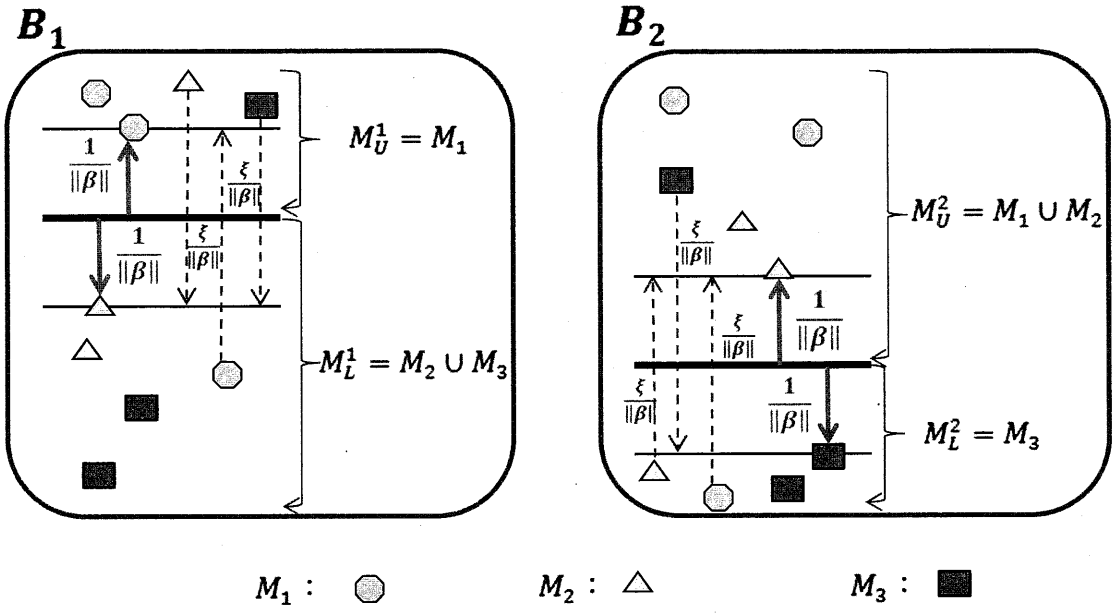

図 1: SVMによる外部格付が3つ存在するとしたときの格付判別のイメージ

このように一般的に格付を J 個判別したいケースであれば，$J-1$ 個判別面を書けばよい．そこでスラック変数を $\boldsymbol{\xi}^h = (\xi_1^h, \ldots, \xi_{J-1}^h) \in \mathbf{R}^{J-1}, h \in \{1, \ldots, H\}$ と再定義した上で，1つ1つの判別面を表現するために集合 B_i を次のように定義する．

$$B_i = \Big\{(\beta_{i,0}, \boldsymbol{\beta}_i, \xi_i^h) \Big| \begin{array}{l} R_i(\boldsymbol{x}^h) + \xi_i^h \geq 1 \text{ for } \forall h \in M_U^i; \\ R_i(\boldsymbol{x}^h) - \xi_i^h \leq -1 \text{ for } \forall h \in M_L^i; \xi_i^h \geq 0 \text{ for } \forall h \in \{1, \ldots, H\}\Big\}. \end{array} \tag{2.7}$$

更にBugera[1]に倣い，$R_{i+1}(\boldsymbol{x}^h) - R_i(\boldsymbol{x}^h) \geq 1, i \in \{1, \ldots, J-2\}$ とする制約式を追加する．この制約を付加することで，分析を行う範囲において，複数ある判別面が交差しないよう格付の逆転現象を防ぐことができる[13]．

なお，用いるモデルは判別面を平行としていないので，厳密にはパラメータ推定の企業データ(トレーニングデータ)でカバーできない範囲で交差する可能性はあるものの同じ標本集団の企業(テストデータ)の格付を判別するのであれば，この制約式を加えることは問題ないと考える．

このため，例えば本論文の分析対象である東証一部上場企業という標本集団に入る企業の中で，格付未発行の企業の格付を新たに推定するような状況であれば，パラメータ推定と格付を判別するデータは，共に東証一部上場と同じスクリーニングをくぐっている企業であり，その意味で同じ標本集団に属するといえ，十分に使用可能と考える．

[13]例を示すと，まずこの制約式をつけたモデルを解いたところ，ある企業 h が $R_1(x^h) = 1.5$ であれば M_1，また $R_2(x^h) = 2.5$ であれば M_1 もしくは M_2 に属することになり，二つの結果から M_1 に属すると判別できる．しかしこの制約式がないと $R_1(x^h) = 1.5$ で M_1，$R_2(x^h) = -2.5$ で M_3 というように，属する格付が判別できなくなる恐れがある．

逆に東証一部上場企業のデータを用いて推定した判別面を用いて，JASDAQ などに属するような新興企業の格付を推定するような状況であれば，パラメータ推定に用いるデータ (トレーニングデータ) と格付を判定したいデータ (テストデータ) が異なる性質を持つことが想定できるので，控えるべきであろう．

以上より本論文で取り扱う SVM のモデルは次の通りとし，以後 (SVM) と呼ぶ．

$$(SVM)\left|\begin{array}{ll}\max & \displaystyle\sum_{i=1}^{J-1}(\|\boldsymbol{\beta}_i\|_Q)^Q + C\sum_{j=1}^{J}\frac{H}{m_j}\sum_{h\in M_j}\sum_{i=1}^{J-1}\xi_i^h \\ \text{s.t.} & (\beta_{i,0},\boldsymbol{\beta_i},\boldsymbol{\xi}^h)\in B_i \text{ for } h\in\{1,\ldots,H\}, i\in\{1\ldots J-1\}, \\ & R_{i+1}(\boldsymbol{x}^h)-R_i(\boldsymbol{x}^h)\geq 1, i\in\{1,\ldots,J-2\}.\end{array}\right.$$

このモデルは線形計画問題もしくは二次計画問題の最小化問題なので，標準的な数理最適化のソフトウェアを用いて解くことができる．

3. 数値実証手順および前提条件

3.1. 変数選択

変数選択については，Yamamoto and Konno[13] で扱われている制約式を用いる．具体的には，モデルで任意に選択する財務指標の数を $S(\leq K)$，回帰係数の下限および上限を表す任意の数を $\underline{\beta}$,$\overline{\beta}$ [14]，0-1 整数変数 $z_k\in\{0,1\}$，$k\in\{1,\ldots,K\}$ を導入し，次のとおり書く．

$$\underline{\beta}z_k\leq\beta_k\leq\overline{\beta}z_k \text{ for } k\in\{1,\ldots,K\},\qquad \sum_{k=1}^{K}z_k=S. \tag{3.1}$$

この制約式の効果は次のとおりで，設定した S の数だけ変数を選択できる．

- $\underline{\beta}\leq\beta_k\leq\overline{\beta}$ if $z_k=1$(=ある財務指標 k を選択したとき)
- $\beta_k=0$ if $z_k=0$(=ある財務指標 k を選択しないとき)

更に信用スコアに対し，似た影響度をもつ指標が複数あればその中で一つだけを選択することを目的とした制約式も追加する．具体的には事前に計算した二つの指標の相関係数 (絶対値換算) が，任意に設定した許容値以上であれば，いずれか一つしか選択しないとする．具体的には説明変数 l と n の相関係数を Cor_{ln}，任意に設定した相関係数の許容値 ρ とすることで次の通り書ける．

$$z_l+z_n\leq 1, \text{ if } |Cor_{ln}|\geq\rho, l\neq n. \tag{3.2}$$

ただし (SVM) のような線形計画もしくは二次計画問題は 0-1 整数変数を導入してもある程度の時間で解けるが，$(Logit)$ のような非線形計画問題は 0-1 整数変数を導入して解を求めることは難しい．このため変数選択を行う場合に限り，目的関数をテイラー展開で二次近似して，整数変数を導入したモデルを用いることとする．

話を簡単にするため $J=3$ のときの逐次ロジットモデルを次の通り書く．なお $y_i=\boldsymbol{\beta}_i^T\boldsymbol{x}^h+\tau_i$ とおく．

$$\left|\begin{array}{ll}\max & \displaystyle\sum_{h=1}^{H}\{\delta^{h1}y_1^h+\delta^{h2}y_2^h-\log(1+\exp(y_1^h))-(\delta^{h2}+\delta^{h3})\log(1+\exp(y_2^h))\} \\ \text{s.t.} & \boldsymbol{\beta}_i,\boldsymbol{\tau}, i\in\{1,2\}.\end{array}\right. \tag{3.3}$$

[14]実証は下限=- 3，上限=+3 として解き，上限と下限に当たるパラメータは無かった．また財務データは基準化し- 3 から+3 の間になるよう操作している．

このとき $\log(1+\exp(y_i^h)), i \in \{1,\ldots,J-1\}$ という非線形関数をテイラー展開で二次近似すると，結果は次のとおりである．

$$\log(1+\exp(y_i^h)) \fallingdotseq \ln(2) + \frac{y_i^h}{2} + \frac{(y_i^h)^2}{8} \tag{3.4}$$

この近似式を式 (3.3) に当てはめれば (*Logit*) は凹型の二次計画最大化問題となり 0-1 整数変数を含むものの、ある程度の問題の大きさであれば標準的な数理最適化のソフトウェアを用いて解くことができる．

なお近似精度の見解を付け加えると，左辺の厳密な関数と右辺の近似した関数を図 2 にあるとおり比較したところ，当然テイラー展開による近似なので当然 $|y_i|$ の値が 0 に近い部分は十分に精度が良いことが確認でき，又絶対値換算で 3 を超えない範囲でもある程度の近似精度を保っていることがわかる．

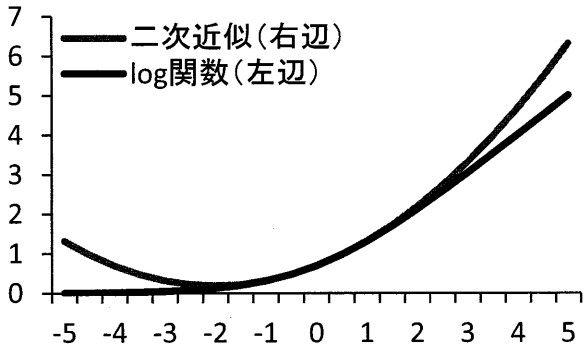

図 2: 逐次ロジットモデルの近似精度（横軸は y_i）

ここで y_i が絶対値換算で 3 を超えない範囲の乖離幅について，図 3 にある，式 (2.1) で示した格付 j 以下に判別される確率 P_j の分布をを見ると，y_i が絶対値換算で 3 を超えない範囲は大体 5% 以下あるいは 95% 以上である．5% 以下の部分では少々の乖離しても結果的に判別されないという判断に影響が出る可能性は低く，また 95% 以上の部分でも少々の乖離が出ても結果的に判別するという判断に大きな影響が出る可能性は低いと考えられる．一方 50%付近という判別に与える影響が大きい範囲は近似精度は十分である．

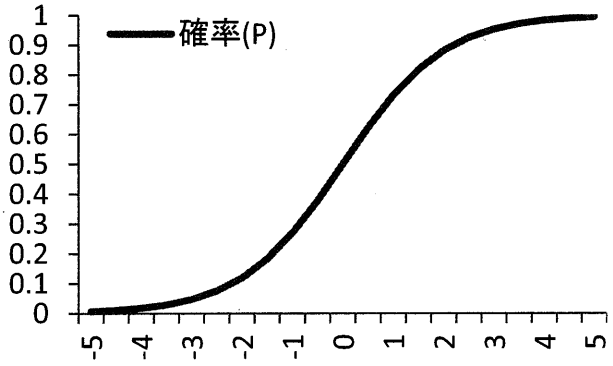

図 3: P_j の分布関数（横軸は y_i）

加えて繰り返しだが近似したモデルは変数選択に利用を留めることと，少々の近似は含むもののある程度の時間を掛ければ限定的な規模の問題であれば解くことができることから，当該近似方法を採用することとした[15].

よって逐次ロジットモデルの実証手順を整理すると，第一段階でテイラー展開による二次近似したモデルを解いて財務指標を選択した上で，第二段階でその財務指標を用いて厳密なモデルである (*Logit*) を解きパラメータを推定するものとする.

3.2. 実証環境と手順

使用データは2011年3月時点でのQuick Astra Managerより取得した東証一部上場企業の財務データおよびR&Iより取得した格付データの内，データが取得できた440社程度を用い，財務指標は(K=)60個である．計算機環境は、プロセッサがintel(R) Core(TM) i5 2.67Ghzでメモリが4GB．使用ソフトウェアはNUOPT(株式会社数理システム)のVer14を使用し，特に計算オプションを与えていない.

また財務データは財務項目毎に標準化し，-3〜+3の値に事前に変換している[16]．この作業はスケーリングの観点から，計算時間およびパラメータ推定の効率化を狙い実施した.

次に分析手順は次の通りである.

- 下図4のとおりデータプールをランダムに220社程度に2つに分け，トレーニングデータでパラメータを推定し，テストで評価する.

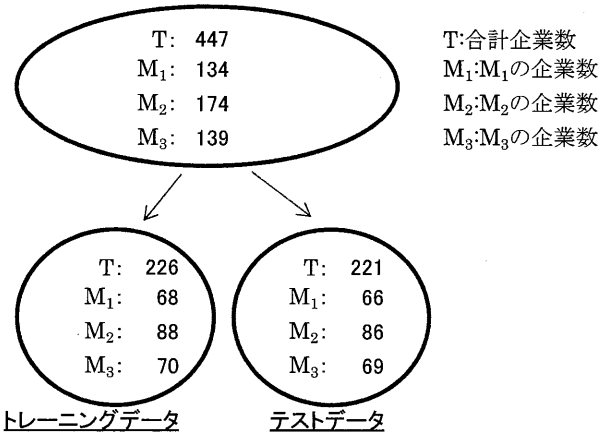

図 4: データ抽出

[15]なお更に言えば，(*Logit*) に整数変数を導入した，一般に混合整数非線形計画問題といわれる問題を解くことができる数理最適化のソフトウェアパッケージは自分の知る限り多くない．ただ，例えばNUOPTのglobalと呼ばれるアドオンパッケージを用いれば求解の可能性は0ではないが現実的な時間内で実行可能解を求めることは非常に厳しいと思われる．事実数理システム [8] では「数十変数程度の小規模でかつ複雑な問題に適しています」とあるし，手元でglobalのアドオンパッケージで試算したところ6時間計算しても局所解すら見つからなかった．ただ，近似モデルであれば混合整数二次計画問題となり，実用的な時間で解を求めることができる.

[16]なお-3より下の値は-3に．+3より大きい値は+3とした.

なお各格付の区分は次表の通り.

表 2: 各格付の区分

R&I の格付	ランク付け	トレーニングデータ	テストデータ
AAA, AA+, AA, AA-, A+	M_1	68	66
A, A-	M_2	88	86
BBB+, BBB, BBB-, BB+, BB	M_3	70	69
合計		226	221

- 再び上述の通り，ランダムにデータを分けて評価を行う. これを全部で 5 回実施した.

3.3. 評価方法

的中率の評価方法を整理する. 実際に格付毎の集合 M_1, M_2, M_3 に属する各企業 h において，係数パラメータベクトル $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ を用い，下表のケースが成立したとき的中したと判定する.

表 3: モデル別の的中率の評価一覧

	$(Logit)$		(SVM)	
ランク付け	β_1	β_2	β_1	β_2
M_1	$\beta_1^{\mathrm{T}} x^h > -\tau_1$	—	$R_1(x) > 0$	$R_2(x) > 0$
M_2	$-\tau_1 \geq \beta_1^{\mathrm{T}} x^h$	$\beta_2^{\mathrm{T}} x^h > -\tau_2$	$0 \geq R_1(x)$	$R_2(x) > 0$
M_3	$-\tau_1 \geq \beta_1^{\mathrm{T}} x^h$	$-\tau_2 \geq \beta_2^{\mathrm{T}} x^h$	$0 \geq R_1(x)$	$0 \geq R_2(x)$

4. 実証結果

この節では数値実証の結果を示す.

まず「4.1. 設定値」では下記の任意に定める設定値に注目し，実証する.

- (SVM) のみ
  - Q : (SVM) のノルムの形式
  - C : (SVM) の調整項
- 双方のモデルに共通
  - ρ : 財務指標同士の相関係数の許容値
  - S : 双方のモデルの変数選択における財務指標選択数

次に，「4.2. 個別実証」で格付毎の的中率と選択した財務指標を見ていく.

4.1. 設定値

まず (SVM) にかかる値の Q, C について，表 4 で $Q = 1$ のとき，表 5 では $Q = 2$ のときの C を 0.01, 0.02, 0.1, 0.5, 1 に変化させたとき結果を見る. なお $S = 5, \rho = 0.5$ とし，的中率および計算時間は試行した 5 回の平均値を示す.

また表の見方について補足すると，M_1，M_2，M_3 は各格付における的中率を指し，平均値は M_1，M_2，M_3 の的中率の平均値，標準偏差は M_1，M_2，M_3 の的中率の標準偏差を指す.

表 4: (SVM) の $Q=1$ のときの C 別の的中率 (単位:%)

調整項 (C)		0.01	0.02	0.1	0.5	1
A. トレーニング	M_1	60.88	80.00	81.76	82.35	83.53
	M_2	79.55	74.77	68.18	66.36	68.64
	M_3	61.43	66.86	78.29	81.14	82.00
	平均	67.29	73.88	76.08	76.62	78.06
	標準偏差	10.62	6.62	7.06	8.90	8.19
B. テスト	M_1	58.79	76.06	80.30	79.39	78.48
	M_2	76.28	73.72	63.95	61.40	62.33
	M_3	64.35	69.57	73.62	73.91	76.81
	平均	66.47	73.12	72.63	71.57	72.54
	標準偏差	8.94	3.29	8.22	9.23	8.89
A-B	M_1	2.09	3.94	1.46	2.96	5.04
	M_2	3.27	1.05	4.23	4.97	6.31
	M_3	-2.92	-2.71	4.66	7.23	5.19
	平均	0.81	0.76	3.45	5.05	5.51

表 5: (SVM) の $Q=2$ のときの C 別の的中率 (単位:%)

調整項 (C)		0.01	0.02	0.1	0.5	1
A. トレーニング	M_1	64.71	74.71	81.47	81.76	82.65
	M_2	78.41	70.68	67.05	68.41	67.05
	M_3	68.29	72.00	78.86	80.86	81.14
	平均	70.47	72.46	75.79	77.01	76.95
	標準偏差	7.11	2.05	7.69	7.46	8.61
B. テスト	M_1	62.12	71.82	80.00	79.39	78.79
	M_2	75.35	69.77	63.95	61.63	61.63
	M_3	65.51	69.86	71.88	72.17	74.78
	平均	67.66	70.48	71.95	71.07	71.73
	標準偏差	6.87	1.16	8.02	8.93	8.98
A-B	M_1	2.58	2.89	1.47	2.37	3.86
	M_2	3.06	0.91	3.09	6.78	5.42
	M_3	2.78	2.14	6.97	8.68	6.36
	平均	2.81	1.98	3.85	5.95	5.21

まず的中率についてテストの結果をみると Q の違いに関係なく，C を変化させたとき 0.01 より 0.02 のほうが平均値は高くなっているが，0.02 より大きくしてもほぼ横ばいである．しかし 0.02 と，それより C を大きくさせたケースとの差異を格付毎に見てみると，全体的に M_1，M_3 は 4～6%程度増加し，M_2 は 7～10%程度減少していることがわかる．また C が 0.02 のとき格付毎の的中率のばらつきを示す標準偏差が一番小さいことがわかる．

このように C の変化によって結果が変化するのは，ソフトマージン SVM の構造の問題と考える．改めて C の持つ意味を整理すると，この値は最小幾何マージン ((SVM) の第一項) の最大化とペナルティ ((SVM) の第二項から C を除いた値) の最小化を同時に実施する際，どちらに比重を置くか考慮する値で，C を変化したときの各項の一例を次に示す．

表 6: (SVM) の各項の値の比較例

	0.01	0.02	0.1	0.5	1
(SVM) の第一項	2.22	3.62	6.19	10.45	13.93
(SVM) の第二項から C を除いた値	594.84	476.08	408.91	389.33	383.22
(SVM)	8.17	13.14	47.08	205.11	397.15

この結果から C を大きくすることは，モデルの中でペナルティの最小化に重きを置くことと同じ効果が働くことが推測できる．なぜなら目的関数の値においてペナルティの値の比重が大きくなるのでペナルティの値の最小化のほうが，最小幾何マージンの最大化より，目的関数の最小化に与える影響が大きいためである

このため，ペナルティの値を構成するスラック変数 $\boldsymbol{\xi}_i$ の推定を優先するといった，最小幾何マージンの最大化をあまり考慮しない最適化は，判別面を構成するパラメータ $\boldsymbol{\beta}_i$ の推定が甘くなるといえる．その結果が C を 0.02 より大きくしたときのテストデータの的中率の低下と，トレーニングデータとテストデータの的中率の乖離幅の増加につながっていると推察する．また C を大きくするとトレーニングデータの的中率のみが大きくなったのはペナルティ最小化が優先され，$\boldsymbol{\beta}_i$ の推定より，ペナルティの最小化を優先し，各スラック変数 $\boldsymbol{\xi}_i$ が小さくなったためと判断する．

このように実行可能解の算出には C に 0 より大きい値を入れる必要があるが，あまりに大きな値を入れると $\boldsymbol{\beta}_i$ において信頼性の高い結果が導出できない．この点を踏まえ検証する中で最適な値を判断する必要がある．

次に Q の違いについて詳しくみるため表 7 で目的関数値，計算時間，的中率の平均値を比較した．

表 7: $Q=1$ と $Q=2$ の比較一覧

調整項 (C)		0.01	0.02	0.1	0.5	1
目的関数	$Q=1$	7.58	12.21	41.88	175.64	339.75
	$Q=2$	6.18	10.78	41.62	179.22	345.37
計算時間	$Q=1$	6.36	11.74	61.97	560.93	672.23
(単位:秒)	$Q=2$	366.94	511.09	1653.30	2310.32	2568.59
テストの平均的中率	$Q=1$	66.47	73.12	72.63	71.57	72.54
(単位:%)	$Q=2$	67.66	70.48	71.95	71.07	71.73

まず目的関数に与える影響にQが与える影響は小さいことがわかる．違いは最小幾何マージンにかかる項が1ノルムか2ノルムかの違いによる程度のもので，これは単位のオーダーが変わるものではないことから特に違和感ある結果ではない．次に計算時間については同じCの値のとき$Q=1$のほうが$Q=2$よりも非常に短い．この結果は0-1整数変数を含んだ線形計画問題と二次計画問題では前者のほうが速く解けることが一般に知られていることから，特に違和感はない．最後にテストの平均的中率はCの値に関係なく，的中率に大きなずれはなく，むしろ1〜2%高いものとなっている．このように大きな乖離がみられなかったことは目的関数に大きな違いがみられなかったことも踏まえれば，特に違和感ない結果である．これらの結果から$Q=1$および$C=0.02$が最適と判断した．

さてここからは(SVM)と$(Logit)$の双方のモデルについて取り扱う設定値を変化させた結果を論じる．まずρの値を0.1〜1まで0.1刻みとしたときの各格付の的中率の平均値は下図5の通りである．この検証結果の狙いは，ρを低くするほど0-1整数変数にかかる制約が強く効き実行可能解が減る一方で，計算時間の削減が期待できることからどのρの値が解の精度と計算時間の観点から最適といえるのかを実証することである．また計算時間については図6の通りである．

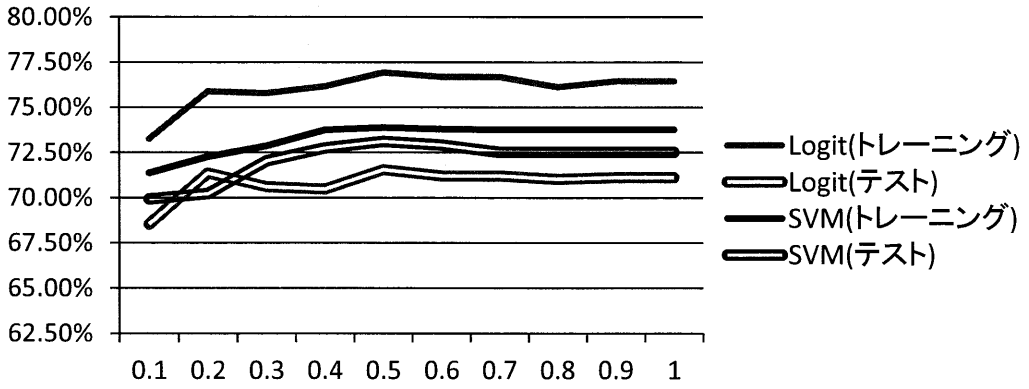

図 5: 各モデルのρ別の的中率

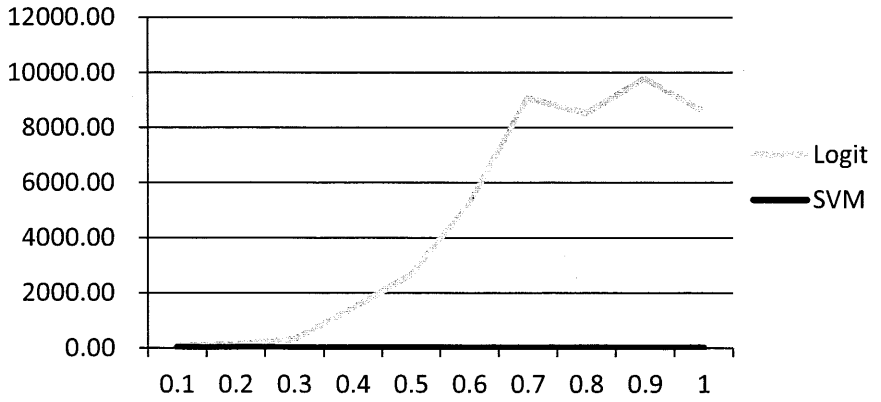

図 6: 各モデルのρ別の計算時間 (単位:秒)

双方のモデルとも0.1から0.4までは特にばらつきがあるが，$(Logit)$，(SVM)共に0.5以

降になると安定している．またテストの結果で一番高い的中率のケースが $\rho = 0.5$ であることを確認した．またトレーニングとテストの的中率の乖離は全体的に ρ を変えても大きな変化が起きなかったことも確認した．

また (SVM) は $C = 0.02$ であればどのケースでも 10 秒程度で解けていることを確認した．対して $(Logit)$ は ρ が 0.7〜1.0 で大体 8500 秒から 10000 秒程度と多くの時間を要している．なおテストの的中率が一番高かった $\rho = 0.5$ のときも 2700 秒程度と (SVM) が 10 秒程度に対して相当の乖離がみられた．これは (SVM) が線形計画問題，$(Logit)$ が (変数選択時は) 二次計画問題という問題の構造に加え，表 7 にもあるとおり C を小さくすると計算時間が短くなる (SVM) の特性から起きたものといえる．言い換えれば $(Logit)$ が大幅に時間を要するというよりは，(SVM) の問題の構造と与えた設定値 $(C = 0.02)$ であれば非常に速く解けるモデルだと判断できる．

これらを考慮し的中率と計算時間の観点から ρ は 0.5 が最適と判断できる．

最後に，S の値を 1〜15 まで 1 刻みとしたときの的中率は図 7 の通りである．また計算時間については図 8 の通りである．この検証結果で見たいのは，少ない財務変数で信用スコアを表現できるのが望ましいことから，どの S の値が解の精度と計算時間の観点から最適といえるのかを実証することが狙いである．

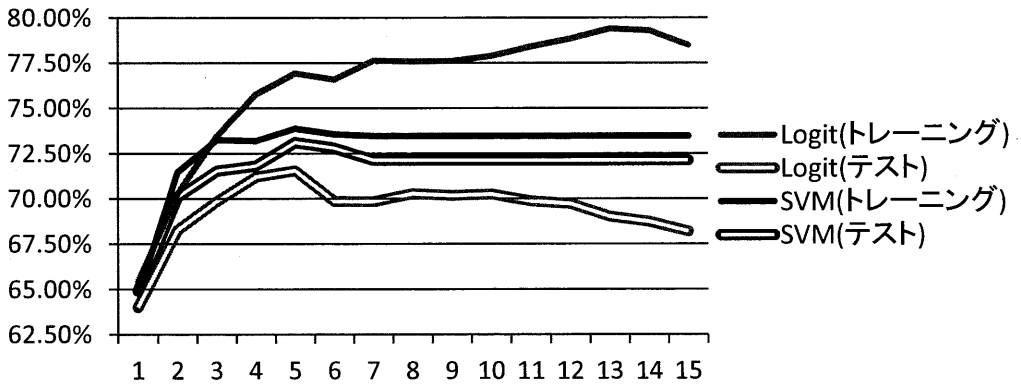

図 7: 各モデルの S 別的中率

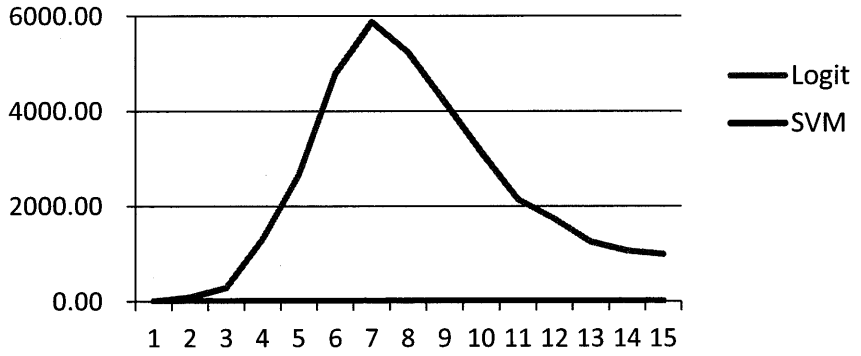

図 8: 各モデルの S 別の計算時間 (単位:秒)

テストの的中率に注目して見ると，Sが1から5の範囲のとき上昇し，特に5が一番高くなっている．またその後は下降もしくは一定となっている．

ここでモデル毎にもう少し見てみると，(SVM)はSを大きくしていくと7以降的中率が変化しなくなることがわかる．これは一般的に$Q=1$(1ノルム)の元で推定した行列は0になる値が多い（疎になる傾向をもつ）ことから，β_iの各値が0になりやすくなったものと思われる．また計算時間は10秒程度で解けている．

対して$(Logit)$は(SVM)と違ってS個必ず選択する傾向を持つことと，似た効果を持つ財務変数が二つあるときは一つしか選択しないとする制約を付与していることで，一定以上Sが大きくなるとあまり適切でない財務指標も選択してしまったと思われる．これが効いて財務変数が多くなると，トレーニングの的中率は増加できたもののテストの的中率は減少してしまったものと判断する．また合わせて計算時間がSが7となるまで上昇しているが，それ以降減少しているのも，制約が強く効いたためであろう．このため$\rho=0.5$のときは，Sが5程度の個数であれば問題はないが，多めに選択する際はρの値も十分に注意する必要がある[17]．もっとも，信用スコアはある程度絞った財務変数で表現するのが望ましいことを考えると，$S=5$は全く悪くない結果といえるだろう．一方で計算時間はSによって振れ幅があるが現実的な5個のとき2700秒，最大で7個のときに6000秒と，(SVM)と比較すると相当程度時間を要することがわかる．

これらを考慮し，的中率と計算時間の観点からSは5が最適と判断できる．

4.2. 個別実証

この節では前節で確認結果から$S=5, \rho=0.5, C=0.02, Q=1$としての結果を見る．

まず表8でモデル毎に的中率が何ノッチずれたかをテストの結果で見る．見方は縦が実際の格付で横がモデルで判別した格付である．例えば縦軸M_1→横軸M_2の24.24%は，M_1に属する企業の内M_2に判別した的中率を示す．また下線部が的中したケースで，それ以外は誤判別となったケースである．

表 8: テストにおける各モデルの的中率および誤判別率

	$(Logit)$			(SVM)		
実際 \ 判別	M_1	M_2	M_3	M_1	M_2	M_3
M_1	72.12%	24.24%	4.55%	76.06%	21.52%	2.42%
M_2	10.23%	70.93%	19.30%	9.77 %	73.72%	16.51%
M_3	2.32%	26.09%	71.59%	2.32 %	28.12%	69.57%

この結果から的中率はM_1, M_2は(SVM)のほうが高く，一方でM_3は$(Logit)$が高いことがわかる．ここでM_1であるものをM_3と誤判別することよりも，その逆のほうが将来的に懸念される損失は大きいと見込めるが[18]，誤判別率はM_3をM_2に誤ったケースは(SVM)

[17]ただし$(Logit)$で，$\rho=1$のケースでこの問題を解こうとすると，計算時間が非常にかかり，S=6から15については幾つかのケースを試したところ2日間(172800秒)計算しても最適解を導出できなかった．このため，この制約は計算時間の観点から非常に有益といえるので，ρの値への検討を踏まえることは十分に意味をもつといえる．

[18]例えばクレジットカードローンの審査者の立場を想定し，M_1を承認，M_2をさらなる検討が必要，M_3を否認と区分分けしたときを考える．本来M_1と判別するところをM_3と誤判別することは，承認できたものを否認することになるので，会社として将来に損失を生むことは無い．一方でM_3と判別するところをM_1と誤判別することは場合は本来否認となるものを承認したこととなり将来の損失が発生する恐れがある．

のほうがやや高いものの，M_1 にまで誤判別したケースに差はなかったので，大きな誤りは発生しなかったといえる．

次に試行した5回の内の一つのケースで係数の値とその有意性を見てみる．なお，本論文では変数選択をモデルの枠内で組み込んでいることから，算出した係数の結果はそのまま採用した．通常の線形回帰分析などではユーザーが何らかの視点を持って説明変数を選択し，その上で算出した結果を吟味して取捨選択を行うのが一般的であろうが，モデルの中でその取捨選択も合わせて行っているためそのような作業は省略している．その一方で，本論文ではモデルの違いによって選ばれる財務指標の選択や係数についての違いや信頼性が持てるかという点を確認するため，表9でこれを整理した．

表 9: モデル別の係数一覧例

	$(Logit)$		(SVM)	
財務指標	β_1	β_2	β_1	β_2
自己資本比率	0.5449	0.7907	0.0779	0.2103
自己資本対資本金比率	-0.0621	0.6128	0	0.0833
売上高当期利益率	1.6471	1.8803	0	0
総資産利益率	-0.5461	-1.0393	0	0
自己資本伸び率	0	0	-0.0709	-0.1121
1株当たり自己資本	0	0	0.5253	0.4550
期末時価総額(普通株ベース)	1.7960	1.0176	1.1249	1.0393

「自己資本比率」は，大きいほど企業の財務の健全性を示す指標であるので，双方のモデルで算出した係数が"+"であることから信頼性は高い。また「自己資本対資本金比率」も同様の効果を持つ指標であるが，$(Logit)$ の β_1 で値は大きくないものの，"- "となり少々信頼性は低い。ただ他の箇所は0以上となっていることから当該財務指標の採用に強い違和感を感じるものではないと判断する．

「売上高当期利益率」は、大きいほど企業の収益性が良いことを示す指標であるので，採用された $(Logit)$ の係数を見ると"+"であることから信頼性は高い．一方，「総資産利益率」も，同様の効果を持つ指標であるが"- "となっており，信頼性は低い．

「自己資本伸び率」は，一般に大きいほど企業の財務状況の健全性が良いことを示す指標であるが，採用された SVM の係数を見ると，値自体はそれほど大きくないものの"- "となっており少々信頼性は低い．一方，「1株当たり自己資本」も同様の効果を持つ指標で採用された (SVM) の係数を見ると，"+"となっており信頼性は高い．

「期末時価総額(普通株ベース)」は，大きいほど企業価値が高いことを示す指標であるので，双方のモデルで算出した係数が"+"であることから信頼性は高い．

総じて個別に見たときに一部の財務指標の採用に対して違和感はまったく無いわけではないが，全体感として双方のモデルとも概ね良好な指標を組み込んだといってよいと判断した．

また双方のモデルで選択した財務指標の上位を，選択した財務指標順に表10のとおり並べた．(データの試行回数5回の内，2回以上選択した指標をまとめている.)

回数に若干の差はあるが双方のモデルで選択されている指標がいくつかあることが確認できる一方で，次3点の違いも確認した．一つに双方のモデルで選択回数に大きな差がみら

表 10: 各モデルの選択した財務指標一覧

$(Logit)$		(SVM)	
財務指標	選択数	財務指標	選択数
自己資本対資本金比率	4	自己資本対資本金比率	5
期末時価総額 (普通株ベース)	4	1 株当たり自己資本	☆5
総資産利益率	●3	期末時価総額 (普通株ベース)	5
自己資本比率	2	自己資本比率	3
自己資本伸び率	2	自己資本伸び率	3
財務レバレッジ	2	財務レバレッジ	2
1 株当たり自己資本	☆2	---	-

れる☆でマークした「1 株当たり自己資本」で，(SVM) は 5 回に対し $(Logit)$ は 2 回であった．二つに片方である程度選ばれているが，もう片方では全く選ばれていない●でマークした「総資産利益率」で，$(Logit)$ は 3 回に対し (SVM) は 1 回も選択されていなかった．三つに (SVM) は 5 回選択された指標が 3 つあるが $(Logit)$ にそういった指標は無かった．

5. 結論と今後の課題

本論文では，逐次ロジットモデルと SVM について説明し，外部格付が 3 つとしたケースでの比較実証を行った．

その結果 SVM は逐次ロジットモデルよりも，概ね格付別および全体平均での的中率を上回る結果となった．この結果はモデルの枠組みの違いによるものと考えており，具体的にはロジットモデルは与えられたデータ情報と統計学の枠組みに基づいた（ロジスティック分布を前提とした）組み立てに対し，SVM は与えられたデータ情報のみを用いかつ特に分布の前提を必要とせずに判別線とサンプルの距離に注目して判別するといった違いが効いたものであろうと考えている．

又、財務指標の選択を行えるよう 0-1 整数変数を導入した上での計算時間を比較すると (SVM) のほうが $(Logit)$ より非常に速い時間で解けることが確認できた．この点は (SVM) の持っている実用的な側面といえるだろう．

なお SVM の取り扱いで注意すべき点に，ユーザーが任意に設定すべき C の決め方がある．なぜならこの値は只の調整項のため，選択如何によって的中率や選択する指標が大きく異なる可能性がある．このため表 4 のような検証を十分に行った上で適切と思われる値を探索することが必要であるが，その一方で C の最適な値をどのように決めるかは今後の研究課題といえる [19]．

最後に個別企業の信用力を計量するモデルは，碓井 [11]，JCR[5] などからも推察するに，今はロジットモデルが最も広く扱われているモデルの 1 つで今後もこの分野の主流であろうが，今回取り上げた SVM はそれと同等以上の判別力をもっている．このため本論文の結果が個別企業の信用力測定に用いるモデル選択の参考などに少しでも役立てていただければ幸いである．

[19] 対応の一つとして SVM を書き換えたモデルで Gotoh[3] が E μ-SVM と呼ばれるモデルで，CVaR(conditional value at risk) 最小化問題の枠組みを用い，設定値を確率水準と意味のある値にに書き直している．また田中 [9] も E μ-SVM とは異なる視点で，CVaR の枠組みを導入し同じく，確率水準と取り扱うモデルを提示している．

A. 使用した財務指標一覧

用いた財務指標は次の通り 60 個である. これらの指標を用いたのは十分なデータ数が確保できたためである.

No.	財務指標	No.	財務指標
1	流動比率	31	自己資本伸び率
2	当座比率	32	総資産伸び率
3	固定比率	33	売上高伸び率
4	固定長期適合率	34	営業利益伸び率
5	自己資本比率	35	経常利益伸び率
6	負債比率	36	当期利益伸び率
7	有形固定資産比率	37	1 株当たり売上高
8	自己資本対資本金比率	38	1 株当たり経常利益
9	財務レバレッジ	39	1 株当たり利益
10	売上高営業利益率	40	1 株当たり利益 (企業発表)
11	売上高経常利益率	41	1 株当たり当期利益
12	売上高当期利益率	42	1 株当たり EBITDA
13	売上高事業利益率	43	1 株当たり配当金
14	売上高利払後事業利益率	44	1 株当たり自己資本
15	EBITDA マージン	45	1 株当たり CF
16	売上高減価償却費率	46	1 株当たり営業 CF
17	総資産営業利益率	47	CF マージン
18	総資産経常利益率	48	CF 対流動負債比率
19	総資産利益率	49	CF 対負債比率
20	自己資本経常利益率	50	CF 対営業利益比率
21	自己資本利益率 (ROE)	51	CF 対当期利益比率
22	資本金収益率	52	営業 CF マージン
23	企業利潤率	53	営業 CF 対流動負債比率
24	使用資本利益率 (ROCE)	54	営業 CF 対負債比率
25	総資産回転率	55	営業 CF 対自己資本比率
26	流動資産回転率	56	営業 CF 対設備投資比率
27	固定資産回転率	57	期末 EV/EBITDA
28	有形固定資産回転率	58	期末 EV
29	稼動有形固定資産経常利益率	59	期末株式時価
30	稼動有形固定資産純利益率	60	期末時価総額

参考文献

[1] V. Bugera, S. Uryasev, and G. Zrazhevsky: Classification Using Optimization: Application to Credit Ratings of Bonds. E.J. Konthoghiorghes, et al(Eds.), *Computational Methods in Financial Engineering*,(2008), 211–239.

[2] 大和証券: 平成24年3月末連結自己資本規制比率に関するお知らせ
www.daiwa-grp.jp/data/current/press-3185-attachment.pdf,(2012).

[3] J. Gotoh and A. Takeda: A linear classification model based on conditional geometric score. *Pacific Journal of Optimization*,**1**,(2005), 277–296.

[4] 木島正明, 小守林克哉: 信用リスク評価の数理モデル (朝倉書店,1999).

[5] 日本格付研究所: JCR大企業モデル（デフォルト率推定モデル）の概要
www.jcr.co.jp/coc/11-04-21.pdf,(2011).

[6] Y. Okada and H. Konno: Failure Discrimination by Semi-Definite Programming Using a Maximal Margin Ellipsoidal Surface., *J. of Computational Finance*, **12** (2009), 63–78.

[7] M. Saito and H. Konno: Classification of Companies Using Maximal Margin Ellipsoidal Surfaces., *Computational Optimization and Applications*, **55** (2013), 469–480.

[8] 数理システム: NUOPT/SIMPLEマニュアル,(2011).

[9] 田中克弘, 中川秀敏: CVaR最小化に基づくSVMによる格付判別モデルの提案と評価. *JAFEE 2012*冬季大会予稿集, (2012), 109–120.

[10] V.G. Tony, B. Bart, G. Joao, and V.D. Peter: A Support Vector Machine Approach to Credit Scoring., *MENDELEY*,**1** (2010), 73–82.

[11] 碓井 茂樹: 内部格付制度と信用リスク計量化
http://www.boj.or.jp/announcements/release_2012/data/rel120626a8.pdf, (2012).

[12] V.N. Vapnik: *The nature of statistical learning theory*, (Berlin,Springer-Verlag,1995).

[13] R. Yamamoto and H. Konno: Choosing the best Set of Variables in Regression Analysis using Integer Programming., *J. of Global Optimization*, **44** (2009), 273–282.

[14] 山下 浩, 田中 茂: サポートベクターマシンとその応用
www.msi.co.jp/vmstudio/materials/svm.pdf,(2001).

[15] 安川武彦: 平行性の仮定と格付データ：順序ロジットモデルと逐次ロジットモデルによる分析. 統計数理,**50-2** (2002), 201–215.

田中 克弘
E-mail： pwd5re227i@hotmail.co.jp

ABSTRACT

A METHOD OF CORPORATE CREDIT RATING CLASSIFICATION BASED ON SUPPORT VECTOR MACHINE AND ITS VALIDATION IN COMPARISON OF SEQUENTIAL LOGIT MODEL.

Katsuhiro TANAKA Hidetoshi NAKAGAWA
Hitotsubashi University
Graduate School of International Corporate Strategy.

In this study we present a method of multiple discriminant analysis based on support vector machine (SVM) and to apply it to corporate credit rating for financial risk management. In anticipation of improving accuracy ratio of classification of companies with estimated credit rating, we introduce a optimization problem to estimate linear discriminant functions mixed of margin maxitimization and 0-1 integer variables to choose the best set of variables. We show validity of the method to some extent through some comparative analyses between the proposed SVM method and a sequential logit model approach that is one of the most popular statistical models.