半正定値計画法を用いた中小企業の倒産予測

研究代表者 **今野　浩**（中央大学理工学部）
共同研究者 **鎌倉　稔成**（中央大学理工学部）
共同研究者 **渡辺　則生**（中央大学理工学部）
共同研究者 **川代　尚哉**（中央大学理工学部）
共同研究者 **越塚　知幸**（中央大学大学院理工学研究科博士後期課程）

1　はじめに

過年度の研究で，我々は財務データを用いた中小企業の倒産予測モデルを提案し，これによって従来より精度の高い予測が可能であることを示した。

このモデルは，従来実務上広く用いられてきたロジット・モデルの中の倒産強度関数を財務データの 1 次式から 2 次式に拡張した上で，等倒産強度面が楕円面もしくは放物面となるための条件，すなわち半正定値条件を追加したものである。

2 次の倒産強度関数を導入するメリットは，財務指標間の相互関係を記述できることである。ところが一般の 2 次式を採用すると，パラメータの自由度が大きすぎるためオーバー・フィッテングが起り，予測精度は線形モデルに比べて却って劣化することが知られている。そこで我々は，財務指標の性質を考慮した上で，倒産強度関数の等高面が楕円面または放物面となるように 2 次式のクラスを限定することによって，オーバー・フィッテングを回避することを考え，データを用いて判別精度が従来より数パーセント程度改良されることを示した。たとえ数パーセントの精度向上であっても，実務上その意味するところは極めて大きい。（なおこの方式は，すでに株式会社日本格付研究所において実装され，良い成果をあげている）

しかしこのモデルの場合，パラメータの半正定値条件の下で非線形な対数尤度関数を最大化しなくてはならないため，線形ロジット・モデルに比べて大量の計算が必要となるという欠点がある。過年度に我々が開発した手法を用いれば，数千企業から 1 万社程度の企業を対象とするモデルまで解くことができるが，数万企業を対象とする問題を解くにあたっては，より効率的な手法を考案することが必要となる。

そこで今年度は，超大型問題を解くための近似解法を提案し，それによって十分に最適解に近い解が得られることを示した。この方法は，取り扱っている関数の性質上，等倒産強度面は全空間で楕円面とならなくても，財務データがカバーする領域において近似的に楕円面を構成していればそれで十分である，という推論に基いてのものであるが，今回の数値実験によってその推論の正しさが裏付けられた。

なおこの方法はより大きな数万社レベルの問題を解くためには，一層の改良が必要である。

2　半正定値ロジット・モデル

n 種類の財務指標を用いて倒産予測を行うものとし，x_{ij} を第 i 企業の第 j 財務指標の値とする。過去の一定期間に倒産した企業の集合を M_1，倒産しなかった企業の集合を M_0 とする。我々の目的は，現在存続している企業の財務データを $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ としたとき，その企業が一定期間に倒産する確率 $f(\boldsymbol{x})$ をデータ・マイニング・アプローチによって求めることである。

ここで我々が採用するのは，2 次ロジット・モデル

$$f(\boldsymbol{x}) = \frac{exp\{Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x})\}}{1 + exp\{Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x})\}}$$

である。ここで倒産強度関数 $Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x})$ は

$$Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x}) = \alpha_0 + \sum_{j=1}^{n} \alpha_j x_j + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha_{jk} x_j x_k$$

で与えられる 2 次関数である。

2 次ロジット・モデルは，($\alpha_{kj} = \alpha_{jk}$ を仮定して）全体で $(n+1)(n+2)/2$ の未知パラメータを含んでいる。従って，$n = 10$ のとき，その数は 66 個に達する。このため，2 次ロジット・モデルを用いて最尤推定を行うと，データに対してモデルが過剰に反応するオーバー・フィッテングが起り，予測制度を下げる結果になる。

そこで我々はこの現象を避けるために，$\boldsymbol{A} = (\alpha_{jk} \in R^{n \times n})$ のクラスを半正定値行列の範囲に限ることによって，モデルの説明力を高めるとともに自由度を減らすことを考えた。

この結果導かれる対数尤度関数は

$$L(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}) = \Pi_{i \in M_1} \frac{exp\ \{Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x}_i)\}}{1 + exp\ \{Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x}_i)\}} \times \Pi_{i \in M_0} \frac{1}{1 + exp\ \{Z(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}, \boldsymbol{x}_i)\}}$$

とおいて

$$(P)\quad \left| \begin{array}{ll} 最大化 & lnL(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}) \\ 条件 & \boldsymbol{A}\ \text{は半正定値行列} \end{array} \right.$$

となる。幸い，この問題の目的関数は $(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A})$ の凹関数となるので，切除平面法による非線形関数最大化を適用することによって解くことができる。

3　半正定値ロジット・モデルの近似解法

良く知られているように $\boldsymbol{A}$ が非負定値であるためには，n 次元の単位球面 B_n 上の全てのベクトル $\boldsymbol{y}$ に対して，$\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} \geq 0$ となることが必要かつ十分である。従って，問題 (P) は次の半無限最適化問題：

$$(P')\quad \left| \begin{array}{ll} 最大化 & lnL(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}) \\ 条件 & \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} \geq 0\ \ {}^{\forall}\|\boldsymbol{y}\| \in B_n \end{array} \right.$$

と等価である。これを厳密に解くために，我々は半無限計画問題を解くための常套手段である緩和法を採用した。

緩和法

条件 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} \geq 0$ を外した制約なし最大化問題を解き，その最適解を $(\alpha_0^0, \boldsymbol{\alpha}^0, \boldsymbol{A}^0)$ とする。もし $\boldsymbol{A}^0$ が半正定値になっていれば，$(\alpha_0^0, \boldsymbol{\alpha}^0, \boldsymbol{A}^0)$ は問題 (P) の最適解である。そうでないならば $\boldsymbol{A}^0$ の負の固有値に対応する固有ベクトルを $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_{k_1}$ とし，新たな問題

$$(P_1)\quad \left| \begin{array}{ll} 最大化 & lnL(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}) \\ 条件 & \boldsymbol{y}_l^T\boldsymbol{A}\boldsymbol{y}_l \geq 0, \quad l = 1, 2, \cdots, k_1 \end{array} \right.$$

を解く。この問題の最適解を $(\alpha_0^1, \boldsymbol{\alpha}^1, \boldsymbol{A}^1)$ とし，$\boldsymbol{A}^1$ が半正定値であれば，$(\alpha_0^1, \boldsymbol{\alpha}^1, \boldsymbol{A}^1)$ が問題 (P) の最適解である。一方，そうでないときは，$\boldsymbol{A}^1$ の負の固有値に対応する固有ベクトルを $\boldsymbol{y}_{k_1+1}, \boldsymbol{y}_{k_1+2}, \cdots, \boldsymbol{y}_{k_2}$ として，新たな条件

$$\boldsymbol{y}_l^T\boldsymbol{A}\boldsymbol{y}_l \geq 0, \quad l = k_1 + 1, \cdots, k_2$$

を追加した問題を解き，このプロセスを繰り返す。 ∎

この方法によって，十分な反復を繰返せば (P) の最適解 $(\alpha_0^*, \boldsymbol{\alpha}^*, \boldsymbol{A}^*)$ が生成されるが，収束までに大量の計算が必要となるのが欠点である。

そこで我々は，問題 (P) の近似解を求めるため

$$\left| \begin{array}{ll} 最大化 & lnL(\alpha_0, \boldsymbol{\alpha}, \boldsymbol{A}) \\ 条件 & \boldsymbol{x}_j^T\boldsymbol{A}\boldsymbol{x}_j \geq 0, \quad j = 1, 2, \cdots, M_0 \end{array} \right.$$

を解くことにする。$\boldsymbol{x}_j$ は第 j 企業の財務データである。もし $\boldsymbol{x}_j$ が全空間に万遍なく分布していれば，この問題の最適解は問題 (P) の近似解を与えるはずである。

一方 $\boldsymbol{x}_j$ の分布に偏りがある場合には，この問題の最適解 $\boldsymbol{A}^*$ は非負定値とは限らないが，少なくとも $\boldsymbol{x}_j$ が分布する領域での等倒産強度面は楕円面または放物面に近いものになるものと予想される。

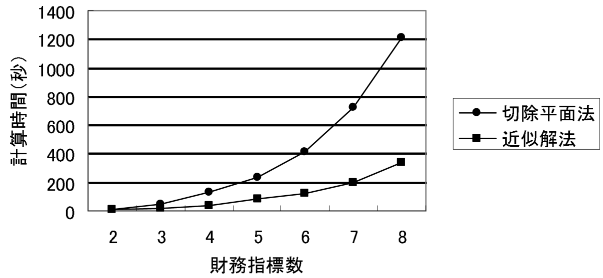

図 1　計算時間の比較

4　計算結果

我々は，1998 年から 2000 年までの非上場製造業における中小企業 7800 社の財務データを用いて数値実験を行った。10%の企業は翌 12ヶ月以内に倒産している。計算機環境は，Personal Computer: CPU; Pentium IV 853MHz, Operating System; Vine Linux 2.1 CR, RAM; 1024MB となる。なお，線形制約の下での凹関数最大化問題は NUOPT Version 6(株式会社数理システム) を利用した。

図 1 から分かるように，提案した近似解法は切除平面法と比較して非常に高速に問題が解ける。財務指標数が増加するに従って計算時間の差は大きくなり，指標数が 8 においては 4 倍程度の計算時間の差が存在する。切除平面法による計算時間は，指標数が大きくなるに従って指数的に増加する。

得られたモデル比較の為に，財務指標数 8 の場合において算出された係数を表 1 に示す。定数項（α_0），1 次の係数（$\boldsymbol{\alpha}$）は非常に近い値となっており，2 次の係数行列（$\boldsymbol{A}$）においては，ほぼ全ての要素の符号が一致しており，等倒産強度面の形状も近いことが分かる。また，最適値の誤差は 1 %未満であった。以上より，提案した方法によって得られたモデルは非常に良い近似となっていることが分かった。

さらに，算出された 2 次の係数行列（$\boldsymbol{A}$）における固有値を表 2（財務指標数 8 の場合）に示した。表 2 より，それぞれのモデルにおける行列 $\boldsymbol{A}$ の固有値は類似しており，特に，大きい順に 3 番目までの固有値はほぼ同じ値となっている。また，近似解法における行列 $\boldsymbol{A}$ における最小固有値の大きさは最大固有値の大きさの 5 %程度となっている。

5　まとめ

過年度の研究においては，企業の倒産確率をより精度

表 1　モデルの比較

次元	モデル	A								α	α0
		0.0153343	0.0111067	0.0282616	0.036353	0.007506	−0.02039	−0.02347	−0.02007	−0.98824	−5.65748
		0.0111067	0.0551532	0.0754336	0.034207	0.03524	0.000919	0.006874	0.046039	1.14921	
		0.0282616	0.0754336	0.16721	0.03317	−0.04423	0.030024	0.011427	0.117169	0.380266	
		0.0363529	0.0342074	0.0331702	0.139323	0.024277	−0.1175	−0.04519	0.037616	−0.41656	
		0.00750567	0.0352398	−0.044225	0.024277	0.063808	−0.04233	−0.01806	−0.10765	0.294263	
		−0.0203924	0.000918713	0.0300238	−0.1175	−0.04233	0.109139	0.069531	−0.02512	−0.40257	
		−0.0234726	0.0068739	0.0114273	−0.04519	−0.01806	0.069531	0.218848	−0.06493	−0.04672	
	近似解法	−0.0200687	0.0460392	0.117169	0.037616	−0.10765	−0.02512	−0.06493	0.433152	−0.95038	
		0.00735234	0.0067285	0.014536	0.02787	0.001076	−0.01844	−0.01167	−0.0005	−0.99171	−5.6231
		0.0067285	0.0488236	0.0702357	0.030501	0.002829	−0.00138	0.021919	0.043249	1.15673	
		0.014536	0.0702357	0.149675	0.031986	−0.0452	0.020191	0.013136	0.124008	0.374398	
		0.0278695	0.0305014	0.0319864	0.144519	0.026	−0.11317	−0.04859	0.028246	−0.42154	
		0.00107636	0.00282878	−0.0451968	0.026	0.068993	−0.03404	−0.01715	−0.10398	0.304087	
		−0.0184397	−0.00138284	0.0201905	−0.11317	−0.03404	0.108792	0.074668	−0.01622	−0.38647	
		−0.0116703	0.0219194	0.013136	−0.04859	−0.01715	0.074668	0.211875	−0.07325	−0.02605	
8	切除平面法	−0.0005035	0.0432486	0.124008	0.028246	−0.10398	−0.01622	−0.07325	0.416179	−0.87958	

表 2　固有値の比較

固有値	1	2	3	4	5	6	7	8
近似解法	0.5328676	0.3226400	0.1985361	0.1270544	0.0603309	−0.0046901	−0.0057764	−0.0289953
切除平面法	0.5207960	0.3191849	0.1881917	0.0908355	0.0372022	−0.0000001	−0.0000006	−0.0000007

良く算出するモデルを考案した。しかし，実務上大規模なデータを利用したモデルの構築が必要とされている。そこで，本年度は，さらに大きな問題を解く解法の研究を行った。現在までの研究により，効率良く非常に近似したモデルの構築が可能となった。しかし，今後，それぞれのモデルにより推定される倒産確率の比較と検証が必要である。また，本研究で提案した近似解法においては，企業数に対応した数の制約式が導入されているが，企業群をいくつかにクラスタリングすることによって，各クラスタの代表値に対応する制約のみを導入することによって計算時間は著しく減少すると思われる。

謝辞

本件研究は 2004 年度，中央大学理工学研究所共同プロジェクトの一環によるものであることを付記する。

参 考 文 献

[1] 今野　浩，武　黛，“半定値計画法による倒産確率推計”，日本応用数理学会論文誌，12 (2002) 121–134.

[2] Konno, H. and Kawadai, N. and Tuy, H., “Cutting Plane Algorithms for Nonlinear Semi-Definite Programming Problems with Applications,” to appear in *J. of Global Optimization*.

[3] Konno, H., Kawadai, N. and Wu, D., “Estimation of Failure Probability using Semi-definite Logit Model,” ISE02–03, Department of Industrial and Systems Engineering, Chuo University, June, 2002 (to appear in *J. of Computational Management Science*).

[4] Konno, H., et al., “Prediction of Bankruptcy of Small to Medium Scale Companies via Semi-Definite Programming,” 中央大学理工学研究所年報, 10 (2003) 53–56.