

修士論文

対象の値動きの方向性に着目した経済時系列予測への
サポートベクターマシンの応用に関する研究

Research regarding the Application of Support Vector Machines
to Predict the Direction of Price Changes in Economic Time Series

平成20年2月4日 提出
指導教員 古関 隆章 准教授

東京大学大学院 情報理工学系研究科
電子情報学専攻 66432

中田 貴之

Takayuki Nakata

内容梗概

時系列の予測にかかわる研究はこれまで数多く行われており、それらの多くはいかに時系列の実際の値との誤差を小さく予測するかといった点に焦点を当てていた。それは時系列の予測という視点から見ると当然のことである。しかし株の取引では収益を上げることが最終的な目標であり、そのためには誤差を小さくするよりも値の動きの方向一致率を高める方がより重要であるとする報告がされている。

そこで本研究ではパターン認識の手法であるサポートベクターマシン (SVM: Support Vector Machines) を用い、値動きの方向一致率を評価指標として予測を行った。その結果、SVMでの短期予測および中期予測はインデックス運用に有用であり、また手法には汎用性が認められるという結論を得た。

Abstract

Today, time series data are predicted using various methods. The main technique currently used to identify a time series has been adjusting the coefficient parameter in order to minimize the disparity between predicted data values. This is a logical method from the viewpoint of time series data prediction, but it does not suit the needs of economic time series prediction. With economic time series, like stock index data, the final goal of prediction is to invest more efficiently. To accomplish efficient investment, it is more important to raise the hit ratio in predicting the direction of price changes rather than to minimize the incongruity in data trends.

This research uses Support Vector Machines (SVM), a method of pattern recognition, and the direction-of-price-change hit ratio, as an index for evaluating economic time series data prediction. The empirical results detailed in this thesis reveal that short term or midterm predictions by SVM are useful for index investments. This method can be generalized for investing in many different financial institutions.

目次

第 1 章	序論	5
1.1	研究の背景	5
1.2	研究の目的	5
1.3	構成	5
第 2 章	時系列	6
2.1	時系列について	6
2.2	時系列の分類	6
2.3	経済時系列	7
2.3.1	日経平均株価	8
2.3.2	海外の株価指数・為替レート	8
2.3.3	その他の時系列	9
第 3 章	サポートベクターマシン	10
3.1	線形分離可能なデータの場合	12
3.2	線形分離不可能なデータの場合	14
3.3	非線形な超平面で分離する手法	16
第 4 章	日経平均株価の短期予測	18
4.1	サポートベクターマシンを時系列の予測に応用する手法	18
4.2	サポートベクターマシンの最適なパラメータの選択	19
4.2.1	従来研究の問題点	19
4.2.2	本研究でのパラメータ探索法	19
4.3	一日後の値の上下予測	21
4.4	擬似トレーディングによる評価	28
4.4.1	売買手数料を考慮せず、終値のみを用いた場合	28
4.4.2	売買手数料を考慮せず、始値と終値を用いた場合	28
4.4.3	売買手数料を考慮し、始値と終値を用いた場合	30
第 5 章	日経平均株価の中期予測	32
5.1	二週間後の値の上下予測	32
5.2	擬似トレーディングによる評価	39
5.2.1	売買手数料を考慮せず、始値と終値を用いた場合	39
5.2.2	売買手数料を考慮し、始値と終値を用いた場合	40

第 6 章 結論	42
6.1 本研究のまとめ	42
6.2 今後の課題	42
参考文献	43
発表文献	45
謝辞	46
付 録 A 補足理論	47
A.1 双対問題と双対定理	47
A.2 ラグランジュ緩和	47
付 録 B Nelder-Mead 法	50
B.1 Nelder-Mead 法の概略	50
B.2 Nelder-Mead 法のアルゴリズム	50

Table of contents

Chapter 1	Introduction	5
1.1	Background of This Research	5
1.2	Purpose of This Research	5
1.3	Composition of This Paper	5
Chapter 2	Time Series	6
2.1	About Time Series	6
2.2	Classification of Time Series	6
2.3	Economic Time Series	7
2.3.1	Nikkei Average	8
2.3.2	Stock Price Index of Foreign Countries and Exchange Rate	8
2.3.3	The Other Time Series	9
Chapter 3	Support Vector Machines	10
3.1	The Linearly Separable Case	12
3.2	The Linearly Non-Separable Case	14
3.3	The Method to Separate by Non-Linear Hyperplane	16
Chapter 4	Short Term Prediction of Nikkei Average	18
4.1	The Method to Apply Support Vector Machines to Prediction of Time Series	18
4.2	Selection of The Best Parameter of Support Vector Machines	19
4.2.1	Problems of Researches in The Past	19
4.2.2	The Parameter Search Method of This Research	19
4.3	Prediction of Up-or-Down of Price of One Day Later	21
4.4	Evaluation by Pseudo-Trading	28
4.4.1	The Case that Closing Prices are Only Used with No Trading Commissions	28
4.4.2	The Case that Closing Prices are Only Used with Trading Commissions	28
4.4.3	The Case that Opening Prices and Closing Prices are Used with Trading Commissions	30
Chapter 5	Mid Term Prediction of Nikkei Average	32
5.1	Prediction of Up-or-Down of Price of Two Weeks Later	32
5.2	Evaluation by Pseudo-Trading	39
5.2.1	The Case that Opening Prices and Closing Prices Only Used with No Trading Commissions	39

5.2.2	The Case that Opening Prices and Closing Prices are Used with Trading Commissions	40
Chapter 6	Conclusion	42
6.1	Summary	42
6.2	Future Tasks	42
	Reference	43
	List of Publications	45
	Acknowledgement	46
Appendix A	Supplementation Theory	47
A.1	Dual Problem and Dual Theorem	47
A.2	Lagrange Relaxation	47
Appendix B	Nelder-Mead Method	50
B.1	About Nelder-Mead Method	50
B.2	Algorithm of Nelder-Mead Method	50

第1章 序論

本章ではまずはじめに研究の背景と目的について述べ、最後に本論文の構成を紹介する。

1.1 研究の背景

時系列の予測はこれまで自己回帰や移動平均などから線形モデルを仮定し、その後にデータとの差が最小となるように係数パラメータを調整することで行われてきた [1]。しかし、複雑な要因がからみあって形成されている経済時系列のような非線形な時系列を正確に予測することは難しい。

そして、計算機の大幅な進歩は大規模データを扱うことを容易にし、これに伴って多様な手法を用いての経済時系列予測の研究が盛んに行われるようになった。ニューラルネットワーク [2] や遺伝的アルゴリズム [3]、遺伝的プログラミング [4] という手法を用いて、未来の経済時系列、とりわけ株価の値を予測するものはそれぞれ一定の成果をあげていたが、値に対する位相遅れが目立っていたり、教師信号にオーバーフィッティングしてしまったりしていた。

これらの多くはいかに時系列の実際の値との誤差を小さく予測するかといった点に焦点を当てている。それは時系列の予測という視点から見ると当然のことである。しかし株の取引では収益を上げることが最終的な目標である。

そこで近年では株価の値そのものを予測するのではなく、売買のポイントを予測したり [5, 6]、や売買ルールを作成する研究 [7, 8] が増えてきている。

1.2 研究の目的

株式データのような経済時系列においては、未来の値動きを予測し、実際に効率的に投資が行えることが最終目標である。本研究ではその点に的を絞り、効率的に投資が行える売買タイミングの意思決定に資する経済時系列予測について考える。

実際に効率的な投資を行うためには値の誤差を小さく予測するよりも値の動きの予測方向一致率を高める方がより重要であるとする報告がされている [9]。そこで本研究ではパターン認識の一手法であるサポートベクターマシン (SVM: Support Vector Machines) を用い、値動きの方向一致率を評価指標として予測を行う。

1.3 構成

本論文の構成は以下の通りである。まず第2章で時系列全般を説明し、第3章で Support Vector Machine の理論について述べる。そして、第4章と第5章では提案手法として、一日後の値の上下を予測する手法と大きな変動を予測する手法を説明し、それぞれについて擬似トレーディングによって評価を行う。最後に第6章でまとめと今後の課題を述べる。

第2章 時系列

本章ではまず時系列について大きく概略とその分類について述べた後、本研究にて予測対象となる経済時系列である日経平均株価について説明し、最後に海外市場の株価指数、為替相場の時系列及びその他の時系列について記す。

2.1 時系列について

時間の経過とともに不規則に変動する現象の記録が時系列である [10]。気圧、気温や雨量などの気象情報、地震波の記録、株価や為替レートなどの経済現象の記録、脳波や心電図などの医学データなど我々の生活に身近なところだけでも数多くの時系列がある。Fig.2.1 に例として、太陽の黒点数の年次変化 [11] を示した。

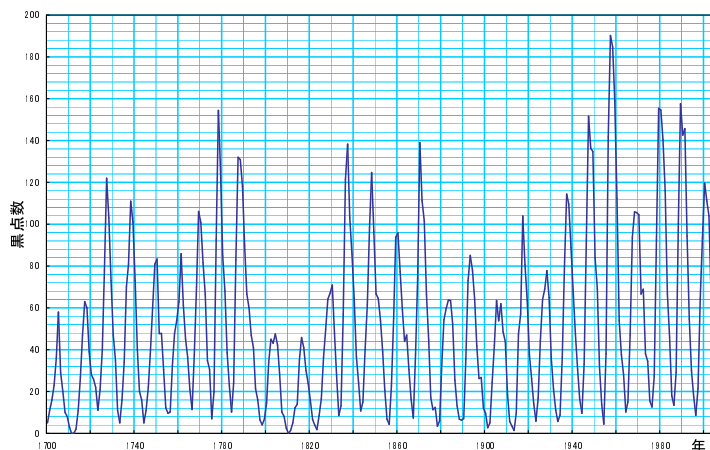


図 2.1: 太陽黒点数の年次変化

2.2 時系列の分類

上記の通り、時系列にはさまざまな種類のものがある。これらの時系列は以下のようにいろいろな観点から類別することができる。

(1) 連続時間時系列と離散時間時系列

アナログレコーダー等で連続的に記録されたデータは連続時間時系列と呼ばれる。これに対し、1時間おきに計測された気圧のようにある時間間隔で観測されたものを離散時間系列と呼ぶ。

(2) 一変量時系列と多変量時系列

各観測時点で1種類の情報だけが得られたものが一変量時系列であり、2種類以上の情報を同時に記録したものが多変量時系列である。ただし、ある現象と同時に起こる現象はいくらでもあるので、時系列が一変量であるか多変量であるか、また多変量の場合でもどのような変数を同時に考慮するかは解析の対象によって一意的に決まるものではない。

(3) 定常時系列と非定常時系列

一見したところ不規則な現象であるが、時間的に変化しない一定の確率的モデルの実現値とみなせる時系列は定常時系列と呼ばれる。一方、気温のように平均が時間とともに変動していたり、地震波のように平均まわりの変動の仕方が時間的に変化しているものを非定常時系列という。

(4) 線形時系列と非線形時系列

線形なモデルの出力として表現できるような時系列は線形時系列と呼ぶ。これに対し、非線形なモデルが必要なものは非線形時系列と呼ぶ。

また、時系列は大きく分けて4つの方法で解析される。

A. 記述

時系列を図示したり、標本自己共分散関数、ピリオドグラムなどの基本的な記述統計量を用いて時系列の特徴を簡潔に表現すること。

B. モデリング

与えられた時系列に対し、その変動の仕方を表現する時系列モデルを構築し、時系列の確率的構造を解析すること。

C. 予測

時系列が互いに相関を持つことを利用して、現在までに得られた情報から今後の変動を予測すること。

D. 信号抽出

時系列から目的に応じて必要な信号や情報を取り出すこと。

本研究では「予測」を行っているが、これは決してCだけではなく、AからDすべての考え方を取り入れたものである。

2.3 経済時系列

時系列の中でも特に、経済現象に関連する時系列を経済時系列と呼び、株価、為替相場、金利などの値動きなどが挙げられる [1]。経済時系列はきわめて不規則な変動を示し、一般に非定常かつ非線形な時系列であると考えられている。また、さまざまな経済時系列のデータが同時に存在することから多変量とみなすこともできる。

2.3.1 日経平均株価

日経平均株価は東京証券取引所第一部上場の銘柄株式を対象とした株価指数である。東京証券取引所における一日の取引時間は、午前9時から11時と12時半から15時となっている。Fig.2.2のように午前9時に付いた値をその日の始値、15時に付いた値を終値、一日の中で最も高い値を高値、最も安い値を安値とし、四本値と呼ばれるこの四種の値を一日のデータとして容易に入手することができる [12]。

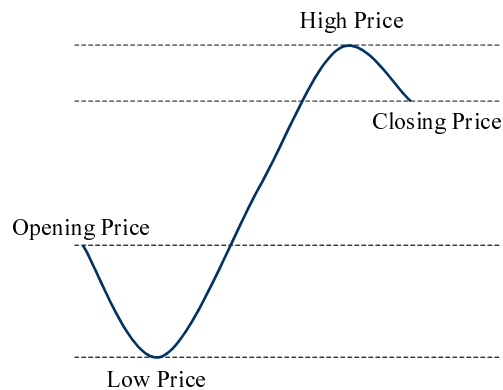


図 2.2: 株価の四本値

2.3.2 海外の株価指数・為替レート

海外の株価指数と為替レートについても日経平均株価と同様に始値、終値、高値、安値の4本値を同様に入手できる。以下ではそれらの中でも代表的なもので、特に日経平均株価との繋がりが深いと考えられるアメリカ市場の株価指数二つとヨーロッパ市場の株価指数を一つ、そして円に対する為替レート(円相場)について説明する。図 2.3

1. ダウ平均株価 (DJIA: Dow Jones Industrial Average)

ダウ・ジョーンズ社がアメリカのさまざまな業種の代表的な銘柄を選出し、公表している株価指数。ニューヨーク証券取引所上場銘柄を主な対象としている。ダウ平均株価には3種類あるが、最も有名なものは「工業株30種平均株価」でありこれが一般的にダウ平均と呼ばれている。

2. ナスダック総合指数 (Nasdaq: Nasdaq Composite Index)

アメリカ合衆国にある世界最大の新興企業向け株式市場であるNASDAQの株価指数。ニューヨーク証券取引所、NASDAQともに取引時間は日本時間の23時半から翌6時までである。

3. FTSE100種総合株価指数 (FTSE100: Financial Times Stock Exchange 100 Index)

ロンドン証券市場の株価指数であり、欧州を代表する株価指数。ロンドン証券市場に上場している銘柄のうち時価総額で上位100銘柄で構成されている時価総額加重平均型株価指数である。

4. 円相場

円に対する為替レート、つまり外貨の相対的価値は通常外貨1単位に相当する円の額で表示され、特に米ドル、ユーロ、英ポンドとの比較によって示される。それぞれ、ドル円 (USD/JPY: USD-JPY Exchange Rate)、ユーロ円 (EUR/JPY: EUR-JPY Exchange Rate)、ポンド円 (GBP/JPY:

GBP-JPY Exchange Rate) と呼ばれる。これらの相場は土日を除く 24 時間取引が行われている。終値にはニューヨーク証券取引所の取引が終わった時間の値が用いられることが多い。

ここに挙げた 6 つの時系列に、日経平均株価を加えた 7 つの時系列を 2002 年から 2007 年までの期間で -1 から 1 に正規化してプロットしたものが図 2.3 である。

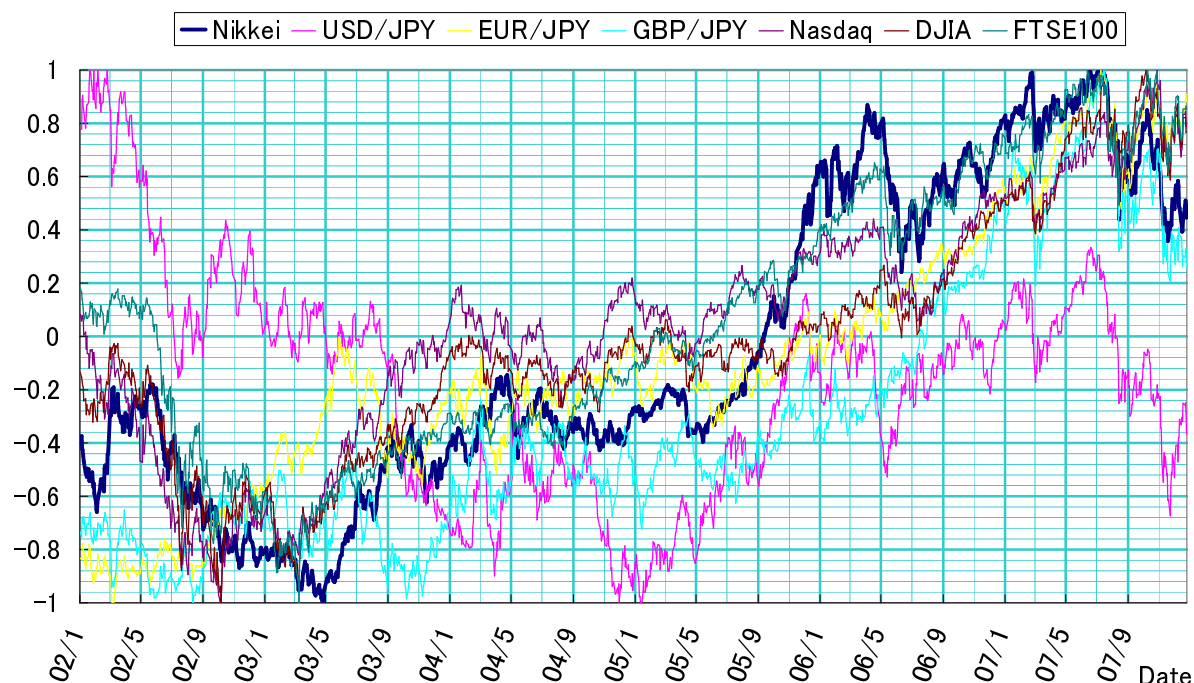


図 2.3: 正規化された各種時系列

2.3.3 その他の時系列

その他の時系列データとして、内閣府経済社会総合研究所が月次で発表している機械受注指標や景気動向指数といったマクロ経済指数、商品市場で扱われている原油や金などの先物相場の値などがある。これらはいずれも年単位での経済動向、ひいては日経平均株価を予測する際には欠かせないものであり、また 1 日や 2 週間といった短中期の日経平均株価の動きに対しても影響を与えていると考えられている。しかし、本研究ではその寄与の大きさは先に挙げた海外の株価指数や為替レートの方がより大きいと考え、その他の時系列については考慮しないこととした。

第3章 サポートベクターマシン

サポートベクターマシン (SVM: Support Vector Machines)[13, 14, 15, 16] とはパターン認識の一手法である。パターン認識とは既にクラス分けされているデータを用い、未知の認識対象を計測して得られた特徴量の集まりからその対象がどのクラスに属するかを判定する識別器を開発することである。

ここで簡単な具体例として、毒キノコを見分ける方法を取り上げる。とある村に老齢のキノコ名人がおり、今年収穫されたキノコのうち 100 本の毒キノコと 100 本の毒の無いキノコを選び出したまま旅に出てしまった。村人たちは残りの大量のキノコに対して毒の有無を見分けなければならない。毒の有無がわかっている 200 本の学習用のキノコを利用して、毒キノコを見分ける最良の方法を導き出すこと、これがパターン認識の扱う課題である。

各キノコの特徴のうち、柄の長さ、柄の太さ、笠の面積、重さ、アミノ酸含有量など定量的に扱えるものをキノコの特徴量と呼ぶ。今、 d 個のこうした特徴量に着目したとすると、あるキノコの特徴は特徴空間と呼ばれる d 次元空間の 1 点 x として表現することができる。この d 次元ベクトル x をパターンと呼び、特に学習用キノコから得られるパターンを学習パターンと呼ぶ。毒のないキノコの集合をクラス 1、毒キノコの集合をクラス 2 としたとき、解決すべき問題は属するクラスが不明なパターン x のクラスを判定する識別規則を学習パターンを利用して求めることであり、これは特徴空間上での 2 クラスの境界を決定することに相当する。

多重パーセプトロンなどのパターン認識手法では学習の際に経験リスクの最小化だけを行う。つまり学習データに対してのみ学習結果を保証する。そのため汎化的な性能が低くなり、極端に学習データにのみ特化してしまう過学習の状態に陥ってしまう問題点があった。一方 SVM は構造的リスク最小化の原理に基づいていて経験リスクと汎化性能の両面から学習の最適化が行われるため、未知の対象に対してもより正確に識別できる。図 3.1(a) に二次元の例を挙げたが、この場合白丸と黒丸を識別する境界は無数にある。SVM 以外の他手法では学習データに対してのみ結果を保証するため、この境界の中で優劣はつけられない。そのため図の中で曲線で示したような過学習になる可能性もある。

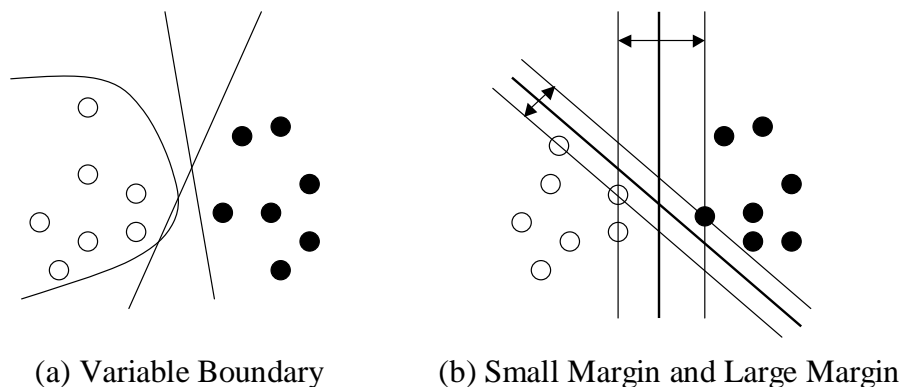


図 3.1: 2 クラスにパターンを分類する識別境界

特徴空間上における2クラスのパターン分布において、線形分離可能な場合には Fig.3.1(b) のように一般に学習パターン全てを正しく識別する超平面が複数存在する。識別境界候補となるそうした超平面の周囲には学習パターンが存在しない空白の領域が存在し、SVM では複数の識別境界候補の中でこの領域の間の距離であるマージンを最大にする超平面を最良とみなす。SVM の学習、すなわち学習パターンからマージンを最大にする最適識別関数を求める過程を以下では数学的に定式化して説明する。

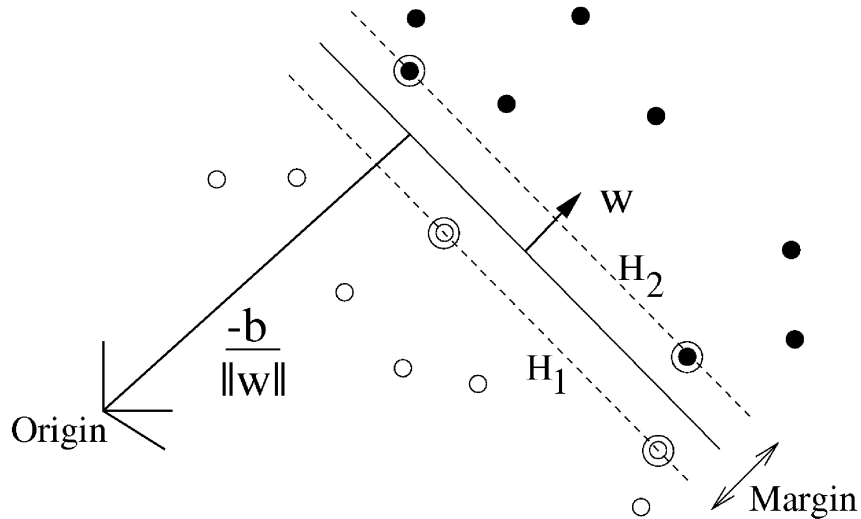


図 3.2: 線形分離可能な場合の識別平面

3.1 線形分離可能なデータの場合

図 3.2 のように線形分離可能な正負二つのクラス ($y_i = \pm 1, i = 1, \dots, n$) に分けられた n 個の学習パターン $\mathbf{x}_i \in \mathbb{R}^n (i = 1, \dots, n)$ を超平面で分離する場合、識別関数 $f(\mathbf{x})$ は次のように表される。

$$f(\mathbf{x}) \equiv \text{sign}(g(\mathbf{x})) = \text{sign}(\mathbf{w}^t \mathbf{x} + b) \quad (3.1)$$

ここで \mathbf{w} は重みをあらわすパラメータのベクトルで、重みベクトルと呼ばれ、 b はバイアス項と呼ばれるパラメータである。そして n 個の学習パターン \mathbf{x}_i の満たすべき条件を

$$g(\mathbf{x}) = \mathbf{w}^t \mathbf{x}_i + b \begin{cases} \geq 1 & (y_i = 1) \\ \leq -1 & (y_i = -1) \end{cases} \quad (3.2)$$

$$\Leftrightarrow y_i(\mathbf{w}^t \mathbf{x}_i + b) - 1 \geq 0$$

とする。点 \mathbf{x}_i から平面 $g(\mathbf{x}) = 0$ までの距離は $\frac{|g(\mathbf{x}_i)|}{\|\mathbf{w}\|}$ であるから、式 (3.2) は識別境界から距離 $\frac{1}{\|\mathbf{w}\|}$ の範囲内、すなわち平面 $g(\mathbf{x}) = 1$ と $g(\mathbf{x}) = -1$ の間に学習パターンが存在しないことを意味する。平面 $g(\mathbf{x}) = \pm 1$ 間の距離は $\frac{2}{\|\mathbf{w}\|}$ である。以下ではこの平面 $g(\mathbf{x}) = \pm 1$ 間の距離のことをマージンと呼ぶ。先に述べたように、このマージンを最大にするような超平面を SVM では最良とみなす。マージンを最大にする超平面は、扱いやすくなるように $\|\mathbf{w}\|^2$ と y_i を用いて考えて

$$\begin{aligned} \min_{\mathbf{w}, b} \quad & G(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t.} \quad & y_i(\mathbf{w}^t \mathbf{x}_i + b) - 1 \geq 0 \end{aligned} \quad (3.3)$$

の解により決まる。

この最適化問題に対応する双対問題に変換する。双対問題への変換の手順については付録 A に記した。ラグランジュ乗数 $\alpha_i (\geq 0, i = 1, \dots, n)$ を用いて、制約条件を緩和すると

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i(\mathbf{w}^t \mathbf{x}_i + b) - 1) \quad (3.4)$$

となる。この関数は \mathbf{w} と b について最小化される。そこで \mathbf{w} および b に関する偏微分から停留条件

$$\frac{\partial L(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = \mathbf{0} \quad (3.5)$$

$$\frac{\partial L(\mathbf{w}, b, \alpha)}{\partial b} = \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.6)$$

を得る。すなわち、

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad (3.7)$$

$$0 = \sum_{i=1}^n \alpha_i y_i \quad (3.8)$$

という関係を得る。これらを式 (3.4) を変形した式 (3.9) に代入する。

$$\begin{aligned} L(\mathbf{w}, b, \alpha) &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w}^t \mathbf{x}_i + b) - 1) \\ &= \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i y_i \mathbf{w}^t \mathbf{x}_i - b \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i \end{aligned} \quad (3.9)$$

$$\begin{aligned} L(\alpha) &= \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j - \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j + \sum_{i=1}^n \alpha_i \\ &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j \end{aligned} \quad (3.10)$$

$$s.t. \quad \alpha_i \geq 0 (i = 1, \dots, n), \sum_{i=1}^n \alpha_i y_i = 0$$

このようにして式 (3.10) を最大化するという双対問題に変換することができる。また、これを解くと多くの α_i が 0 となる。式 (3.7) より、0 ではない α_i に対応する \mathbf{x}_i のみから \mathbf{w} が決まるため、 $\alpha_i > 0$ を満たすものだけが識別超平面の決定に寄与する。このような \mathbf{x}_i はサポートベクターと呼ばれ、図 3.2、図 3.3 では二重丸で示した。

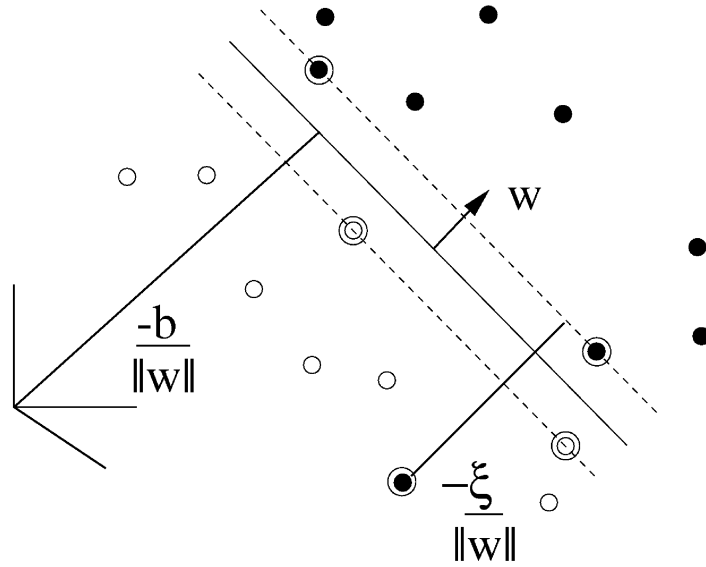


図 3.3: 線形分離不可能な場合の識別平面

3.2 線形分離不可能なデータの場合

図 3.3 のように線形分離不可能な場合、式 (3.2) を満たす w は存在しない。そこで正の変数 $\xi_i (i = 1, \dots, n)$ を導入し、学習パターンに対して若干の誤りを許すように条件を緩め、

$$g(\mathbf{x}) = \mathbf{w}^t \mathbf{x}_i + b \begin{cases} \geq 1 - \xi_i & (y_i = 1) \\ \leq -1 + \xi_i & (y_i = -1) \end{cases} \quad (3.11)$$

とする。これをソフトマージン法と呼ぶ。そして最小化問題を

$$\begin{aligned} \min_{\mathbf{w}, b, \xi} G(\mathbf{w}, \xi) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t. } & y_i(\mathbf{w}^t \mathbf{x}_i + b) - (1 - \xi_i) \geq 0, \xi_i \geq 0 \end{aligned} \quad (3.12)$$

と定める。式 (3.12) の右辺第 1 項は線形分離可能な場合と同様にマージンを大きくとるためのものであり、一方第 2 項はマージンからはみだしたパターンに対するペナルティ項である。係数 C は第 1 項と第 2 項のバランスを決める定数で、問題に応じて適切に与える。 C が小さければ平面 $g(\mathbf{x}) = \pm 1$ 間の距離は大きくなる。最適な C を定めることが重要であるといえる。

3.1 節と同様に最適化問題を双対問題へと変換する。ラグランジュ変数 $\alpha_i, \beta_i (\geq 0, i = 1, \dots, n)$ を用いて

$$L(\mathbf{w}, b, \xi, \alpha, \beta) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i (y_i(\mathbf{w}^t \mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^n \beta_i \xi_i \quad (3.13)$$

とする。 \mathbf{w}, ξ, b に関する偏微分から

$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial \mathbf{w}} = \mathbf{w} - \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = \mathbf{0} \quad (3.14)$$

$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial \xi_i} = C - \alpha_i - \beta_i = 0 \quad (3.15)$$

$$\frac{\partial L(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial b} = \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.16)$$

という停留条件を得る。つまり、

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \quad (3.17)$$

$$C = \alpha_i + \beta_i \quad (3.18)$$

$$0 = \sum_{i=1}^n \alpha_i y_i \quad (3.19)$$

という関係を得る。式 (3.18) 及び、 $\beta_i \geq 0$ より、 $0 \leq \alpha_i \leq C$ となる。また、これらを式 (3.13) を変形した式 (3.20) に代入すると、

$$\begin{aligned} L(\mathbf{w}, b, \xi, \alpha, \beta) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w}^t \mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^n \beta_i \xi_i \\ &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i y_i \mathbf{w}^t \mathbf{x}_i - \sum_{i=1}^n \alpha_i y_i b + \sum_{i=1}^n \alpha_i - \sum_{i=1}^n \alpha_i \xi_i - \sum_{i=1}^n \beta_i \xi_i \\ &= \frac{1}{2} \|\mathbf{w}\|^2 + \sum_{i=1}^n (C - \alpha_i - \beta_i) \xi_i - \sum_{i=1}^n \alpha_i y_i \mathbf{w}^t \mathbf{x}_i - b \sum_{i=1}^n \alpha_i y_i + \sum_{i=1}^n \alpha_i \end{aligned} \quad (3.20)$$

$$\begin{aligned} L(\alpha) &= \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j - \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j + \sum_{i=1}^n \alpha_i \\ &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^t \mathbf{x}_j \end{aligned} \quad (3.21)$$

$$s.t. \quad 0 \leq \alpha_i \leq C (i = 1, \dots, n), \sum_{i=1}^n \alpha_i y_i = 0$$

となり、双対問題である α のみの最大化問題へと帰着できる。この双対問題は変換前の最小化問題と比較して見通しが立てやすくなっているため、簡単な勾配登りのアルゴリズムやニュートン法により解を得ることができる。

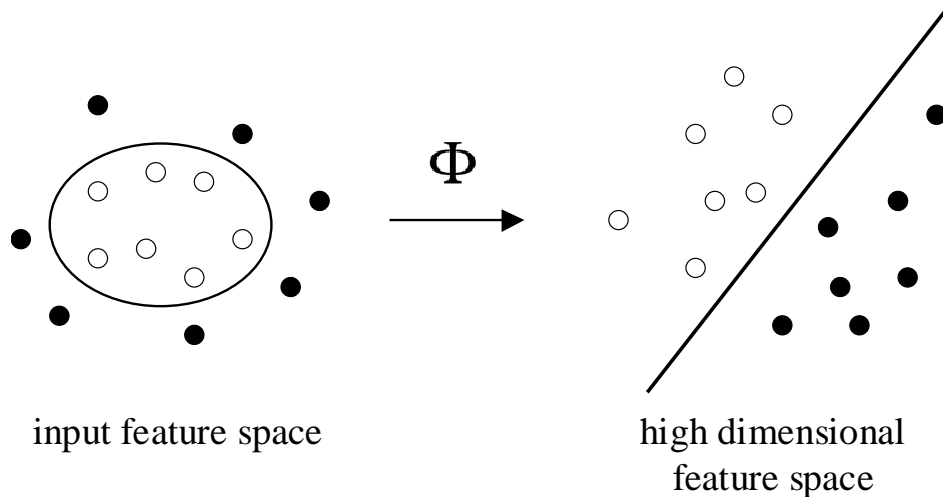


図 3.4: 高次元空間への写像

3.3 非線形な超平面で分離する手法

ソフトマージン法を用いたとしても、本質的に非線形で複雑な識別問題に対しては必ずしも良い性能の識別器を構成できるとは限らない。そこで前処理として Fig.3.4 で示すように非線形な写像を用いてより高次元の空間に写像を行い、線形分離性を高めることが考えられる。

$$\Phi : \mathbf{R}^n \rightarrow \mathbf{R}^q \quad (n < q) \quad (3.22)$$

このような写像を行い、写像先の空間 \mathbf{R}^q で線形識別を行えば、実質的に元の空間で非線形識別を行っているのと同じことになる。

\mathbf{R}^q において SVM を適用すると、識別関数は式 (3.1) より

$$f(\Phi(\mathbf{x})) = \mathbf{w}^t \Phi(\mathbf{x}) + b \quad (3.23)$$

となる。ここで \mathbf{w} は式 (3.17) のように書けるから、

$$f(\Phi(\mathbf{x})) = \sum_{i=1}^n \alpha_i y_i \Phi(\mathbf{x})^t \Phi(\mathbf{x}_i) + b \quad (3.24)$$

また、同様に学習の問題も

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \Phi(\mathbf{x}_i)^t \Phi(\mathbf{x}_j) \quad (3.25)$$

$$s.t. 0 \leq \alpha_i \leq C (i = 1, \dots, n), \sum_{i=1}^n \alpha_i y_i = 0$$

となる。

ここで元の空間で定義されるカーネル関数 $K(\mathbf{x}_i, \mathbf{x}_j)$ を用意して、 Φ は式 (3.26) の条件を満たすと仮定する。

$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^t \Phi(\mathbf{x}_j) \quad (3.26)$$

このとき、式 (3.24) や式 (3.25) は式 (3.26) を満たす K の関数として書くことができ、 Φ を陽に含まない。したがって、内積が計算できれば最適な識別器を構成することが可能であるため、非線形写像によって変換された空間での特徴ベクトル Φ を陽に計算する必要はない。このように高次元に写像しながら、写像された空間での特徴の計算を実際には避ける方法を「カーネルトリック」と呼んでいる。ここで式 (3.26) のように表せるカーネル関数を正定値カーネルと呼ぶ。 $K(\mathbf{x}_i, \mathbf{x}_j)$ が \mathbb{R}^n 上の正定値カーネルであるとは、次の条件を満たすことである。

- 対称性

$$K(\mathbf{x}_i, \mathbf{x}_j) = K(\mathbf{x}_j, \mathbf{x}_i) \quad (3.27)$$

- 正定値性

任意の実数 c_1, \dots, c_n に対し、

$$\sum_{i,j=0}^n c_i c_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad (3.28)$$

上記の条件を満たす正定値カーネルとしては、式 (3.29) で定義される多項式カーネルや式 (3.30) のガウシアンカーネルが知られている。なかでも、ガウシアンカーネルは非線形に写像が可能であるため特徴量が少なくすむこと、パラメータが一つであることという二つの長所を持っているため、本研究ではガウシアンカーネルを利用した。

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i^t \mathbf{x}_j + 1)^p \quad (3.29)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (3.30)$$

第4章 日経平均株価の短期予測

本章では日経平均株価の短期予測として、一日後の値の上下をSVMを用いることにより予測した。これは日経平均株価の過去のデータのうち、得られる最小の単位が一日であるからである。SVMを時系列の予測に応用する手法、およびSVMのパラメータの定め方について述べた後、SVMへの入力データを変化させた時の予測結果を示し、最後にその予測を評価するために擬似トレーディングを行った。

4.1 サポートベクターマシンを時系列の予測に応用する手法

まず、訓練データ、テストデータともに予測対象日の終値が当日の終値より高くなっている日をA、低くなっている日をBとクラスを決定した。そして過去の株価データである訓練データでSVMによりAとBの識別面を学習し、それをテストデータに適用しクラス分類することにより一日後の値の予測を行った。識別面の学習、及びそれをテストデータに適用する流れについては図4.1にそのフローを示した。訓練データにおいてはSVMのパラメータを決めるため、識別誤り率が最も低い点を探索する。このパラメータの探索については4.2節で述べる。

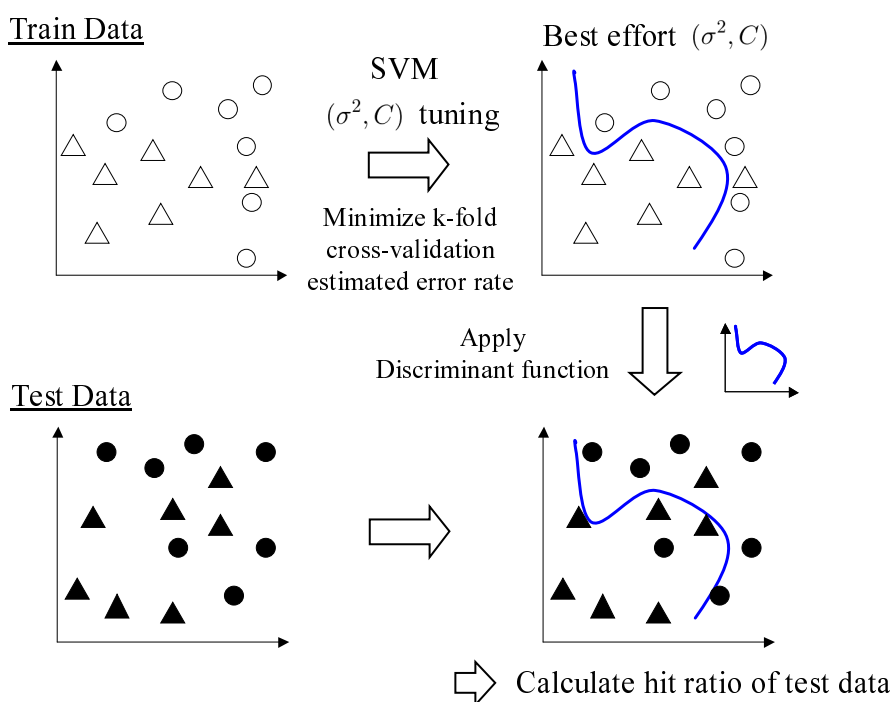


図 4.1: SVM を用いた予測システムの流れ

4.2 サポートベクターマシンの最適なパラメータの選択

SVMを使用するにあたり、パラメータの選択は非常に重要である。ここではSVMを時系列の予測に応用した従来研究におけるパラメータ選択の問題点と、それを踏まえた本研究における最適パラメータの探索法を述べる。

4.2.1 従来研究の問題点

SVMを時系列、とりわけ経済時系列の予測に応用した従来研究としては文献[17]、文献[18]、文献[19]の3つがある。

このうち文献[17]、文献[18]の2つについてはパラメータを既知として値を定めており、そのパラメータに対する予測結果のみについて論じている。SVMのパラメータの決定に際しての最適化の評価は訓練データのみで行うべきであり、その論拠をはっきりと示さなければただテストデータに対して最も性能を良くするパラメータを持ってきたと思われてしまい、結果の正当性に対する疑問が残る。

文献[19]ではパラメータの変化に対しての訓練データ、テストデータそれぞれの識別成功率を掲載している。しかし、訓練データのみでパラメータを決めると明示的に論じておらず、テストデータで最も良い識別成功率をとるパラメータのみについて論じている。

4.2.2 本研究でのパラメータ探索法

本研究において、SVMのパラメータにはマージンを大きくとるということとマージンからはみ出したパターンに対するペナルティとのバランスをとる役割をする式(3.12)の C と式(3.30)のガウシアンカーネルの σ^2 の2つがある。

従来研究の問題点をふまえ、本研究では訓練データでの識別誤り率が最も小さいものを最良のパラメータの組とみなした。また訓練データの識別誤り率を計算する際には図4.2に示すように10分割交差検定法を用いた。これは、訓練データ全体で作成された識別関数を訓練データ全体でテストするだけでは過学習してしまったものが最良とみなされるので、それを避けたいと考えたからである。

訓練データの識別誤り率の概要を知るため、まずは探索空間の2つの独立変数で張られる平面を網の目のように $\ln \sigma^2 = -10, -9, \dots, 10$ 及び $\ln C = -10, -9, \dots, 10$ と切ってその格子点上で総当りをした。これを以下では網の目探索法と呼ぶ。

その後、関数の導関数、つまり勾配がわからなくても最適点を探索できる直接探索法であるNelder-Mead法[20]を用いた。Nelder-Mead法については付録Bに記した。また、Nelder-Mead法の初期点の一つに網の目探索法での最高点を用いることで探索の手間を減らした。

このようにして得られたパラメータの組と訓練データによって作成された識別関数をテストデータに当てはめて識別が成功した率、つまり予測の方向が一致した率を識別成功率とした。また、テストデータの識別成功率については図4.3に示したとおり、網の目のパラメータの組においてそれぞれ作成された識別関数をテストデータに当てはめて全て計算した。

このテストデータの識別成功率のグラフと訓練データの識別誤り率のグラフを比較検討することで、SVMに入力した特徴量の形が適切であったかより視覚的に判断することができる。具体的にいえば訓練データの識別誤り率のグラフとテストデータの識別成功率のグラフが同じような形であるほど、SVMに入力した特徴量の形が適切であったといえる。

Train data

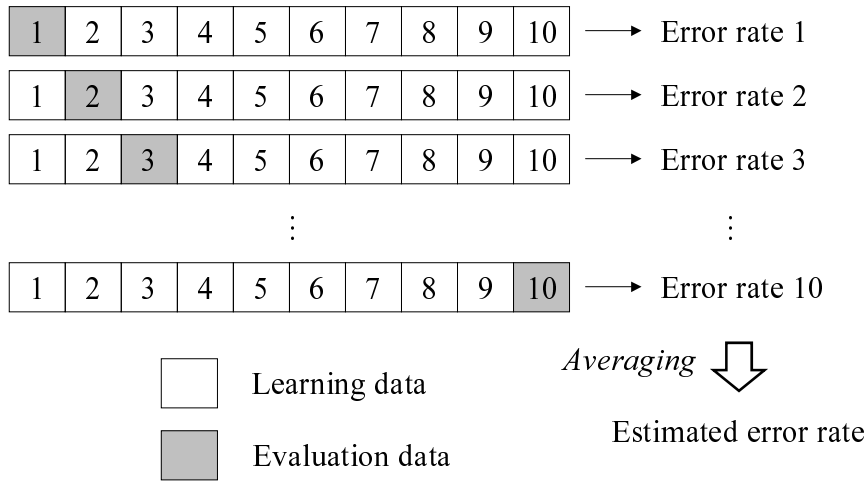


図 4.2: 10 分割交差検定法

Train Data

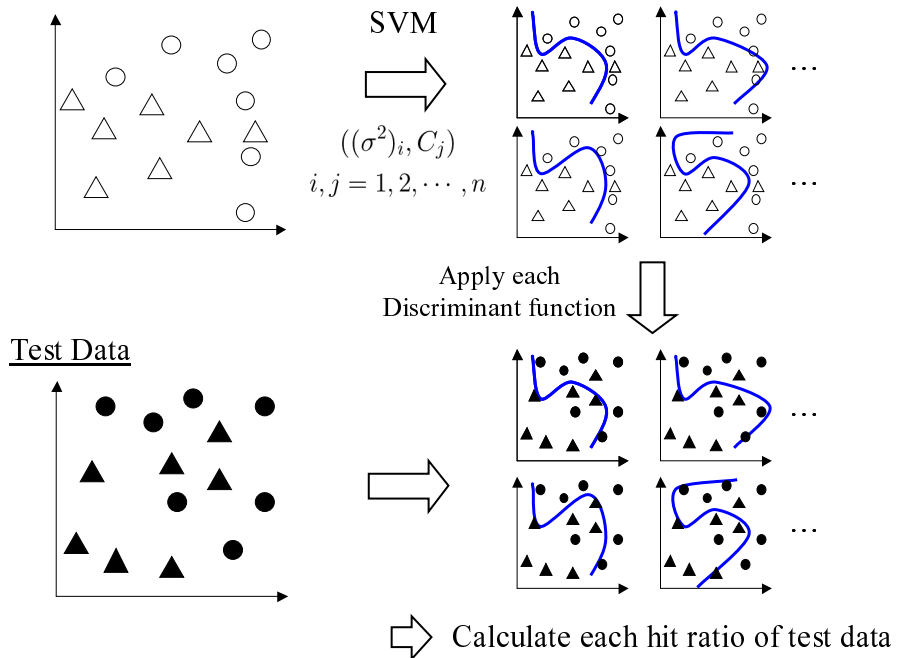


図 4.3: テストデータの識別成功率を計算する方法

4.3 一日後の値の上下予測

この予測に用いる日経平均株価の期間及びサンプル数は表 4.1 の通りである。図 4.4 に示したように、訓練データの期間は上げ相場、テストデータの期間である 2007 年は乱高下相場となっている。異なった性格の相場が良い結果を得られれば、予測手法としての汎用性が高いといえる。

また、SVM のインプットとしては日経平均株価の他に、2.3.2 節で解説したドル・ユーロ・ボンドの円相場、及びナスダック総合指数、ダウ平均株価、イギリスの株価指数である FTSE1000、計 7 つの時系列を用いた。また、これらの時系列と日経平均株価との相関係数を表 4.2 にまとめた。

この 7 つの時系列を時間変動の周波数特性の影響を見るため、以下に示す 4 通りの方法で加工しそれぞれについて予測を行った。以下では n 日目の終値を x_n とする。

$$\begin{aligned} \text{Case A} & \quad x_0 - x_1 \\ \text{Case B} & \quad x_0 - \sum_{i=1}^5 \frac{(11-2i)x_i}{25} \\ \text{Case C} & \quad x_0 - \sum_{i=1}^5 \frac{(6-i)x_i}{15} \\ \text{Case D} & \quad x_0 - \sum_{i=1}^5 \frac{x_i}{5} \end{aligned}$$

一日後の値の上下を予測するという事はすなわち当日と一日後の値の「差」を予測するという事であるから、それぞれの入力には「差」を用いている。Case A は当日と前日の終値の単純な差である。しかし、ただ差をとただけではノイズが大きく有効成分が埋もれてしまうため、Case D では過去 5 日間のデータの平均と当日との差をとっている。これはローパスフィルタの考え方であるが、単純に過去 5 日間のデータを平均しただけでは時定数の遅れが大きいと考えられるため、より近い日に重み付けをして大きくした過去 5 日間の重み付け平均と当日との差をとったものが Case B と Case C である。Case B の方がより近い日に重みをつけている。また、これら 4 つの Case は図 4.5 に図示した。

4.1 節で論じた方法で予測を行い、まず網の目探索法を行った結果及びそれぞれの点におけるテストデータの識別成功率を各々の Case について図 4.6 から図 4.13 に示した。各 Case 網の目探索法において最小の識別誤り率であったパラメータの組とそのパラメータの組におけるテストデータの識別成功率は表 4.3 の通りである。また、その後 Nelder-Mead 法を用いて識別誤り率の最小点とされた際のパラメータの組とその組でのテストデータの識別成功率は表 4.4 のようになった。最もテストデータの識別成功率が良かったのは Case C であり、その時の予測結果は表 4.5 となっている。

表 4.1: 予測に用いる日経平均株価データ

	期間	サンプル数	クラス A	クラス B
Train Data	2002.10 - 2006.4	879	469	410
Test Data	2007.1 - 2007.12	244	124	120

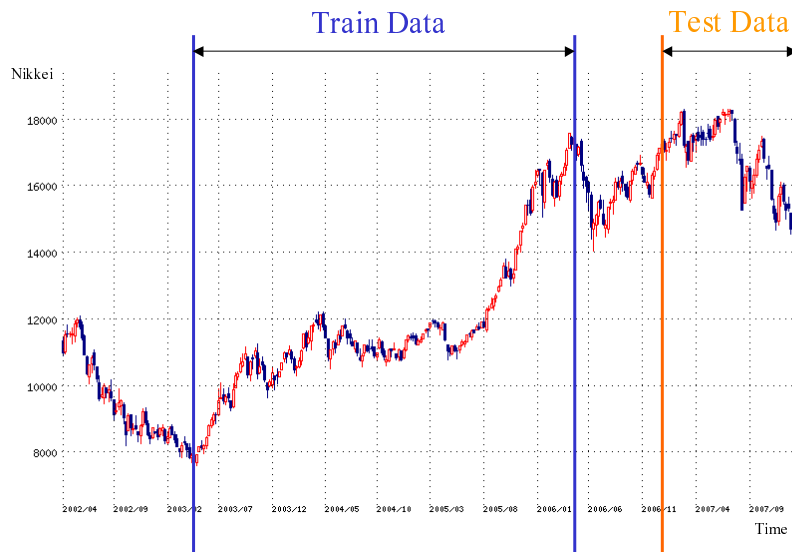


図 4.4: 訓練データ及びテストデータの期間

表 4.2: SVM のインプットに用いた経済時系列およびそれらと日経平均株価の相関係数

	略称	日経平均株価との相関係数
Dow Jones Industrial Average	DJIA	0.374
Nasdaq Composite Index	Nasdaq	0.414
Financial Times Stock Exchange 100 Index	FTSE100	0.313
USD-JPY Exchange Rate	USD/JPY	0.117
EUR-JPY Exchange Rate	EUR/JPY	0.067
GBP-JPY Exchange Rate	GBP/JPY	0.083

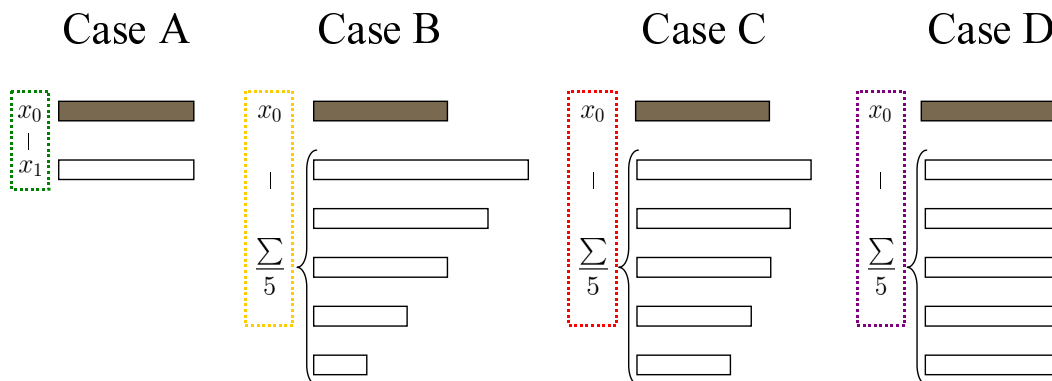


図 4.5: Case A,B,C,D における入力データ加工

表 4.3: グリッドサーチ最高点での識別成功率

	$\ln \sigma^2$	$\ln C$	Error rate	Hit ratio
Case A	3.00	3.00	0.366	0.574
Case B	6.00	6.00	0.380	0.648
Case C	8.00	6.00	0.388	0.668
Case D	8.00	3.00	0.389	0.643

表 4.4: パラメータチューニング後の識別成功率

	$\ln \sigma^2$	$\ln C$	Error rate	Hit ratio
Case A	3.00	3.00	0.366	0.574
Case B	6.60	6.63	0.372	0.660
Case C	7.60	6.00	0.371	0.676
Case D	7.20	2.98	0.381	0.643

表 4.5: パラメータチューニング後の Case C の予測結果

	Class A	Class B
Predicted Class A	104	59
Predicted Class B	20	61

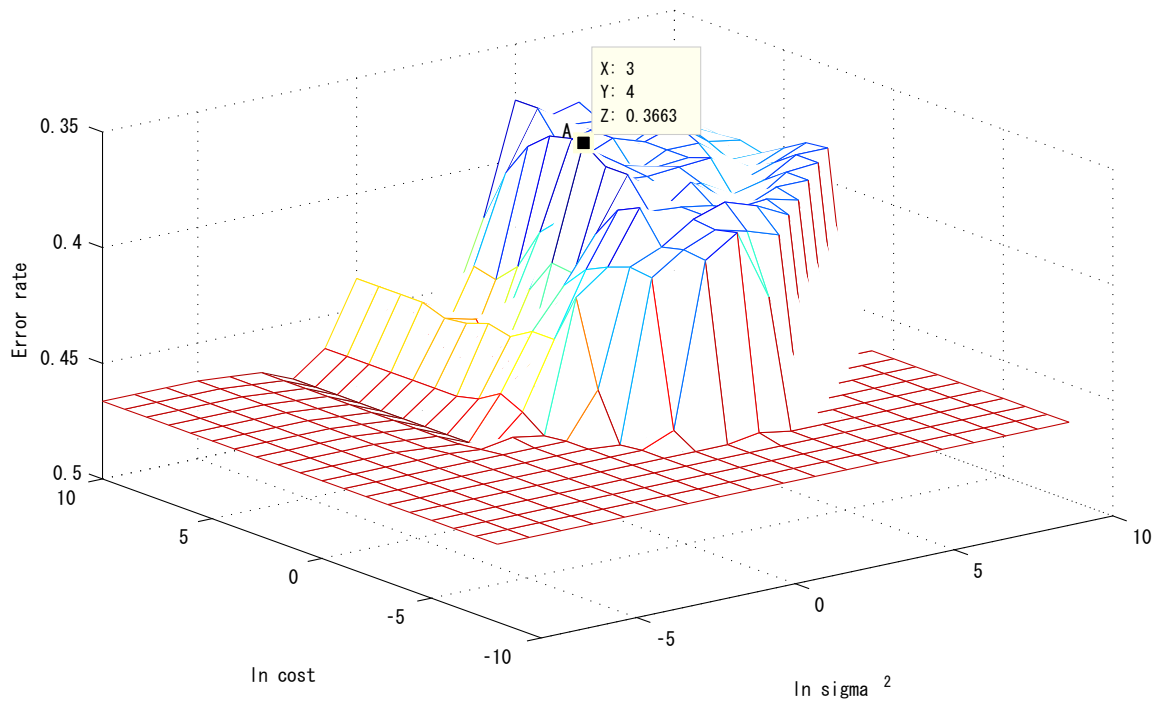


図 4.6: 訓練データの識別誤り率 (Case A)

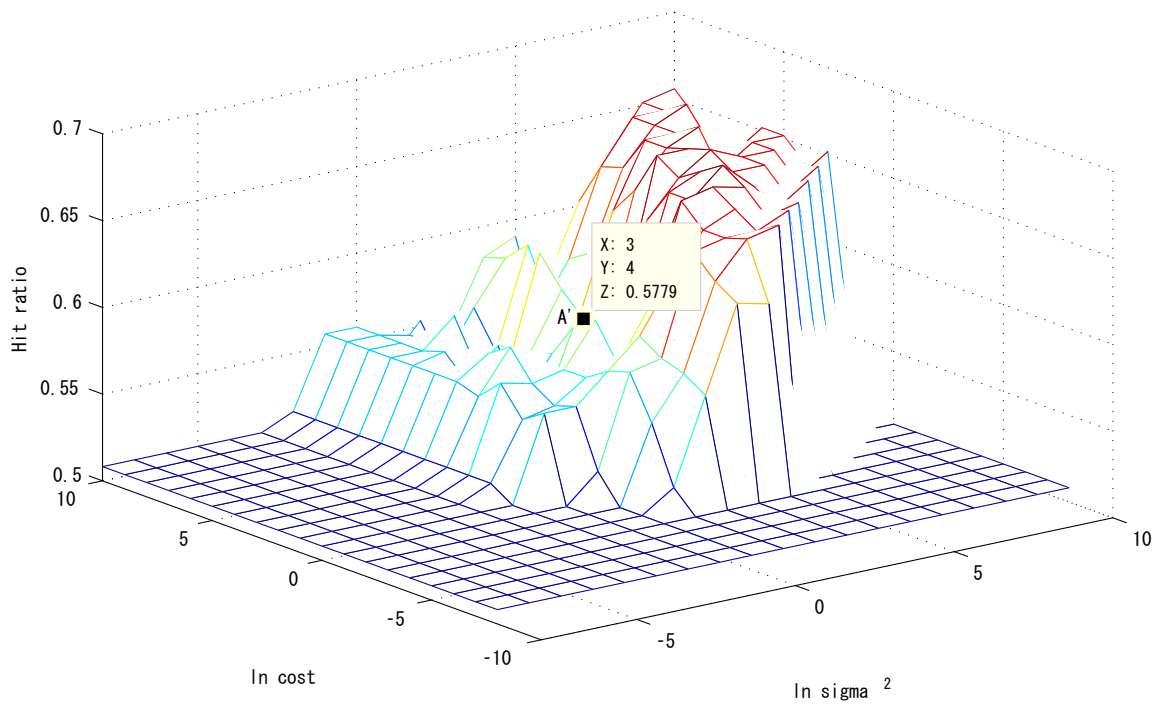


図 4.7: テストデータの識別成功率 (Case A)

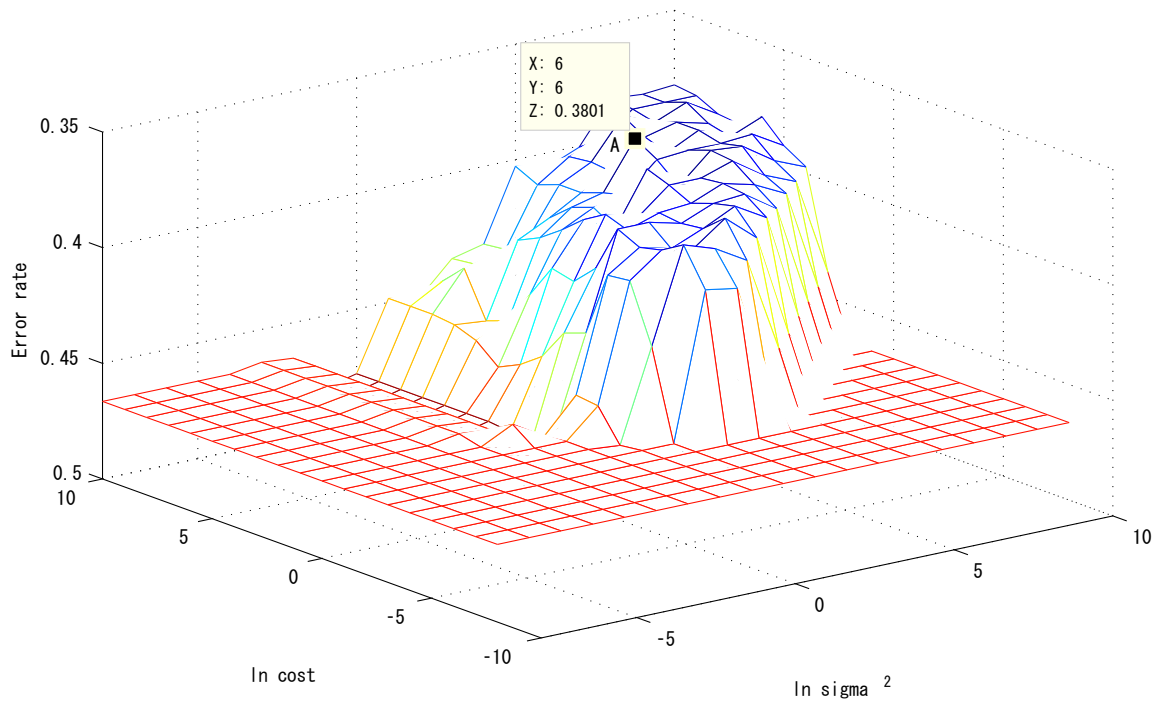


図 4.8: 訓練データの識別誤り率 (Case B)

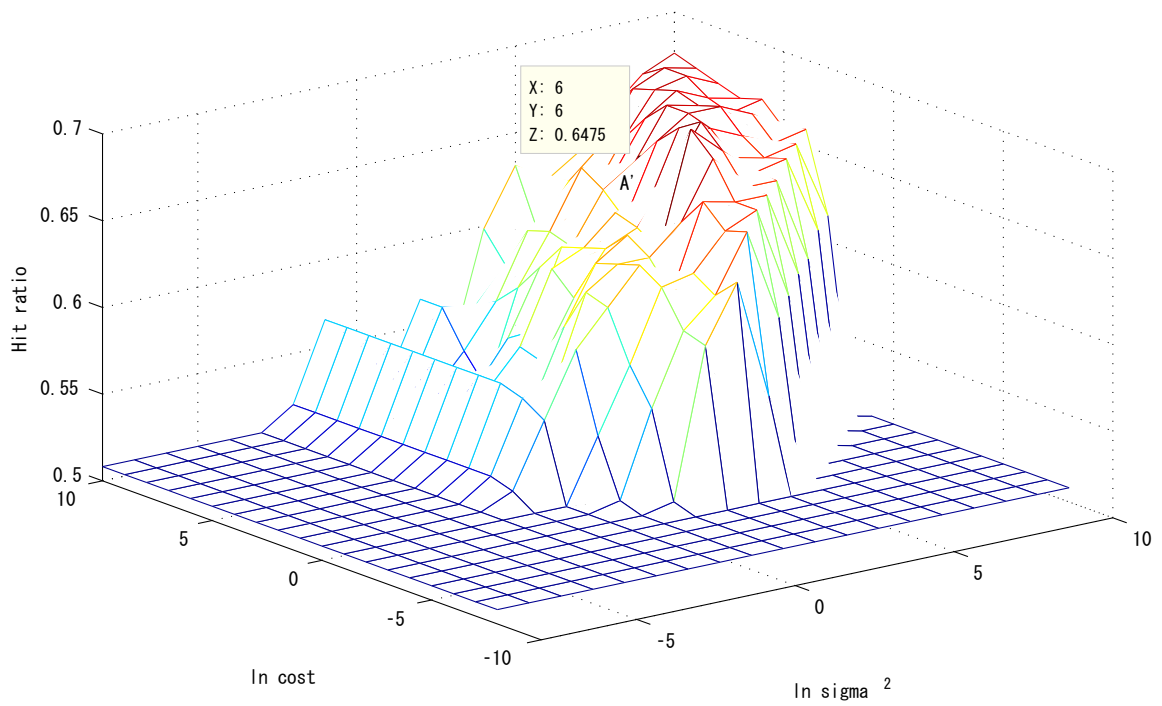


図 4.9: テストデータの識別成功率 (Case B)

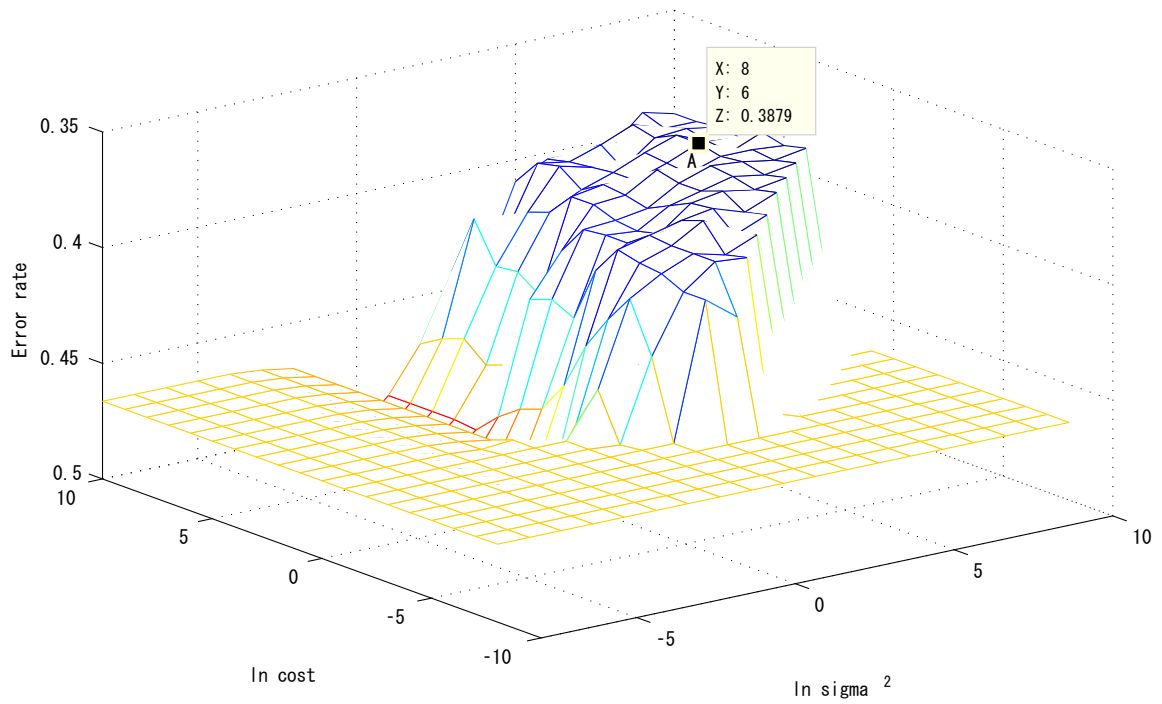


図 4.10: 訓練データの識別誤り率 (Case C)

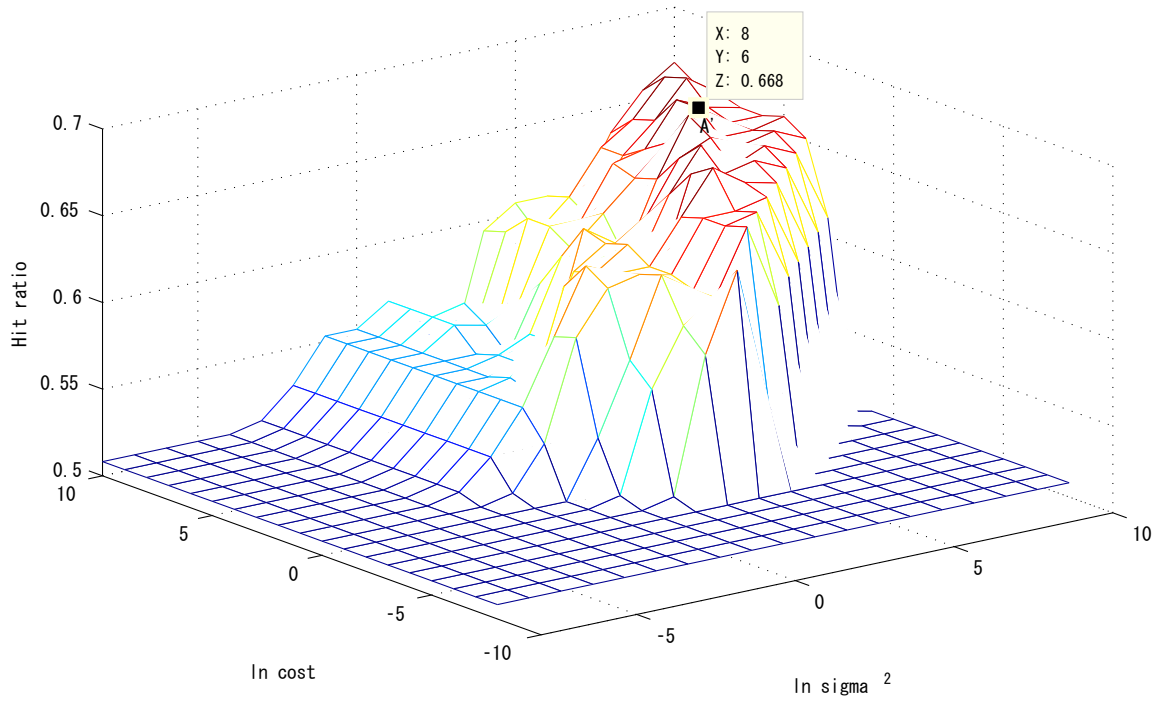


図 4.11: テストデータの識別成功率 (Case C)

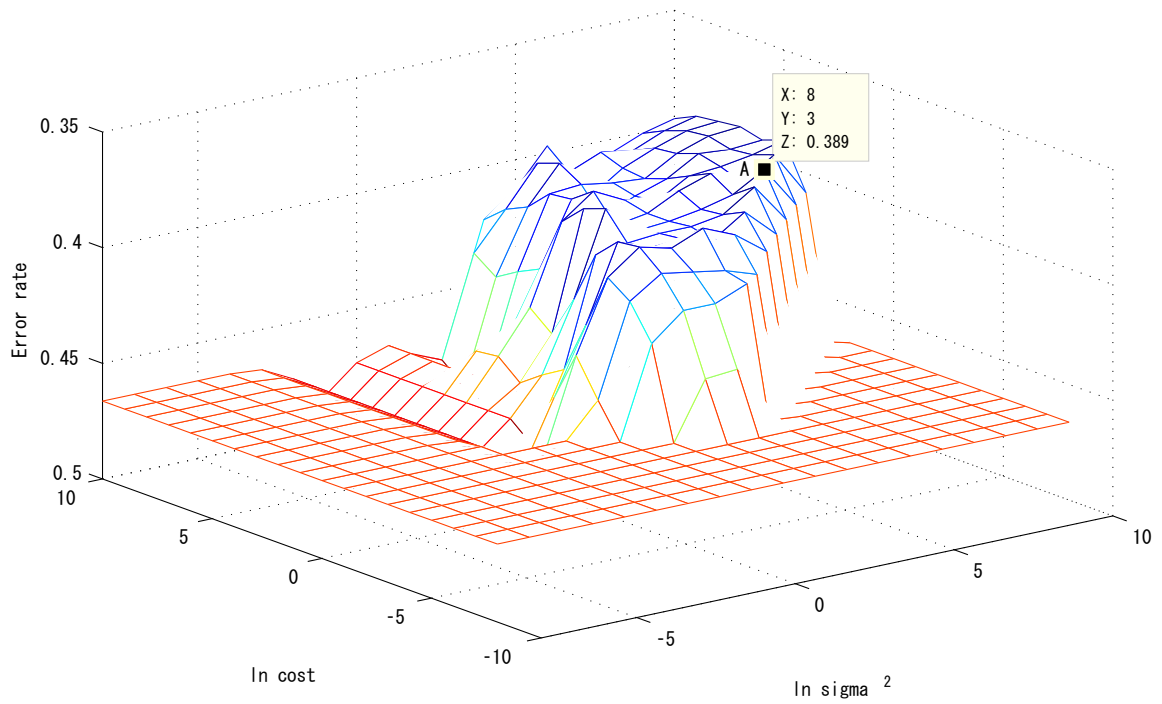


図 4.12: 訓練データの識別誤り率 (Case D)

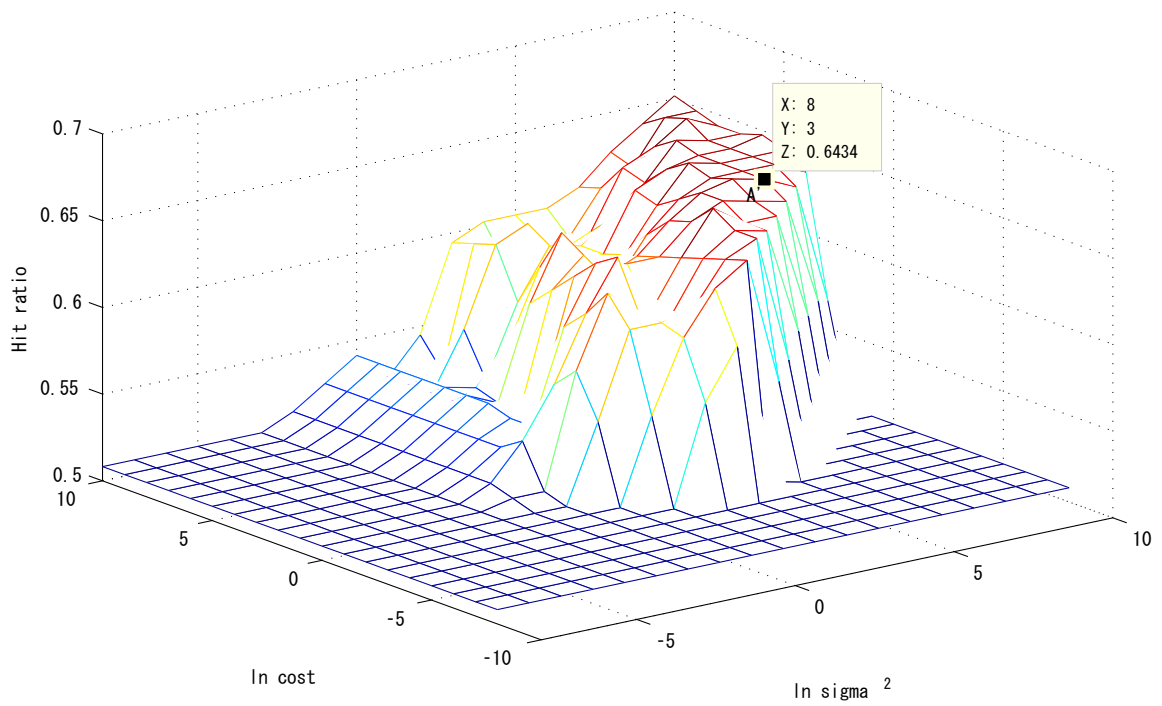


図 4.13: テストデータの識別成功率 (Case D)

4.4 擬似トレーディングによる評価

日経平均株価の擬似トレーディングを行って、この予測の評価を行った。予測の評価を目的とするため、トレーディングはこの予測を全面的に信頼するように行われ、決して利益の最大化を狙ったものにはなっていない。トレーディングの対象としては日経平均株価と同じ値動きをする株を仮定する。

まずは売買手数料を考慮せず単純に終値のみを用いた場合を考えてシンプルに方向一致率の評価を行い、次に現実的なモデルとして、終値だけでなく始値も使って擬似トレーディングを行った場合、さらに売買手数料を考慮した場合について述べる。

4.4.1 売買手数料を考慮せず、終値のみを用いた場合

まず、はじめの擬似トレーディングは以下のルールで行った。

1. 最初の予算は 100 万円
2. 売買の判断は毎日行う。
3. 手元に株を持っていない場合
 - 翌日の終値が当日の終値より「上がる」と予測された場合、当日の終値で手元の資金が許す限り株を買う。
 - 翌日の終値が当日の終値より「下がる」と予測された場合、株を買わず現金のままで資金をもち続ける。
4. 手元に株を持っている場合
 - 翌日の終値が当日の終値より「上がる」と予測された場合、そのまま株をもち続ける。
 - 翌日の終値が当日の終値より「下がる」と予測された場合、当日の終値で手元の株を全て売却する。
5. 売買の際に手数料はかからない。

この擬似トレーディングではシンプルに方向一致率の評価を行うため、終値しか用いていない。その結果は、図 4.14 の通りである。最も良い識別一致率をもった Case C がやはり最も良いトレーディング結果を残し、おおよそ年 80% のリターンを得ることができている。

4.4.2 売買手数料を考慮せず、始値と終値を用いた場合

本研究では当日の終値から翌日の終値が上がるか下がるかという予測を行っており、その予測の精度をよりわかりやすく評価するために、4.4.1 節では終値のみを用いて擬似トレーディングをおこなった。しかし、実際の相場においては予測を行った後にその日の終値で売買をすることはできず、翌日の相場が開くまで売買の機会は待たなければならない。

そこで次は以下のルールで、擬似トレーディングを行った。

1. 最初の予算は 100 万円

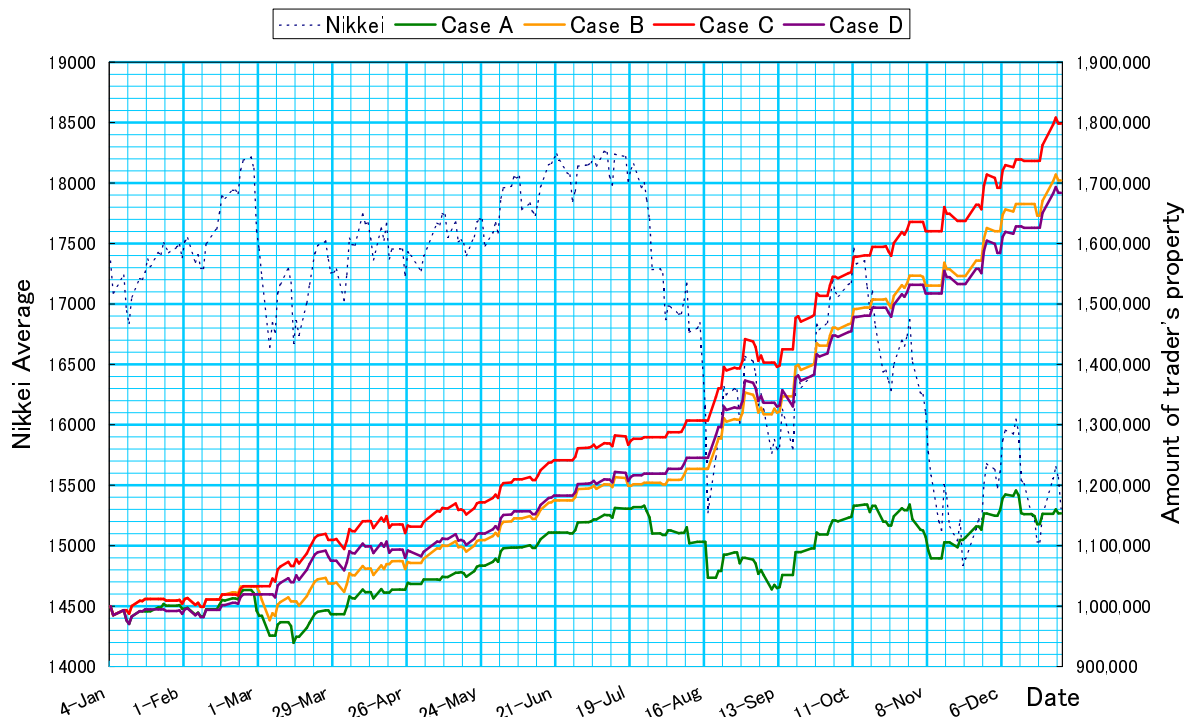


図 4.14: 手数料を考慮せず終値のみを用いた擬似トレーディング結果

2. 売買の判断は毎日行う。

3. 手元に株を持っていない場合

- 翌日の終値が当日の終値より「上がる」と予測された場合、翌日の始値で手元の資金が許す限り株を買う。
- 翌日の終値が当日の終値より「下がる」と予測された場合、株を買わず現金のままで資金を持ち続ける。

4. 手元に株を持っている場合

- 翌日の終値が当日の終値より「上がる」と予測された場合、そのまま株を持ち続ける。
- 翌日の終値が当日の終値より「下がる」と予測された場合、翌日の始値で手元の株を全て売却する。

5. 売買の際に手数料はかからない。

この擬似トレーディングの結果は図 4.15 のようになった。4.4.1 節の擬似トレーディングと比較すると利益が減っているが、それでも Case C では 15% の利益をあげることができている。一般的に株を売買する際には手数料が発生するが、投資信託などを購入する際には手数料が発生しないものもある。投資信託のなかには日経平均株価と連動しほぼ同じ動きをするものもあり、年に 1% 弱かかる信託報酬という別の手数料を考慮すると、約 15% の利益は現実に可能であるといえる。

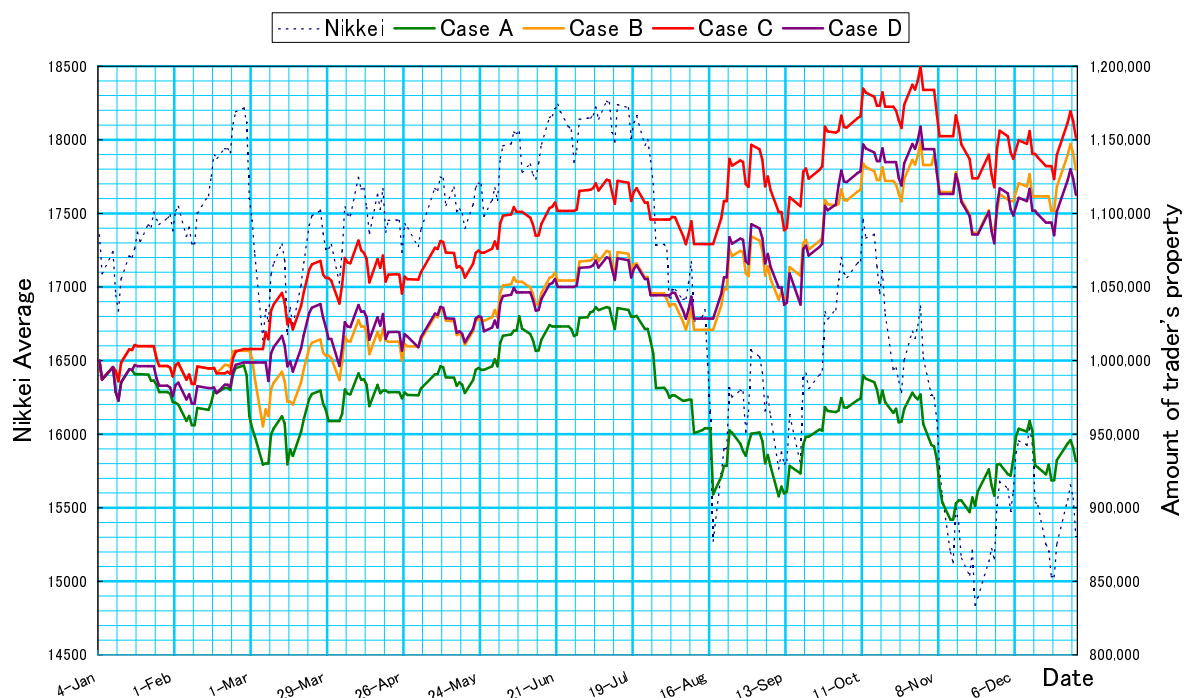


図 4.15: 手数料を考慮せず始値と終値を用いた擬似トレーディング結果

4.4.3 売買手数料を考慮し、始値と終値を用いた場合

最後に売買手数料を考慮し、以下のルールで擬似トレーディングを行った。手数料を決めるにあたっては、インターネット上の比較サイト [21] を参考とした。表 4.6 に 2008 年 2 月現在の代表的なインターネット証券会社各社の取引手数料を示しておく。

1. 最初の予算は 100 万円
2. 売買の判断は毎日行う。
3. 手元に株を持っていない場合
 - 翌日の終値が当日の終値より「上がる」と予測された場合、翌日の始値で手元の資金が許す限り株を買う。
 - 翌日の終値が当日の終値より「下がる」と予測された場合、株を買わず現金のままで資金を持ち続ける。
4. 手元に株を持っている場合
 - 翌日の終値が当日の終値より「上がる」と予測された場合、そのまま株を持ち続ける。
 - 翌日の終値が当日の終値より「下がる」と予測された場合、翌日の始値で手元の株を全て売却する。
5. 売買の際に手数料一律 1000 円がかかるものとする。

このトレーディングの結果は図 4.16 のようになった。手数料を考慮していない 4.4.2 節の擬似トレーディングと比較すると、どの Case も約 8% リターンが小さくなっていることがわかる。

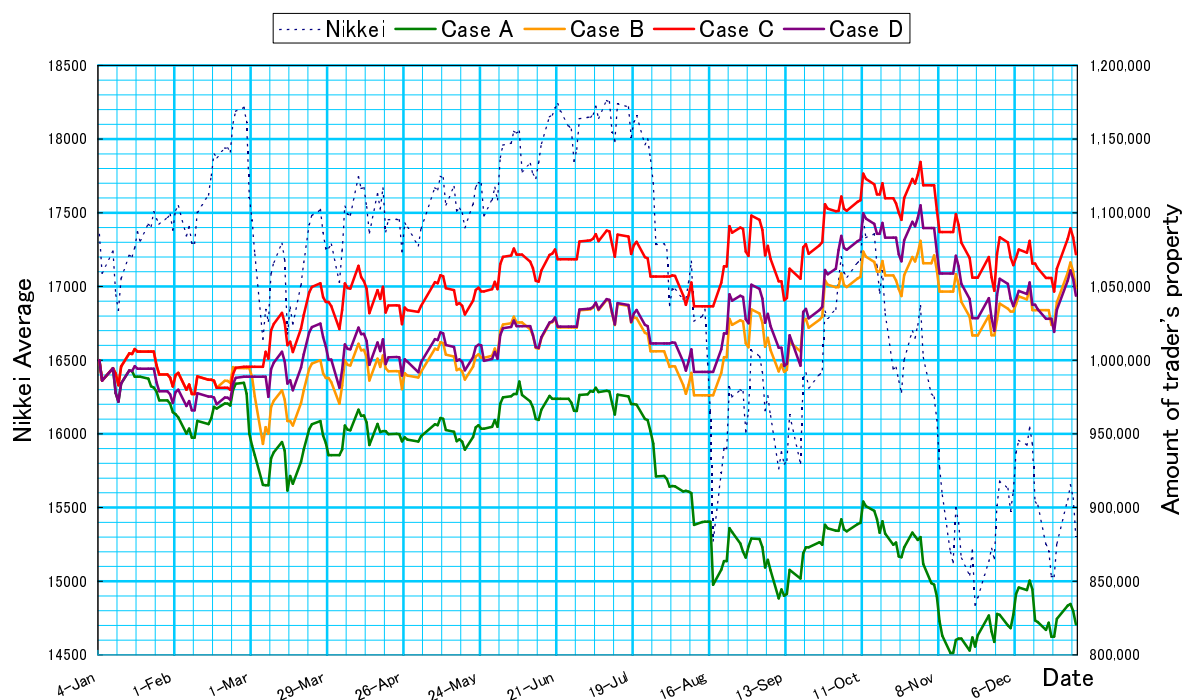


図 4.16: 手数料を考慮し始値と終値を用いた擬似トレーディング結果

最大の利益を得た Case C では 7% の利益をあげている。しかし、この期間で日経平均株価が 12% も下落していることを考えるとこれは決して低い値ではないといえる。

表 4.6: インターネット証券会社各社の取引手数料 (2008 年 2 月現在)

証券会社名	取引手数料			
	売買 10 万円	売買 50 万円	売買 100 万円	売買 300 万円
ジョインベスト証券	100 円	400 円	750 円	1400 円
オリックス証券	315 円	472 円	945 円	2835 円
カブドットコム証券	500 円	1050 円	1050 円	1890 円
マネックス証券	1050 円	1050 円	1050 円	3150 円
SBI イー・トレード証券	200 円	450 円	800 円	1500 円
松井証券	0 円	525 円	1050 円	3150 円
楽天証券	472 円	472 円	840 円	1575 円

第5章 日経平均株価の中期予測

本章では日経平均株価の中期予測として二週間後の値の上下をSVMを用いることにより予測した。これは実際にこのようなシステムに従って売買する人を考えた際に、一日ごとの上下の予測よりも二週間程度といったある程度長いスパンの予測の方がより有用であると考えたためである。また、二週間後と称しているが正確には10営業日後である。土日が休みであることを考えるとほぼ二週間後であることから、以下では二週間後とする。4章と同様にSVMへの入力データを変化させた時の予測結果を示し、その予測を評価するために擬似トレーディングを行った。

5.1 二週間後の値の上下予測

期間及びサンプル数は表5.1の通りであり、期間については5.1節の予測と同じである。

SVMへのインプットは、節で挙げた7つの時系列をどの程度の平滑化が有効であるのかを確かめるため、以下に示す5通りの方法で加工しそれぞれについて予測を行った。以下では n 日目の終値を x_n とする。

$$\text{Case E} \quad x_0 - x_{10}$$

$$\text{Case F} \quad \left(\sum_{j=0}^1 x_j - \sum_{j=0}^1 x_{j+10} \right) / 2$$

$$\text{Case G} \quad \left(\sum_{j=0}^2 x_j - \sum_{j=0}^2 x_{j+10} \right) / 3$$

$$\text{Case H} \quad \left(\sum_{j=0}^3 x_j - \sum_{j=0}^3 x_{j+10} \right) / 4$$

$$\text{Case I} \quad \left(\sum_{j=0}^4 x_j - \sum_{j=0}^4 x_{j+10} \right) / 5$$

表 5.1: 予測に用いる日経平均株価データ

	期間	サンプル数	クラス A	クラス B
Train Data	2002.10 - 2006.4	854	520	334
Test Data	2007.1 - 2007.12	24	13	11

一日後の値の予測と同様に、こちらにも入力には「差」を用いている。Case Eは当日と10営業日後の終値の単純な差である。やはりこれではノイズの影響が大きいと考えられるので、1日後、2日後とそれぞれ平均をとったものどうしの差を用いたものが他のCaseであり、どの程度の平滑化が有効であるかを確かめている。また、これら5つのCaseは図5.1に図示した。

4.1節で論じた方法で予測を行い、まず網の目探索法を行った結果及びそれぞれの点におけるテストデータの識別成功率を各々のCaseについて図5.2から図5.11に示した。各Case網の目探索法において最小の識別誤り率であったパラメータの組とそのパラメータの組におけるテストデータの識別成功率は表5.2の通りである。また、その後Nelder-Mead法を用いて識別誤り率の最小点とされた際のパラメータの組とその組でのテストデータの識別成功率は表5.3のようになった。今回の予測ではどのCaseにおいてもNelder-Mead法を用いた後のパラメータでの識別成功率は網の目探索法で求められたパラメータでの識別成功率と変化はなかった。最もテストデータの識別成功率が良かったのはCase FとCase G、Case Hであり、その時の予測結果は表5.4、表5.5、表5.6となっている。

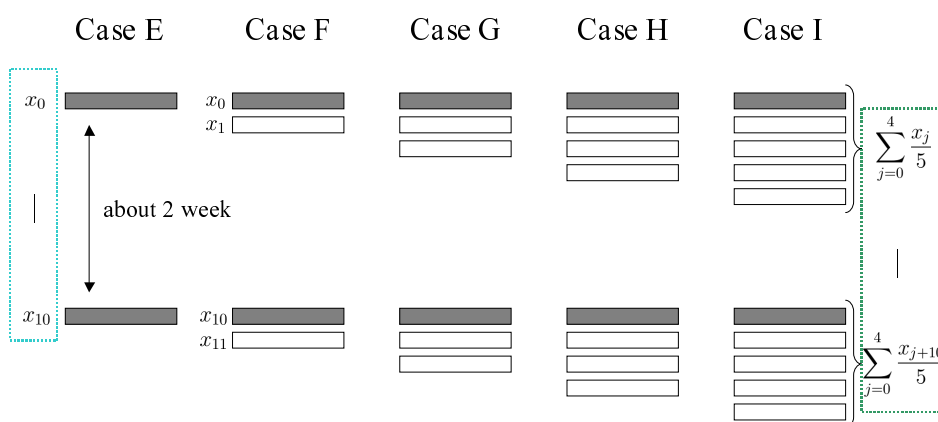


図 5.1: Case E,F,G,H,Iにおける入力データの加工

表 5.2: グリッドサーチ最高点での識別成功率

	$\ln \sigma^2$	$\ln C$	Error rate	Hit ratio
Case E	0.00	1.00	0.209	0.583
Case F	1.00	1.00	0.234	0.708
Case G	0.00	1.00	0.204	0.708
Case H	0.00	1.00	0.126	0.708
Case I	0.00	1.00	0.118	0.625

表 5.3: パラメータチューニング後の識別成功率

	$\ln \sigma^2$	$\ln C$	Error rate	Hit ratio
Case E	-0.01	1.09	0.204	0.583
Case F	0.89	0.99	0.158	0.708
Case G	0.03	1.10	0.112	0.708
Case H	0.01	1.03	0.119	0.708
Case I	-0.05	1.10	0.101	0.625

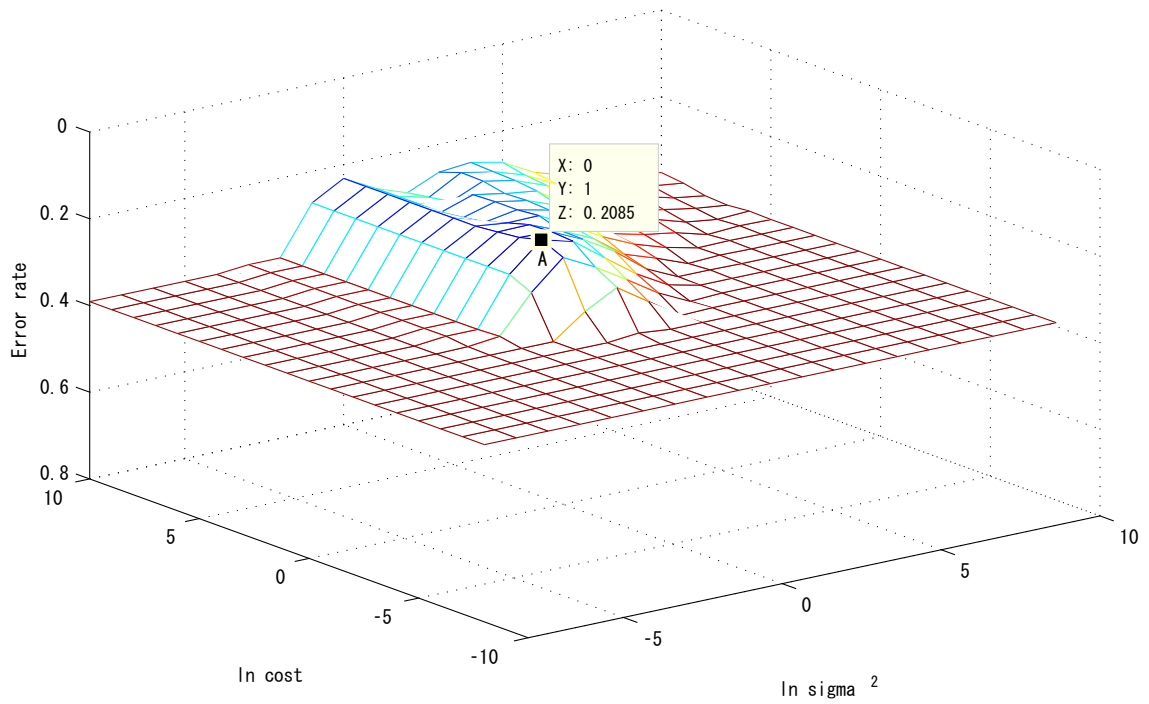


図 5.2: 訓練データの識別誤り率 (Case E)

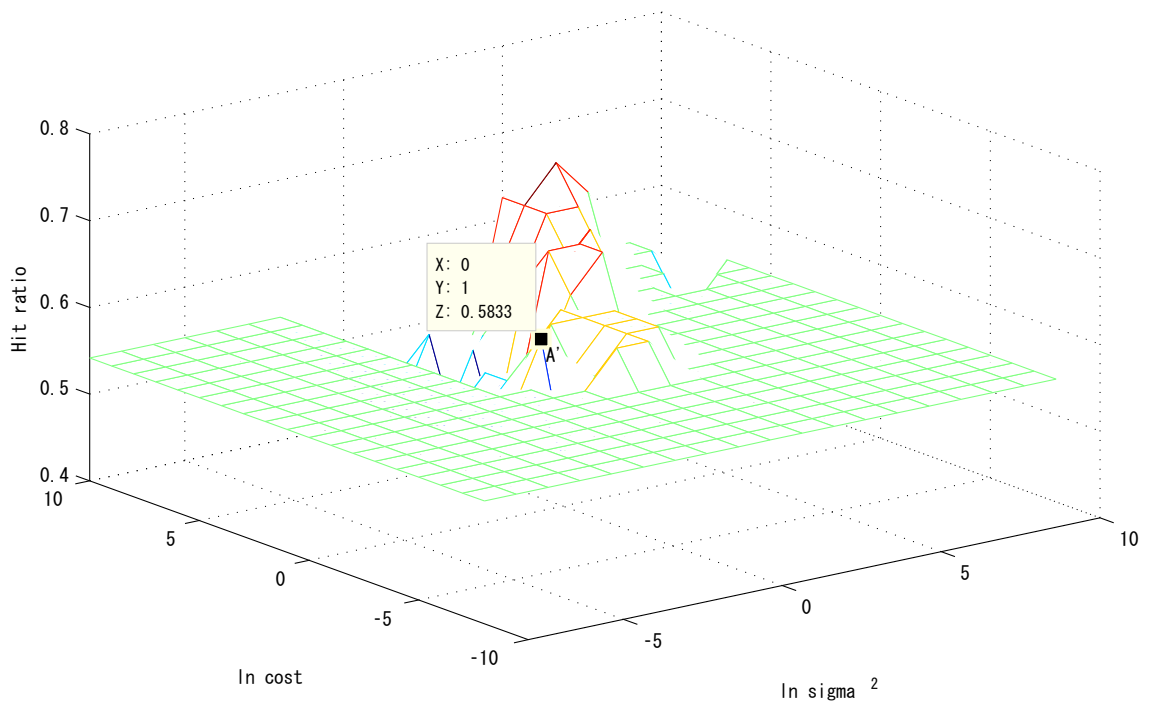


図 5.3: テストデータの識別成功率 (Case E)

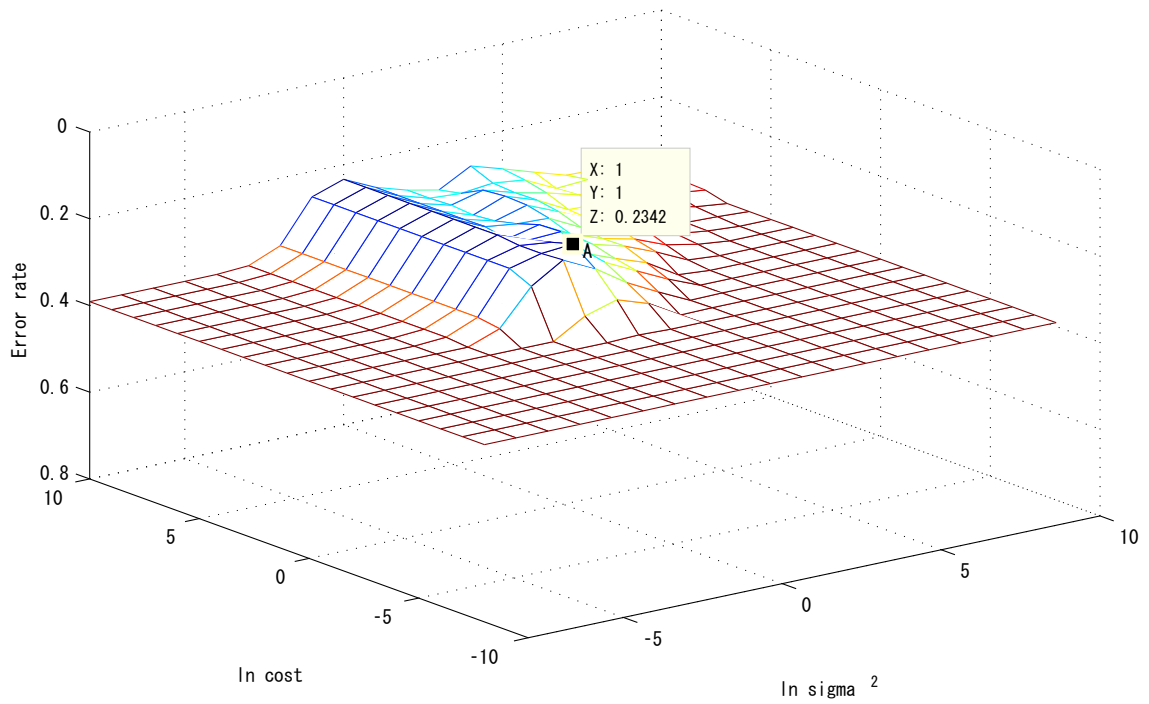


図 5.4: 訓練データの識別誤り率 (Case F)

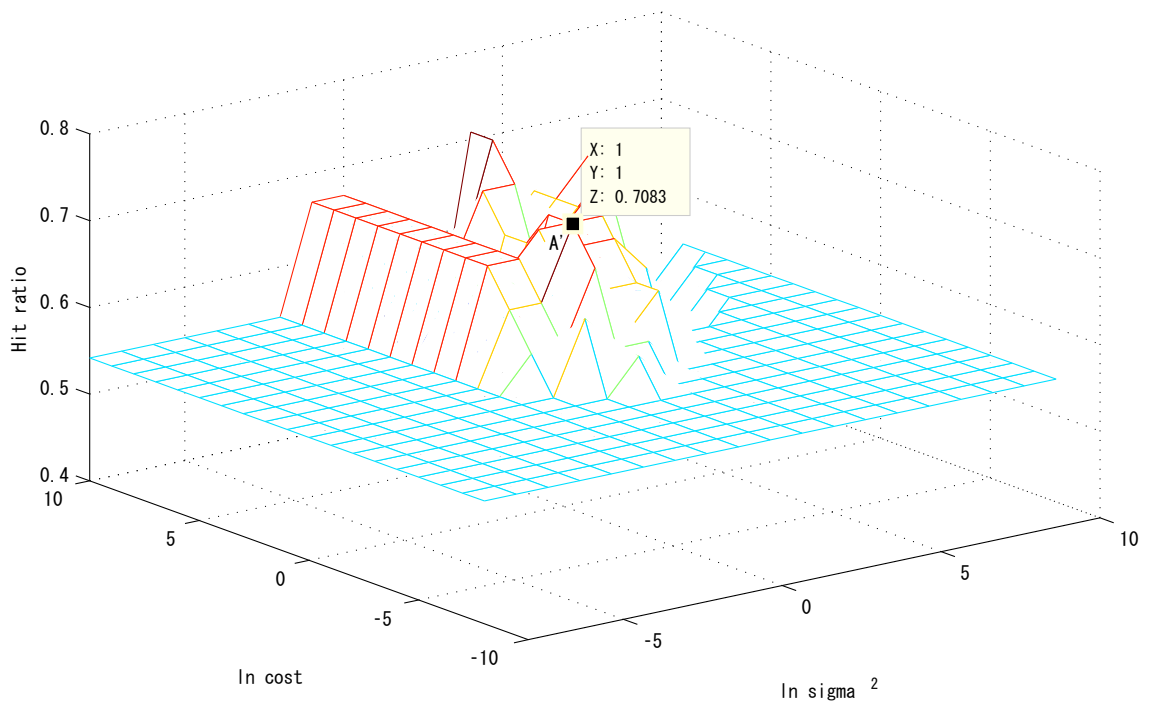


図 5.5: テストデータの識別成功率 (Case F)

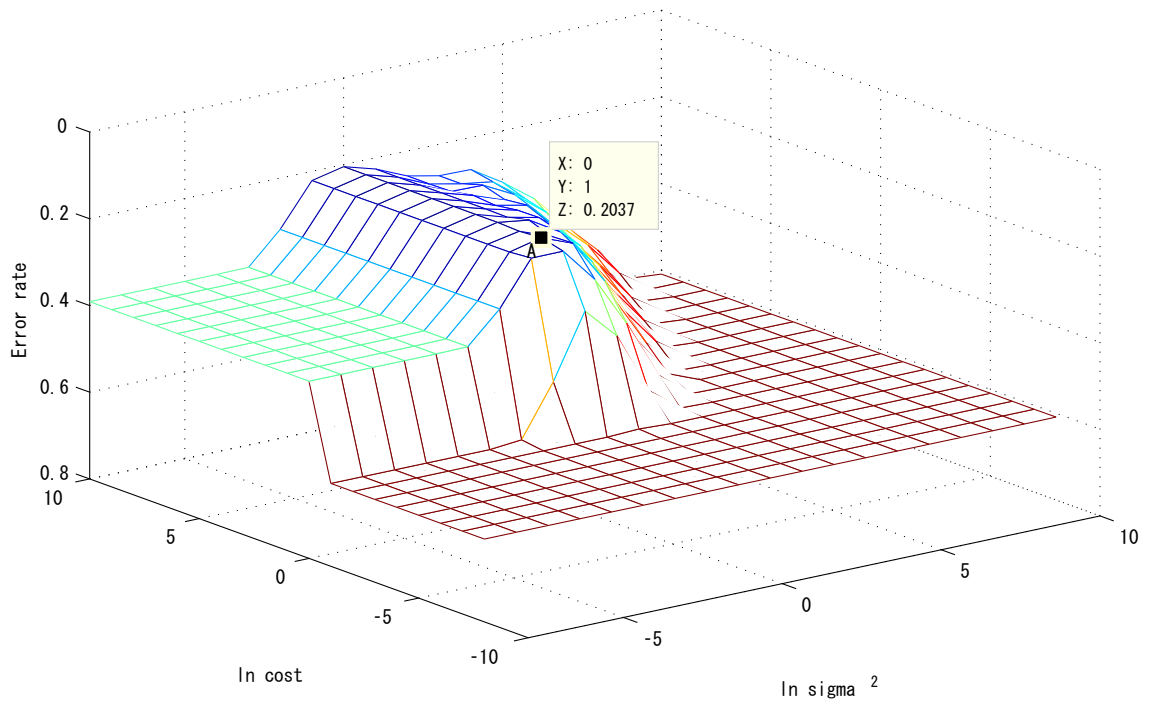


図 5.6: 訓練データの識別誤り率 (Case G)

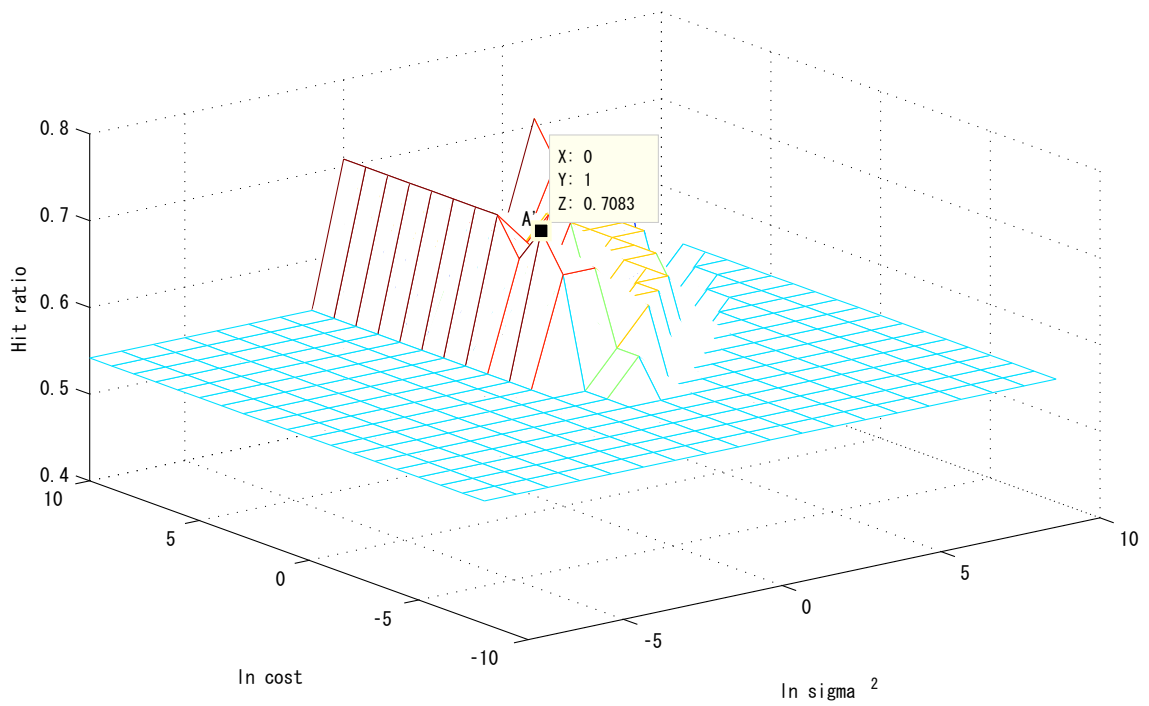


図 5.7: テストデータの識別成功率 (Case G)

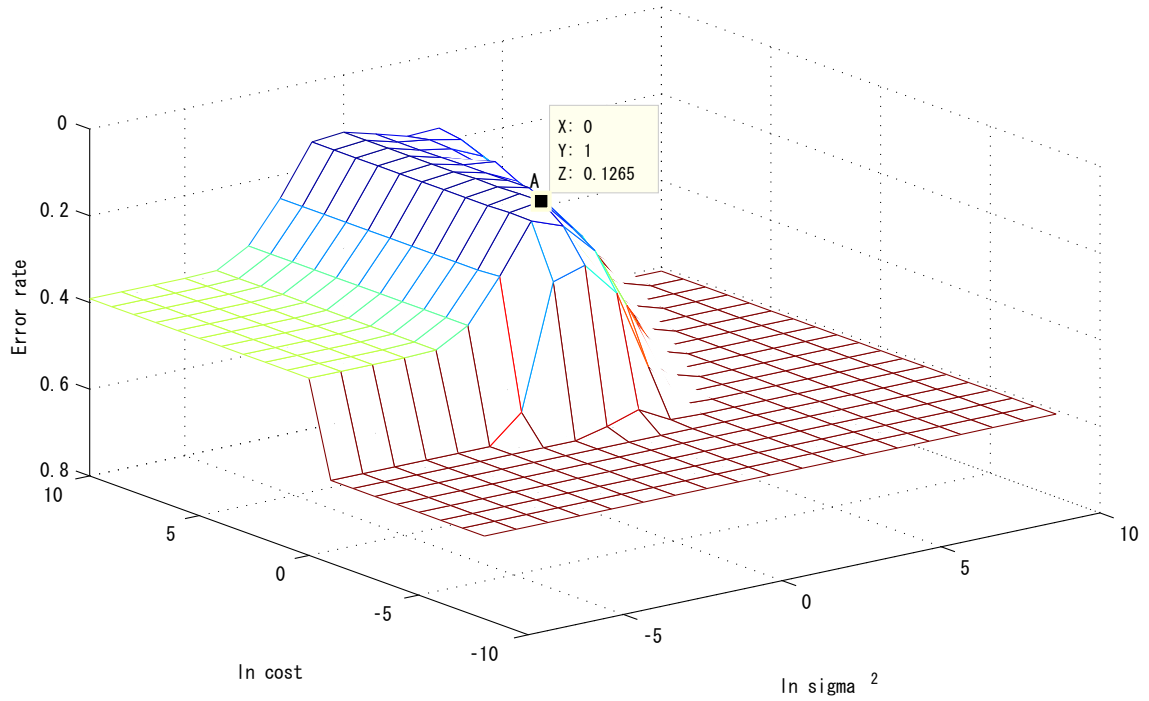


図 5.8: 訓練データの識別誤り率 (Case H)

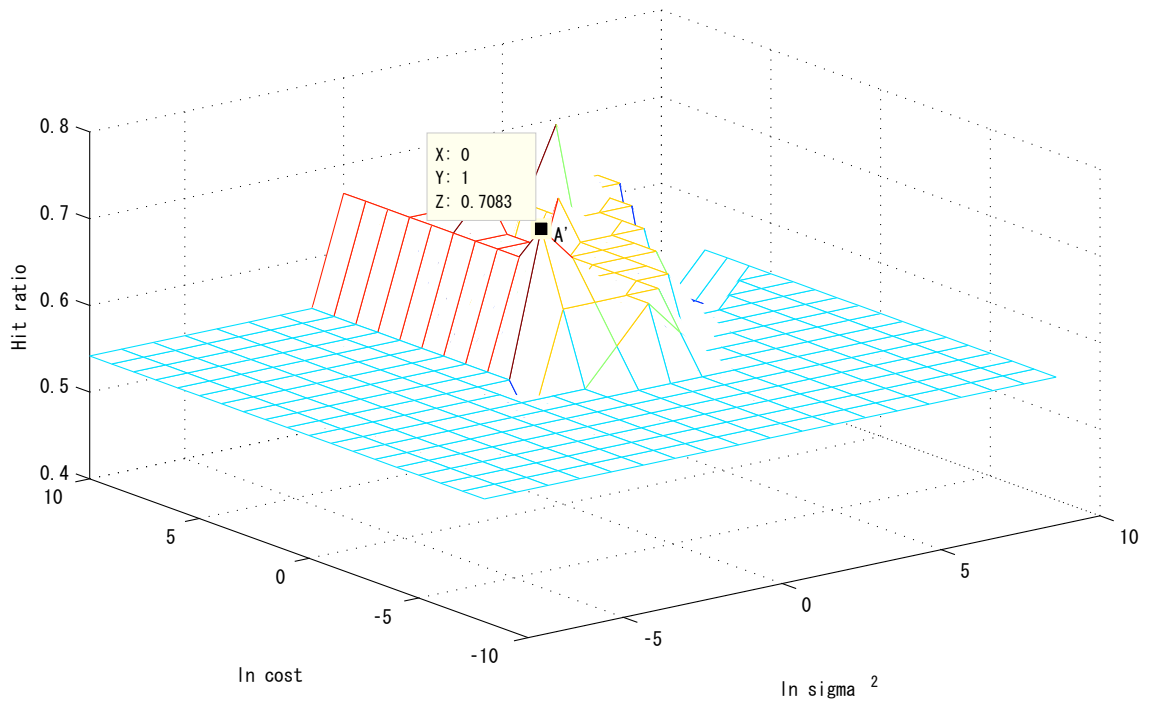


図 5.9: テストデータの識別成功率 (Case H)

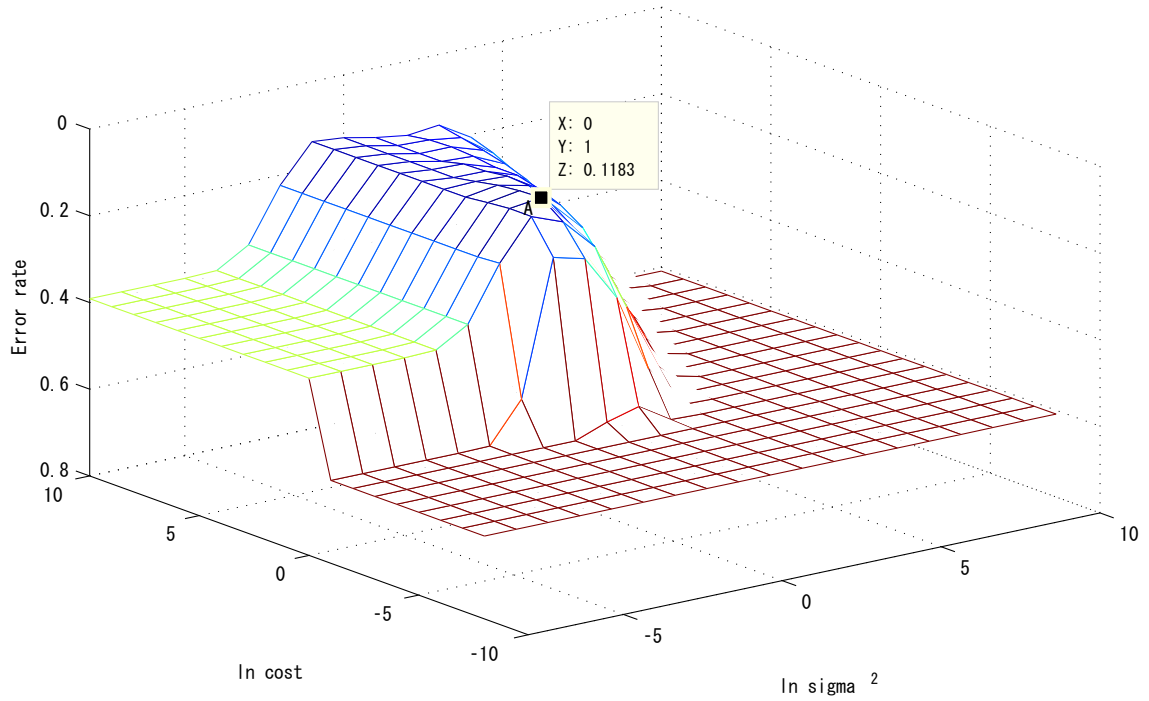


図 5.10: 訓練データの識別誤り率 (Case I)

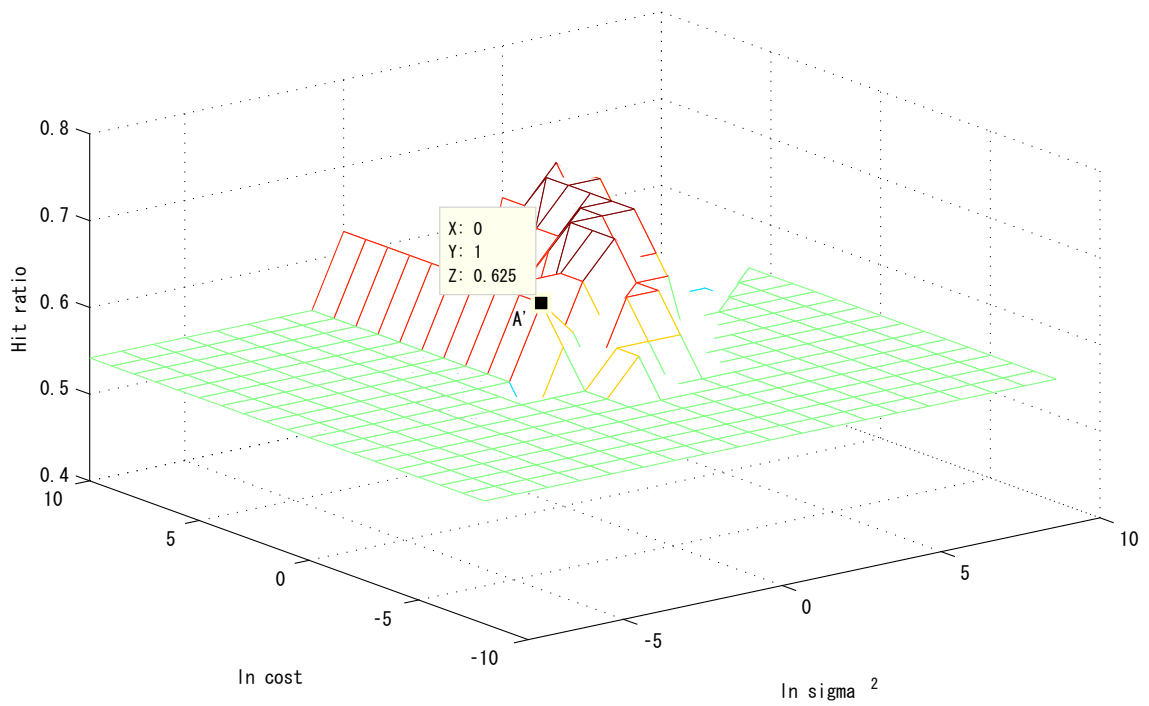


図 5.11: テストデータの識別成功率 (Case I)

表 5.4: パラメータチューニング後の Case F の予測結果

	Class A	Class B
Predicted Class A	12	6
Predicted Class B	1	5

表 5.5: パラメータチューニング後の Case G の予測結果

	Class A	Class B
Predicted Class A	11	5
Predicted Class B	2	6

5.2 擬似トレーディングによる評価

4.4 節と同様に日経平均株価の擬似トレーディングを行って、この予測の評価を行った。今回は常に売買手数料を考慮し、終値のみを用いた場合と終値を始値を用いた場合の 2 パターンについて考える。

5.2.1 売買手数料を考慮せず、始値と終値を用いた場合

二週間後の予測についても、4.4.2 節と同様に始値と終値を用いて以下のルールで擬似トレーディングを行った。

1. 最初の予算は 100 万円
2. 売買の判断は二週間ごとに行う。
3. 手元に株を持っていない場合
 - 二週間後の終値が当日の終値より「上がる」と予測された場合、翌日の始値で手元の資金が許す限り株を買う。
 - 二週間後の終値が当日の終値より「下がる」と予測された場合、株を買わず現金のままで資金を持ち続ける。
4. 手元に株を持っている場合
 - 二週間後の終値が当日の終値より「上がる」と予測された場合、そのまま株を持ち続ける。

表 5.6: パラメータチューニング後の Case H の予測結果

	Class A	Class B
Predicted Class A	12	6
Predicted Class B	1	5

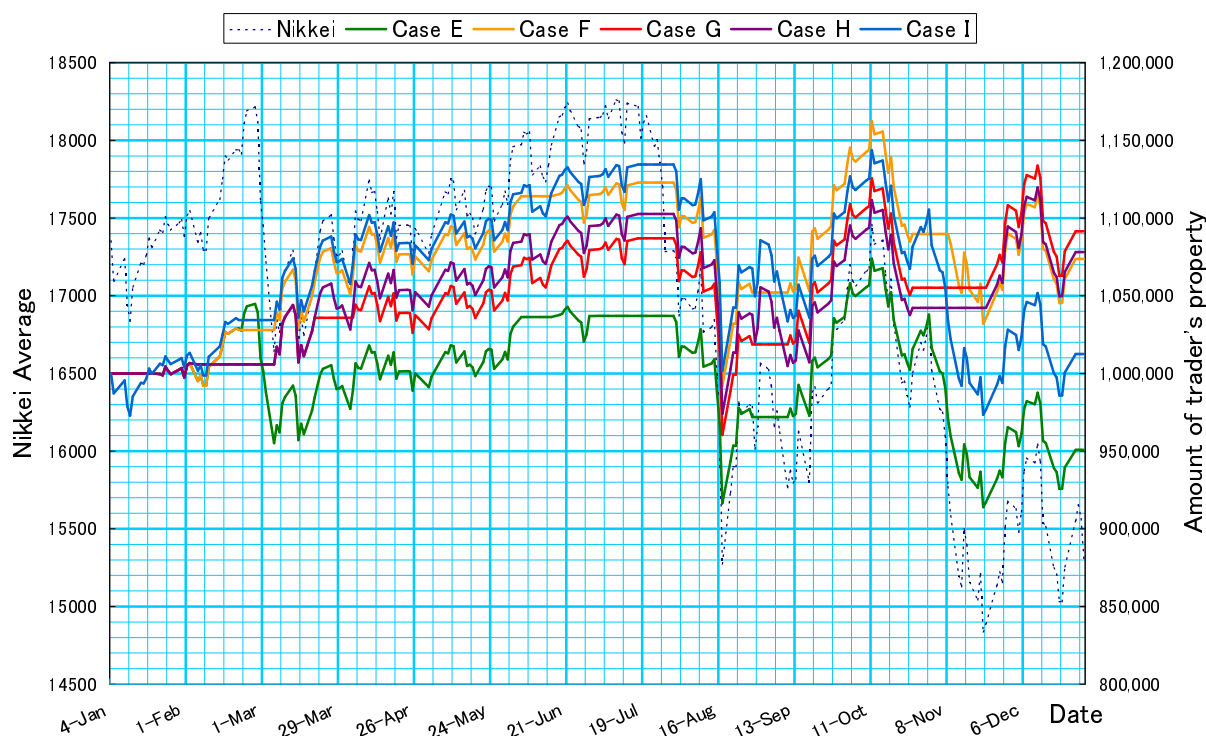


図 5.12: 手数料を考慮せず、始値と終値を用いた擬似トレーディング結果

- 二週間後の終値が当日の終値より「下がる」と予測された場合、翌日の始値で手元の株を全て売却する。

5. 売買の際に手数料はかからないものとする。

図 5.12 にこの擬似トレーディングの結果を示した。同じ識別成功率を持った Case G、Case H、Case I であるが 2% 程度リターン之差がある。全体の中で最も良いトレーディング結果を残したのは Case H で 9% 以上のリターンを得ている。

5.2.2 売買手数料を考慮し、始値と終値を用いた場合

次はさらに売買手数料を考慮し、以下のルールで擬似トレーディングを行った。

1. 最初の予算は 100 万円
2. 売買の判断は二週間ごとに行う。
3. 手元に株を持っていない場合
 - 二週間後の終値が当日の終値より「上がる」と予測された場合、翌日の始値で手元の資金が許す限り株を買う。
 - 二週間後の終値が当日の終値より「下がる」と予測された場合、株を買わず現金のままで資金を持ち続ける。
4. 手元に株を持っている場合

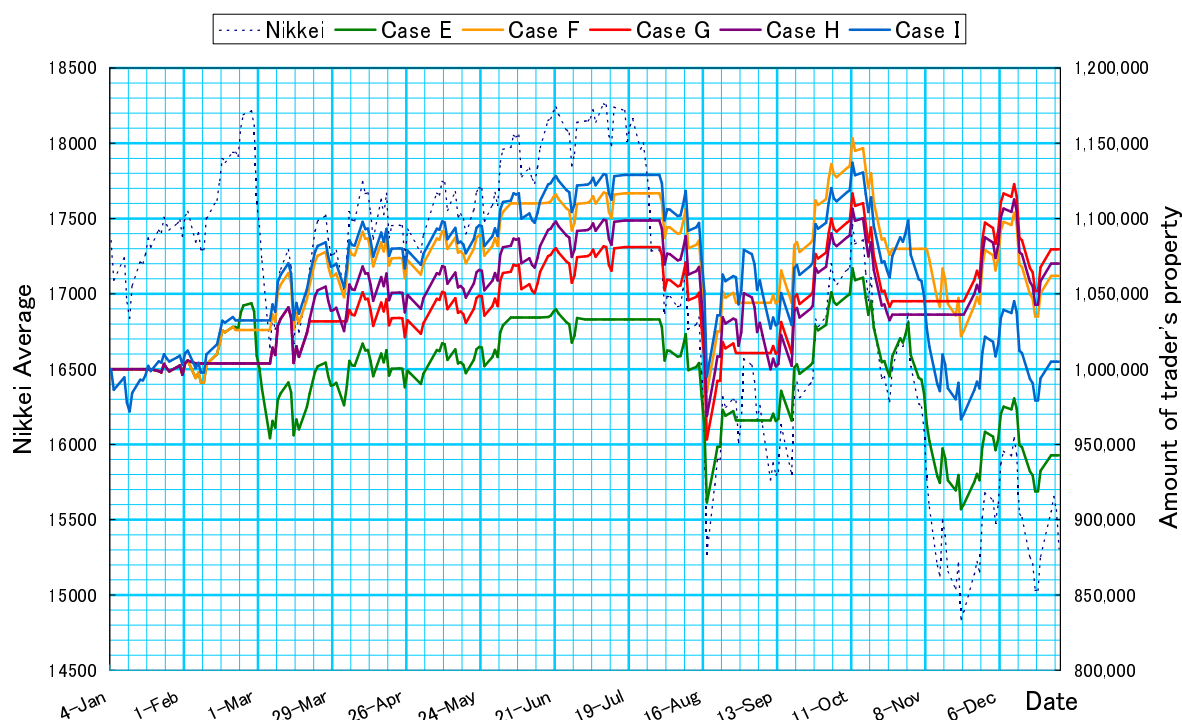


図 5.13: 手数料を考慮し始値と終値を用いた擬似トレーディング結果

- 二週間後の終値が当日の終値より「上がる」と予測された場合、そのまま株を持ち続ける。
- 二週間後の終値が当日の終値より「下がる」と予測された場合、翌日の始値で手元の株を全て売却する。

5. 売買の際に手数料一律 1000 円がかかるものとする。

この結果は図 5.13 の通りである。それぞれの Case のリターンの減少は 1% から 2% 程度であり、1 日の値の予測の場合より手数料をひかれることによるリターンの減少幅が小さいことがわかる。これはそもそも売買の回数が少ないことがその理由といえる。

また、最も良いトレーディング結果を残したのは Case H で約 8% のリターンを得ており、これは手数料を考慮した場合、一日後の値の予測のどの Case よりも高いリターンである。

第6章 結論

6.1 本研究のまとめ

本研究では時系列の予測において、パターン認識の一手法である SVM を用い、値動きの方向一致率を評価指標として短期及び中期の予測を行った。短期の予測である一日後の値の予測では最高で 68% の値動き方向一致率、一年間の擬似トレーディングでは最高で約 7% の利益をあげることができた。また、中期の予測である二週間後の値の予測では最高で 71% の値動き方向一致率、一年間の擬似トレーディングでは最高で一日後の値の予測の利益を上回る約 8% の利益をあげることができた。ただし擬似トレーディングにおいては、売買のタイミングや手数料の仮定が結果に対して大きく影響を与えている。

訓練データとテストデータは意図的にかなり性格の違う相場を選んでおり、またテストデータの期間中に日経平均株価が 12% 下落していることを考慮すると、短期および長期の予測はいずれも手法として汎用性が認められ、またインデックス運用に有用であることがいえる。そして最後に二つの予測を組み合わせた際の擬似トレーディングを行い、最大で約 15% の利益をあげられた。これにより、それぞれの予測が有用であることが確認された。

6.2 今後の課題

今後は、より投資家に役立つという観点から、ただ値の上下を予測するだけでなく大きな変動の前に警鐘を鳴らす役割を果たすような予測についても考慮していきたい。また、文献 [22] では、SVM において識別関数の超平面からの距離をシグモイド関数を用いて確率値に近似することの有効性が実験的に示されており、この確率を利用することで値の上下予測がどの程度信頼できるものなのを示すことができれば、より意思決定に役立つものになるはずである。

参考文献

- [1] 廣松毅, 浪花貞夫. 経済時系列分析. 朝倉書店, 1990.
- [2] Jovina Roman and Akhtar Jameel. Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns. In *Proc. of the 29th Annual Hawaii International Conference on System Sciences*, pp. 454–460, 1996.
- [3] Sam Mahfoud and Ganesh Mani. Financial forecasting using genetic algorithms. *Applied Artificial Intelligence*, Vol. 10, No. 6, pp. 543–565, 1996.
- [4] Hitoshi Iba and Takeshi Sasaki. Using genetic programming to predict financial data. In *Proc. 1999 Congress on Evolutionary Computation (CEC99)*. IEEE Press, 1999.
- [5] 馬場則夫, 西田将人, 甲斐良隆. ニューラルネットを活用した従来型テクニカル分析手法改善の試み. 電気学会論文誌 C, Vol. 126, No. 11, pp. 1324–1331, 2006.
- [6] Shigeo Mori, Kotaro Hirasawa, and Jinglu Hu. A stock price prediction model by using genetic network programming. In *Proc. SICE 2004 Annual Conference*, Vol. 2, pp. 1186–1191, 2004.
- [7] 打出義尚, 伊庭斉志. デイトレードのための遺伝的プログラミングによる売買ルールの生成. 情報処理学会, 第 69 回全国大会論文集, No. 2, pp. 213–214, 2007. 4N-8.
- [8] 葛西大佑, 伊庭斉志. 進化論的計算手法の株式売買ルール生成への応用. 情報処理学会, 第 69 回全国大会論文集, No. 1, pp. 295–296, 2007. 2M-2.
- [9] 納谷元子, 岡部洋一. ニューラルネットワークによる投資戦略支援システムの構築. 電子情報通信学会総合大会講演論文集, 2005 年総合大会講演論文集 情報・システム, No. 1, 2005. D-2-5.
- [10] 北川源四郎. 時系列解析入門. 岩波書店, 2005.
- [11] Royal observatory of belgium. <http://sidc.oma.be/html/sunspot.html>.
- [12] Yahoo! <http://www.yahoo.co.jp>.
- [13] Christopher J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, Vol. 2, pp. 1–47, 1998.
- [14] 津田宏治. サポートベクターマシンとは何か. 電子情報通信学会誌, Vol. 83, No. 6, pp. 460–466, 2000.
- [15] 前田英作. 痛快! サポートベクトルマシン 古くて新しいパターン認識手法 . 情報処理学会誌, Vol. 42, No. 7, pp. 676–683, 2001.

- [16] Nello Cristianini and John Shawe-Taylor. サポートベクターマシン入門. 共立出版, 2005. 大北剛 訳, (原書: An Introduction to Support Vector Machines and other kernel-based learning methods).
- [17] Wei Huang, Yoshiteru Nakamori, and Shu-Yang Wang. Forecasting stock market movement direction with support vector machine. *Computers and Operations Research*, Vol. 32, pp. 2513–2522, 2005.
- [18] Lean Yu, Shouyang Wang, and Kin Keung Lai. Mining stock market tendency using GA-based support vector machines. *Lecture Notes in Computer Science*, Vol. 3828, pp. 336–345, 2005.
- [19] Kyoung jae Kim. Financial time series forecasting using suport vector machines. *Neurocomputing*, Vol. 55, pp. 307–319, 2003.
- [20] 志水清孝. システム最適化理論. コロナ社, 1976.
- [21] 比較.com インターネット証券会社比較. <http://www.hikaku.com/sec/search/com.html>.
- [22] John C. Platt. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. In *Advances in Large Margin Classifiers*, pp. 61–74. MIT Press, 1999.

発表文献

中田貴之, 古関隆章. Support Vector Machine を用いた短期間の経済時系列予測. D-12-133. 電子情報通信学会 2007 年度全国大会

小谷道隆, 柴田知秀, 中田貴之, 黒橋禎夫. 日本語 Textual Entailment のデータ構築と自動獲得した類義表現に基づく推論関係の認識. E-5-4. 言語処理学会 第 14 回年次大会

謝辞

本研究を進めるにあたり、多くの方々のお世話になりました。深く感謝の意を表したいと思います。

指導教員の古関隆章准教授には研究に対する考え方や研究の方向性に対する助言など大変お世話になりました。ありがとうございました。

秘書の南佳子さんには研究室関係の事務作業でお世話になったほか、お隣の席で楽しく会話させていただきました。ありがとうございました。

技術職員の高田康宏さんは技術的な指導に加え、研究室内の明るい雰囲気をも作っていただきました。ありがとうございました。

博士課程の野崎雄一郎さん、同じく博士課程で現在ドイツへ留学中の福正博之さんには先輩として研究のやり方や論文・資料の書き方など様々なことを教えていただきました。ありがとうございました。

同期の鈴木武海君、原和弘君には、研究だけではなく普段の研究室生活から大変お世話になりました。二年間とても楽しかったです。ありがとうございました。

修士課程1年の小幡寛君、熊澤一将君、野田昂志君、Genevieve Marie Pattersonさん、そして卒論生の槻木澤佑公君、道念大樹君とは一年間でしたが同じ研究室の仲間として楽しい時を過ごすことができました。ありがとうございました。

その他研究室生活をともにしたSurosoさん、山口瑛史さん、大山大介さん、Cheung Cheuk Kei Brian君にも感謝の意を表したいと思います。ありがとうございました。

最後に、自分を支えてくれた家族と友人たちに感謝の意を表したいと思います。ありがとうございました。

平成20年2月4日

中田 貴之

付録A 補足理論

SVMの原理を説明する際に不可欠な理論について補足する。

A.1 双対問題と双対定理

一般的な線形計画の問題を考える。

$$\begin{aligned} \text{LP} \quad & \min_{\mathbf{x}} \quad \mathbf{c}^t \mathbf{x} \\ & \text{s.t.} \quad \mathbf{A} \mathbf{x} \geq \mathbf{b} \end{aligned} \tag{A.1}$$

この問題には対となる問題が存在する。対の問題の対はもとの問題になるという関係があるので、この両方の問題を互いに他の双対問題と呼んでいる。上記の問題 LP の双対問題は以下ようになる。

$$\begin{aligned} \text{LP-D} \quad & \max_{\mathbf{y}} \quad \mathbf{y}^t \mathbf{b} \\ & \text{s.t.} \quad \mathbf{y}^t \mathbf{A} = \mathbf{c}^t, \mathbf{y} \geq 0 \end{aligned} \tag{A.2}$$

もとの問題である主問題と、その双対問題については以下の定理が成立する。証明は省略する。

定理 A.1.1 (弱双対定理) 主問題の任意の可能解 \mathbf{x} と、双対問題の任意の可能解 \mathbf{y} に対して

$$\mathbf{c}^t \mathbf{x} \geq \mathbf{y}^t \mathbf{b}$$

が成立する。 ■

定理 A.1.2 (強双対定理) もし主問題か双対問題のどちらか一方に最適解が存在すれば、他方にも最適解が存在して、それぞれの最適解に対する目的関数値は等しい。すなわち

$$\min \mathbf{c}^t \mathbf{x} = \max \mathbf{y}^t \mathbf{b}$$

が成立する。 ■

A.2 ラグランジュ緩和

まず最初に、ラグランジュ緩和の一般形について、次のような最適化問題を用いて説明する。

$$\text{P} \quad \min_{\mathbf{x}} \quad f(\mathbf{x}) \tag{A.3}$$

$$\text{s.t.} \quad g_i(\mathbf{x}) = b_i \quad (i = 1, \dots, k) \tag{A.4}$$

$$h_j(\mathbf{x}) \geq b_j \quad (j = 1, \dots, l) \tag{A.5}$$

上記の問題 P がそのまま解けるような場合は、緩和問題を特に取り上げる必要は無い。しかし問題 P を解くのが困難な場合、これをもう少し簡単な問題に緩和して、この問題の手がかりを得るというアプローチが考えられる。

問題 P の制約式をどれか取り除いてしまった問題は緩和問題と呼ばれる。しかしながら、制約を取り去ってしまうと緩和の度合いが激し過ぎて、緩和問題の最適解が元の問題からかけ離れてしまうことが懸念される。そこで良く用いられる手法がラグランジュ緩和である。

例えば問題 P の等式制約を適当な数値 α_i 倍して目的関数に繰り込んだ問題

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) + \sum_{i=1}^k \alpha_i (b_i - g_i(\mathbf{x})) \\ \text{s.t.} \quad & g_i(\mathbf{x}) = b_i \quad (i = 1, \dots, k) \\ & h_j(\mathbf{x}) \geq b_j \quad (j = 1, \dots, l) \end{aligned} \tag{A.6}$$

について考える。問題 P の任意の許容解 \mathbf{x} に対し、上記の問題の目的関数である式 (A.6) の値と式 (A.3) の関数値は変わらない。ゆえにある解が P の最適解である必要十分条件は、その解が上記の問題の最適解となることである。ここで、上記の問題から等式制約を取り除いた問題

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) + \sum_{i=1}^k \alpha_i (b_i - g_i(\mathbf{x})) \\ \text{s.t.} \quad & h_j(\mathbf{x}) \geq b_j \quad (j = 1, \dots, l) \end{aligned}$$

はラグランジュ緩和問題と呼ばれる。

同様に問題 P の不等式制約を緩和することもできる。数値 $\beta_j (\geq 0)$ を用いて次のような問題を作れば良い。

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) + \sum_{j=1}^l \beta_j (b_j - h_j(\mathbf{x})) \\ \text{s.t.} \quad & g_i(\mathbf{x}) = b_i \quad (i = 1, \dots, k) \\ & h_j(\mathbf{x}) \geq b_j \quad (j = 1, \dots, l) \end{aligned} \tag{A.7}$$

この問題の目的関数値は許容解ならばその値は $f(\mathbf{x})$ より小さめに出ている。このことから、上記の問題は問題 P の緩和問題である。逆に言えば、この性質を保つために β_j は非負でなければならない。また目的関数の第二項は、この不等式制約を破る解についてのペナルティという形で目的関数に反映していると見ることが出来る。

上記の問題化から不等式制約を取り除いてしまったもの

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) + \sum_{j=1}^l \beta_j (b_j - h_j(\mathbf{x})) \\ \text{s.t.} \quad & g_i(\mathbf{x}) = b_i \quad (i = 1, \dots, k) \end{aligned}$$

がラグランジュ緩和問題である。

ラグランジュ緩和問題を作る際に用いた係数 α_i, β_j をラグランジュ乗数と呼ぶ。ラグランジュ乗数の選び方によって、ラグランジュ緩和問題の性能は大きく変わる。そこでできるだけ良いラグランジュ乗

数を求めたいと考える。例えば上記の不等式制約をラグランジュ緩和した問題において、最も良いラグランジュ乗数を求める問題は

$$\max_{\beta} \left[\min_{\mathbf{x}} \left(f(\mathbf{x}) + \sum_{j=1}^l \beta_j (b_j - h_j(\mathbf{x})) \mid g_i(\mathbf{x}) = b_i \quad (i = 1, \dots, k) \right) \right] \quad (\text{A.8})$$

$$s.t. \quad \beta \geq 0 \quad (\text{A.9})$$

という最適化問題となる。

ここで A.1 節の冒頭で提示した問題 LP に対してラグランジュ乗数ベクトル \mathbf{y} を用いてラグランジュ緩和すると

$$\min_{\mathbf{x}} \quad \mathbf{c}^t \mathbf{x} + \mathbf{y}^t (\mathbf{b} - \mathbf{A} \mathbf{x}) \quad (\text{A.10})$$

という制約の無い問題となる。この式を整理すると、

$$\min_{\mathbf{x}} \quad (\mathbf{c}^t - \mathbf{y}^t \mathbf{A}) \mathbf{x} + \mathbf{y}^t \mathbf{b} \quad (\text{A.11})$$

となる。ここで目的関数ベクトルにゼロではない係数が一つでも存在するならば、この問題の最適値はマイナス無限大となってしまう、緩和問題として用をなさない。そのため、ラグランジュ乗数 \mathbf{y} としては、 $\mathbf{c}^t - \mathbf{y}^t \mathbf{A} = \mathbf{0}^t$ を満たすものを選ぶのが望ましい。このとき、上記のラグランジュ緩和問題の最適値は $\mathbf{y}^t \mathbf{b}$ そのものとなる。

以上より、もっとも良いラグランジュ乗数を求めるには、

$$\begin{aligned} \max_{\mathbf{y}} \quad & \mathbf{y}^t \mathbf{b} \\ s.t. \quad & \mathbf{y}^t \mathbf{A} = \mathbf{c}^t, \mathbf{y} \geq 0 \end{aligned} \quad (\text{A.12})$$

を解けば良いことがわかる。A.1 節で示した通り、この問題こそが線形計画問題 LP の双対問題 LP-D に他ならない。そしてこの性質は 2 次の計画問題についても同様である。

問題 LP は不等式制約を持つ問題であるが、その双対問題である LP-D は等式制約と簡単な定義域の制約を持つものとなっており、一般的には問題 LP-D の方がより解きやすい。そして、強双対定理よりそれぞれの最適関数に対する目的関数値が等しいことが保証されている。これらの性質から、本文中でも不等式制約を持つ問題に対してラグランジュ緩和を行い、より解きやすい問題へと変換している。

付録B Nelder-Mead法

Nelder-Mead法とは、非線形最適化に対する直接探索手法の一つである。以下ではその概略とアルゴリズムについて説明する。

B.1 Nelder-Mead法の概略

独立変数 x_1, \dots, x_n を変数とする目的関数 $f(\mathbf{x})$ をある制約条件下で最小化する問題を考える。ただし、 x_{n+1}, \dots, x_m は独立変数の関数である。

Nelder-Mead法の考え方は、 $n+1$ 個以上を頂点とするシンプレックスを作り、このシンプレックスの頂点の中で最悪の目的関数値をとる点の残りの頂点の中心に関する鏡映点は目的関数値を改善させることが期待できるという考え方を取り、またこの鏡映プロセスのほかに目的関数の形状に応じて伸縮、収縮、縮小などの修正動作を加えて最適解の探索を行う手法である。

この手法は、最小化すべき目的関数を変数の関数形として明確に定義する必要が無いばかりか、目的関数の局所的な最小化方向を与える偏微分係数なども計算する必要がなく、ただ任意の変数に対応した目的関数値だけを算出、利用して最適点を探索できるという利点をもっている。

B.2 Nelder-Mead法のアルゴリズム

最初に記号を以下のように定義する。

- \mathbf{x}^i : シンプレックスの頂点 i の座標 (x_1^i, \dots, x_n^i)
- \mathbf{x}^l, f^l : シンプレックスにおける目的関数の最大点、および \mathbf{x}^l における $f(\mathbf{x})$ の値
- \mathbf{x}^h, f^h : シンプレックスにおける目的関数の二番目の最大点、および \mathbf{x}^h における $f(\mathbf{x})$ の値
- \mathbf{x}^s, f^s : シンプレックスにおける目的関数の最小点、および \mathbf{x}^s における $f(\mathbf{x})$ の値
- \mathbf{x}^g, f^g : 最悪点 \mathbf{x}^l を除いたすべての頂点の中心、および \mathbf{x}^g における $f(\mathbf{x})$ の値

シンプレックスの頂点の数を $k(n+1 \leq k)$ とする。 k 個の初期点のうち、1つは自らで与え、残りの $k-1$ 個は制約条件を満たすように、一様乱数を用いて決める。この初期点からなるシンプレックスを何度も更新し、シンプレックスの各頂点の目的関数値 $f(\mathbf{x}^i)$ の標準偏差が一定値を下回ったとき、最適点に収束したと考える。

以下では、シンプレックスの更新の手順について説明する。

シンプレックスの各頂点で目的関数値を比較して、最悪点 \mathbf{x}^l を除き残りの頂点の中心 \mathbf{x}^g に関して、鏡映係数 α を用いて、 $\mathbf{x}^k = (1+\alpha)\mathbf{x}^g - \alpha\mathbf{x}^l$ により鏡映させ、試点 \mathbf{x}^k を求める。鏡映によって得られたこの試点 \mathbf{x}^k の目的関数 f^k が f^s, f^h, f^l と比較してどこに分類されるか、ならびに制約条件を考慮してシンプレックスの伸張、収縮、縮小のいずれかの修正動作をとるかを定める。

1. $f^k \leq f^h$ の場合

x_k が新しいシンプレックスで少なくとも最悪点にならない場合である。

(a) $f^s < f^k \leq f^h$ の場合

x_l のかわりに x_k を採用して探索を続行する。

(b) $f^k < f^s$ の場合

x^k が新しいシンプレックスで最良点であるから、その方向で一層の改善が期待できる。そこで伸張係数 γ を用いて、 $x^e = (1 + \gamma)x^g - \gamma x^l$ によって伸張する。伸張点 x^e における目的関数 f^e を計算する。 $f^e < f^k$ であるとき、 x^l の代わりに x^e をとり、そうでない場合は、 x^k を採用し探索を続ける。

2. $f^h < f^k$ の場合

狭い谷や、最適点近傍で起こる状況で、試点 x^k が新しいシンプレックスにおいても最悪点になる場合である。

(a) $f^l \leq f^k$ の場合

行き過ぎた x^k を戻すのだが、 $f^l \leq f^k$ より x^l 側の内分点のほうがより小さい関数値を与えることが期待できるので収縮率 β を用いて、 $x^c = \beta x^l + (1 - \beta)x^g$ によって x^l 側に収束を行う。

$f^c < f^l$ であれば、 x^l のかわりに x^g を採用し、そうでない場合は全ての頂点を x^s に向かって $\frac{1}{2}$ ずつ移動させて縮小したシンプレックスをつくる。

(b) $f^k < f^l$ の場合

この場合は、 $x^c = \beta x^k + (1 - \beta)x^g$ によって x^k 側に収束を行う。

$f^c < f^l$ であれば、 x^l のかわりに x^g を採用し、そうでない場合は全ての頂点を x^s に向かって $\frac{1}{2}$ ずつ移動させて縮小したシンプレックスをつくる。