

# 製品検査における 1 クラス SVM のパラメータ調整手法

—— クラスタ形状の分析によるガウスカーネルパラメータの選択 ——

## Parameter Tuning Approach of One Class SVM for Product Inspection

— Gaussian Kernel Parameter Selection by Cluster Shape Analysis —

梶谷 和人 田崎 博 中嶋 宏

Kazuto Kojitani Hiroshi Tasaki Hiroshi Nakajima

オムロン株式会社 センシング&コントロール研究所

Sensing and Control Technology Laboratory, OMRON Corporation

**Abstract:** To realize automated product inspection with using single classification modeling, it is necessary to acquire the classification model by using small number of normal sample data. Mahalanobis Taguchi system requires enough number of sample data because the statistical estimation accuracy of data distribution has serious influence on its classification performance. On the other hand, SVM works well with small number of learning data. When the scale parameter of Gaussian kernel is too large, type-2 error will occur; otherwise, type-1 error. In this article, the selection method of Gaussian kernel parameter is proposed by analyzing the cluster shape formed by SVM.

## 1 はじめに

### 1.1 背景

工業製品の品質に対する要求が社会的に高まっている中で、最終検査工程において安定した全数検査を行う自動検査システムの開発が進められている。たとえば、自動車のエンジンやトランスミッションの異音検査装置では、マイクで取得した波形データから検査対象の品質特性を表す複数の特徴を抽出し、正常/異常の判断を行っている。

品質特性には、工程変動や材料に起因する目標品質からのばらつきがあるため、特徴空間において良品の分布はひとつのクラスタをなすと考えられる。一方、不良品は多種多様なので、ひとつのクラスタをなすとは考えにくい。このような非対称な判別問題には、たとえばマハラノビス・タグチ・システム(MTS)のような 1 クラス判別モデルが用いられる。

MTS は、学習サンプルの中心からのマハラノビス距離を識別関数とするため、事前に学習サンプルから分布パラメータを推定しておく必要がある。しかし、実際の製造現場では、たとえば製造ラインの立ち上げ時には、十分なサンプルを収集することができないことが多い。さらに、MTS は学習サンプル数が特徴数以下の場合、正確多重共線性が発生し、原理的にマハラノビス距離を求めることができない。

田崎ら[1]は、少数サンプルからの学習に 1 クラス SVM[2]を使用することで、このような状況下でも判別性能を確保する方法を提案している。1 クラス SVM の識別関数は、学習サンプルから選ばれた少数のサンプル (SV: Support Vector) の関数であり、特徴空間上には関数値が 0 以下となる「サポート (Support)」と呼ばれる領域が構成される。つまり、サポート内を良品、サポート外を不良品と判定す

ることで検査を行う。1 クラス SVM は分布を仮定しないノンパラメトリックな手法なので、パラメータ推定精度に起因する性能低下はなく、学習サンプル数に対する制約もない。

### 1.2 目的

ガウスカーネルパラメータ  $\sigma$  が十分大きい場合、サポートは単一のクラスタとなるが、 $\sigma$  が小さすぎる場合、複数クラスタに分割される可能性がある (図 1)。

工業製品の製造工程は、品質特性の平均値が目標品質を実現するように調整されるのが通常なので、良品は学習サンプルの中心を含む単一クラスタを構成すると考えるのが自然である。

本研究の目的は、単一クラスタで構成されるサポートを実現するガウスカーネルパラメータ  $\sigma$  を求め、製品検査における 1 クラス SVM の判別性能を確保することである。

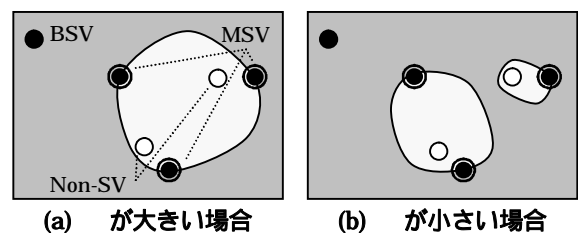
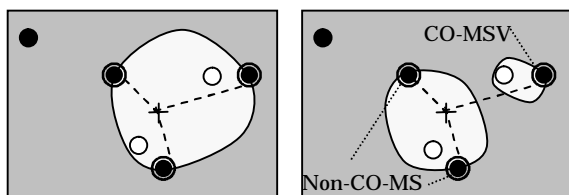


図 1 パラメータ とサポートの形状

## 2 ガウスパラメータ $\sigma$ の調整方法

### 2.1 サポートの形状の分析

特徴空間上の任意の 2 点 ( $x_i, x_j$ ) が別のクラスタに属していれば、それらを結ぶ線分にはサポートを逸脱する区間が存在する[3]。つまり、クラスタ形状が凸であると仮定すれば、識別面と交差する線分がなければ、2 点が同じクラ



(a) 単一クラス (b) 複数クラス  
図 2 中心-MSV 間線分によるサポート形状分析

スタに属していると判定できる。本研究では、識別面との交差の有無を判定する線分として、学習サンプルの中心と識別面上のサポートベクタ (MSV: Margin Support Vector)  $(\bar{x}, x_i^{MSV})$  を結ぶ線分だけを取り上げ、良品クラスタの形状として適当であるかどうか判定に用いる (図 2)。

## 2.2 $\sigma$ の選択基準

よく知られた選択基準として、汎化誤差の最小化を意図した「SV 数最小化基準」がある[4]。一方、提案する基準はサポート形状に着目したものであり、学習サンプルの中心へ至る線分が識別面と交差する MSV(CO-MSV: Crossover MSV)の数が 0 となる最小の  $\sigma$  を選択する。

## 3 シミュレーション実験

### 3.1 実験条件

実験には「Water Pump Dataset」 [5] を使用した。入力には正常なポンプの振動波形データから抽出した 26 の特徴を用いた。少数サンプルを想定し、学習サンプル数  $n = 25, 50$  の 2 通りについて 1 クラス SVM を学習し、別の 96 の正常サンプルと 320 台の異常サンプルで判別性能を評価した。ソフトマージンパラメータ  $\nu$  は 0.05 に固定し、ガウスカーネルのパラメータ  $\sigma$  を変化させながら、SV 数と、CO-MSV 数を測定した。なお、線分の識別面交差の判定には、線分上の 20 点での識別関数値を求め、いずれの点も正である場合に「交差なし」と判断した。

### 3.2 実験結果

図 3 は、 $\sigma$  を変化させたときの SV 数( $n_{SV}$ )、CO-MSV 数( $n_{CO}$ )、オッズ比(Odds Ratio)を示したものである。オッズ比は、良品を不良品と判定する第 1 種誤りと不良品を良品と判定する第 2 種誤りの 2 つの誤判別を考慮した判別性能指標で、以下の式で与えられる。

$$\gamma = 10 \log \frac{n_{11}n_{22}}{n_{12}n_{21}} \quad (1)$$

ただし、 $n_{11}$ 、はそれぞれ良品、不良品を正しく判定した数、 $n_{12}$ 、 $n_{21}$  はそれぞれ、第 1 種、第 2 種誤りの数である。

いずれの学習サンプル数  $n$  においても、オッズ比のピークは SV 数が比較的大きい領域にあり、SV 数最小化基準では適切な  $\sigma$  が選択できないことがわかる。提案手法によって選択された  $\sigma$  (点線) は、オッズ比のピークの近傍に位置し、比較的高い判別性能が実現できることがわかる。

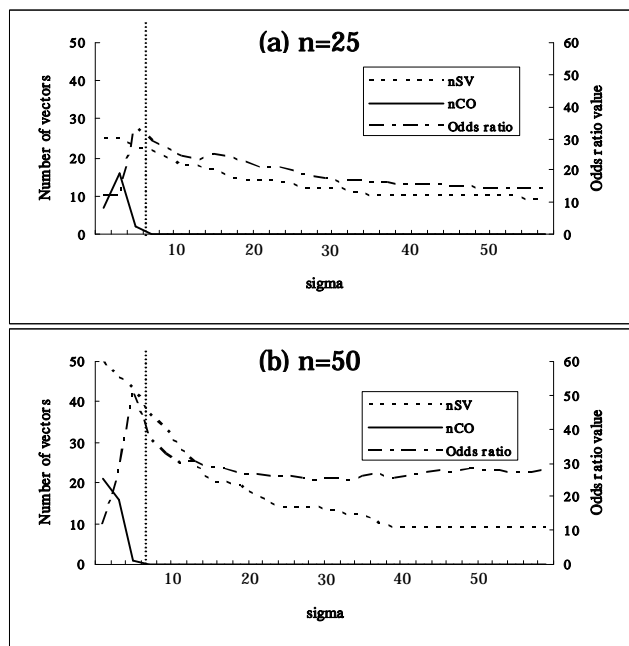


図 3 パラメータ選択指標と判別性能の変化

## 4 まとめ

本稿では、サポート形状を分析することで 1 クラス SVM のガウスカーネルパラメータ  $\sigma$  を決定する手法を提案し、実験によって製品検査における判別性能確保に効果があることを確認した。今後は、BSV 数最大化[5]等他の基準との比較、他のデータセットでの検証、識別面交差判定アルゴリズムの効率化を行い、より実用的な手法となるよう改善を行う。

### 【参考文献】

- [1] 田崎博、梶谷和人、中嶋宏、“自動検査のための漸進的判別モデル更新手法”，第32回知能システムシンポジウム講演論文集, pp. 243-246 (2005).
- [2] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol.13, no.7, pp. 1443-1471 (2001)
- [3] Asa Ben-Hur, David Horn, Hava T. Siegelmann, Vladimir Vapnik, "Support Vector Clustering", *Journal of Machine Learning Research* 2, pp. 125-137 (2001).
- [4] Nello Cristianini, John Shawe-Taylor, "An Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods", Cambridge University Press (2000) (邦訳: サポートベクタマシン入門, 共立出版)
- [5] Runar Unnthorsson, Thomas Philip Runarsson and Magnus Thor Jonsson, "Model selection in one class nu-SVMs using RBF kernels" 16th conference on Condition Monitoring and Diagnostic Engineering Management, 27-29 August (2003). <http://www.hi.is/~runson/svm/>