

サポートベクターマシンを用いた気圧配置検出手法の提案

——西高東低冬型を対象として——

木村広希*・川島英之**・日下博幸***・北川博之**

* 筑波大学大学院生, ** 筑波大学システム情報工学研究科 (計算科学研究センター),

*** 筑波大学生命環境科学研究科 (計算科学研究センター)

気候学研究においては、特定の気圧配置を示す事例を選ぶ必要から、過去の気圧配置を分類することがある。気圧配置の分類は多くの場合、目視で行われており、長期間かけて行うことや複数人で行うことがある。これらの場合、判断にぶれが生じたり、分類結果が研究者の主観に左右されたりする可能性がある。本稿では、この気圧配置の分類に、パターン認識手法であるサポートベクターマシンを用いて、分類を自動化することを提案する。本研究では、日々の気圧配置からの西高東低冬型の検出を目的とし、実験により提案手法の有用性を検討した。JRA-25 のデータを用い、1981~1990 年を学習期間、1991~2000 年を検証期間として実験を行った結果、最良で 90% 以上の適中率で西高東低冬型を検出できることがわかった。これより、学習データの与え方に主観が影響するものの、分類の際の負担を軽減でき、判断基準が変わることなく、ある程度の精度で気圧配置を分類できると考えられる。

キーワード：気圧配置、西高東低冬型、サポートベクターマシン、分類、地上天気図

I はじめに

気候学の研究において、気圧配置は気候の形成を動的にとらえるための重要な基礎となり、その分類を行うことによってこれまでに多くの成果を得てきた。たとえば、気圧配置ごよみの研究 (吉野・甲斐 1975; 甲斐・吉野 1975) や、それに基づく気候変動の研究 (山川 1988, 1989) などがある。気圧配置に関する研究では、西高東低冬型や南高北低夏型などの、ある特徴を持つ気圧配置の事例を、過去の事例から多数分類することがある。この気圧配置の分類は現状では多くの場合、天気図を用いた目視による判別で行われている。このとき、気圧配置の分類は研究者の主観に左右される可能性があり、また、対象とする事例の数が多いほど、この目視による判別という作業は困難なものとなる (宮・日下 2009)。長期間にわたって少しずつ分類することや、作業者を多くすることでこの負担は軽減できる。しかし

ながら、長期間にわたって分類する際には作業時間の経過に伴って判断にぶれが生じる可能性がある。また、作業数を多くする場合には、作業員間における判別基準が必ずしも同一でないという問題が生じる。これらは、とりわけ気候変動の研究の際に大きな問題となる。

一方で、近年計算機の性能が急速に向上し、大量のデータの解析が可能となった。これに伴い、自然科学分野では、蓄積されたデータ資源を有効に活用したさまざまな科学的検証が行われている。気候学分野においても、再解析データなどの格子点データの普及が進み、計算機を用いた大規模データの統計解析や、数値シミュレーションなどが行われている。

そこで、本稿では気圧配置の分類に対してコンピュータサイエンス分野のパターン認識手法を適用することにより、日々の気圧配置を自動的に分類する方法を提案する。これにより、学習データの与え方に主観は残るものの、困難な作業はなくなり、分

類作業の途中で判別基準が変わるという問題も解決できる。

本研究では、日々の気圧配置から西高東低冬型を検出することを目的とし、その手法としてパターン認識手法の一つであるサポートベクターマシン（以下SVM）を用いる。そして、実験により本手法の有用性を検討する。SVMの計算には、SVMのライブラリであるLIBSVM（Chang and Lin 2000）を用いる。

II 対象とする気圧配置と使用データ

1. 気圧配置

西高東低冬型は、南岸低気圧型や日本海低気圧型などと比べて、時間変化が小さく、1時点の天気図からでも気圧配置型をより判断しやすいと考えられる。さらに、西高東低冬型は、日本の西側には高圧部、東側には低圧部が広範囲に存在しているため、気圧分布の特徴も判別しやすいと考えられる。そのため、本研究では西高東低冬型を検出対象とし、日々の気圧配置を、西高東低冬型と非西高東低冬型の2種類に分類することを試みる。

2. 使用データ

パターン認識手法を用いて日々の気圧配置を分類し、その結果の正誤を検討するためには、日々の気圧データと各日の気圧配置型名称データ（例：西高東低冬型）が必要となる。

日々の気圧データにはJRA-25（Onogi et al. 2007）を用い、1981年から2000年における毎日の日本時間9時のデータを対象とする。また、各日の気圧配置型名称データには、吉野（2002）により分類されたデータを用いる。これは、本研究で用いる気圧データと同時刻の気圧配置を分類したものである。

JRA-25の海面更正気圧データは、南北1.25度・東西1.25度メッシュの全球データである。本研究

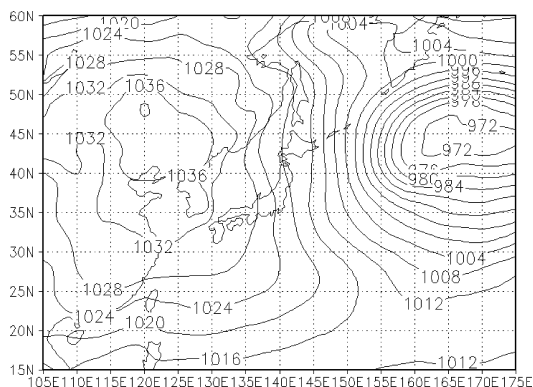


図1 対象領域の海面更正気圧（2006年12月30日9時、西高東低冬型）

Fig. 1 Sea-level pressure in the study area at 09:00 JST December 30, 2006 (winter type)

では、西高東低冬型を特徴づけるシベリア高気圧とアリューシャン低気圧の中心が含まれる、日本周辺地域の気圧データ（ $37 \times 57 = 2,109$ 点）を対象とする。この地域は北緯15~60度・東経105~175度である。図1に、西高東低冬型の事例について対象領域の海面更正気圧分布を示す。

III 研究方法

1. サポートベクターマシン（SVM）

パターン認識とは、観測されたパターンをあらかじめ定められた複数の概念のうちの一つに対応させる処理を示す（石井ほか1998）。これは本研究において、ある気圧データを西高東低冬型か非西高東低冬型の二つのクラスに分類することに対応する。パターン認識の一手法であるSVMは、このような2クラスの分類を行う手法であり、文字認識や音声認識などに幅広く用いられている（津田2000；岡部・佐藤2003）。

本研究では分類にSVMを用いるが、SVMは既存手法であるため、ここでは概要を述べる。詳細については参考文献（津田2000）を参照されたい。

図2にSVMによるクラス分類の概念を示す。SVMは、与えられた学習データの中の、サポートベ

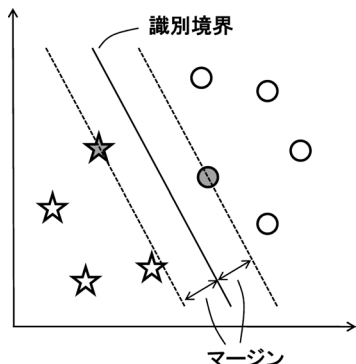


図2 SVMによるクラス分類の概念
○がクラス1の学習データ、☆がクラス2の学習データを示し、網掛けされている学習データがサポートベクトルを示す。

Fig. 2 Concept of classification using SVM

○: training data in class 1 ☆: training data in class 2 shaded training data: support vector

クトル（識別境界近傍に位置する学習データ）と識別境界との距離であるマージンを最大化するように識別境界を構築しクラス分類を行う。図2において、○がクラス1の学習データ、☆がクラス2の学習データを示す。各クラスの網掛けされている学習データがサポートベクトルとなる。ここで、本研究では、毎日9時における2,109点の気圧値を2,109次元ベクトルとして扱い、SVMによる分類を行う。

図2のような線形分離ができない場合には、非線形分離が行われる。図3に非線形分離の概念を示す。これは、非線形な写像 Φ を用いて学習データのベクトルが持つ次元の空間をより高次の空間に写像し、写像先の空間で線形分離を行うことで分離を容易にすることを意味する（津田 2000）。写像 Φ を用いた計算は、本来であれば元の次元より高次元でのベクトル計算を行う必要があるが、カーネルトリックという手法を用いることで高次元でのベクトル計算を避けることが可能となる（津田 2000）。ここで、カーネルトリックとは、 $K(x, x') = \Phi(x) \cdot \Phi(x')$ という高次の空間での内積であるカーネルに対して、

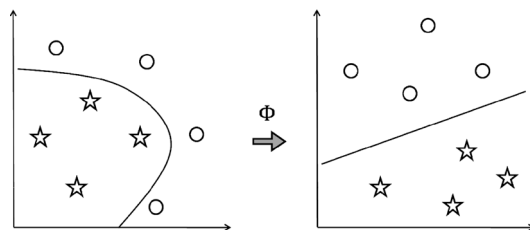


図3 非線形分離の概念（津田（2000）による）
非線形な写像 Φ を用いることで分離が可能になる。

Fig. 3 Concept of nonlinear separation (by Tsuda, 2000)

Nonlinear mapping Φ enables separation of space.

写像 Φ に関する計算 $\Phi(x) \cdot \Phi(x')$ を、カーネル関数 $k(x, x')$ を用いて置き換えることで、カーネル関数のみの計算に変換することを示す（津田 2000）。

カーネル関数は複数あるが、本研究では、代表的なカーネルの一つであり、SVMの計算でよく使われるRBF（Radial Basis Function）カーネルを用いた。以下に、RBFカーネルを示す。

$$K(x, x') = \exp(-\delta \times |x - x'|^2) \quad (1)$$

x, x' はベクトルを表し、 δ はパラメータを表す。計算過程では、 x, x' には2,109次元ベクトルの気圧データが代入される。また、RBFカーネルを用いる際には、パラメータ δ に適切な値を与えることが必要であり、本研究で用いた値は後述する。

以下に、SVMを用いた分類の実行手順を示す。

① 1日分の気圧データ（2,109次元ベクトル）と気圧配置（西高東低冬型または非西高東低冬型）を1組のデータとし、このデータの日数分の集合を学習データとする。

② SVMにより、学習データから識別関数を生成する。

③ 生成した識別関数を用いて、検証データを分類する。

2. 学習データ

SVMによる分類は、学習データから生成された識

別関数で行うため、異なる学習データを用いた場合は分類結果に差が生じると考えられる。そこで、SVM の学習に用いる学習データを 2 種類用意し、それぞれの学習データによる SVM の分類結果を比較する。

まず、SVM で西高東低冬型の分類を行う上で、日々の気圧配置が正例（西高東低冬型）と負例（非西高東低冬型）のどちらに属するかを示すデータが必要となる。本研究では、このデータを吉野（2002）から得た。吉野（2002）では基本的な気圧配置を 15 種類¹⁾に分類している。また、15 種類の型のいずれか一つの型に定まらず、複数の型を持つものを移行型もしくは複合型と分類している。本研究では、吉野（2002）による分類をもとに、表 1 のように気圧配置を区分する。

ここで、吉野（2002）において西高東低冬型に分類されているものを α 、西高東低冬型と他の型から成る移行型・複合型に分類されているものを β 、そして α と β に含まれないものを γ とする。この $\alpha \sim \gamma$ に基づいて、表 2 に 2 種類の学習データを示す。

学習データ 1 は、 α と β を「西高東低冬型」（正例）としている。この学習データを用いることにより、分類に際して、西高東低冬型の特徴を持つ事例

表 1 吉野（2002）に基づくデータの定義
Table 1 Definition of the data based on Yoshino (2002)

α	西高東低冬型
β	西高東低冬型と他の型から形成される移行型・複合型
γ	α と β に含まれないもの

表 2 学習データ
Table 2 Training data

学習データ	正例	負例
1	α, β	γ
2	α	β, γ

を幅広く検出することが期待される。

また、学習データ 2 は、学習データ 1 とは異なり、 β を「非西高東低冬型」（負例）として扱う。これにより、分類の結果として、移行型や複合型ではない、純粋な西高東低冬型が検出できると考えられる。

3. 気圧データの次元削減

本研究で用いる気圧データの格子点数は $37 \times 57 = 2,109$ 点であるため、SVM の計算では、2,109 次元という高次元のベクトル演算を行う。この計算を行う上で、次元数、すなわち格子点数を少なくすることができれば、計算時間を短縮できると考えられる。そこで、分類精度を保ちながら、どの程度まで格子点数を減らすことができるかを調べるために、気圧データの次元削減を行う。この次元削減とは、格子点の集合ごとに平均値を計算し、計算した平均値の集合を新たなデータとするものである。

図 4 に気圧データの次元削減の例を示す。図 4 は縦 10 × 横 13 の格子点を、縦 3 × 横 4 の 12 点の格子点に変換している。まず、格子点データの左上から右に向かって縦 $n \times$ 横 n の n^2 点の格子点集合を複数生成する。一番右の列まで達したら、次の行に移り同じ操作を繰り返す。一番下の行まで終わった後、生成された格子点集合ごとに平均値を計算し、その平均値の集合を新たな気圧データとする。ただし、一番右の列や一番下の行において端数となった点は、新たな格子点の生成の計算には用いないもの

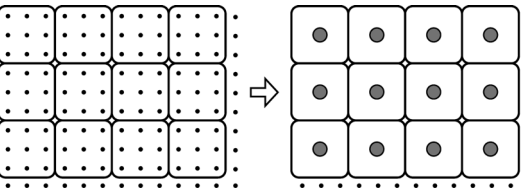


図 4 気圧データの次元削減
縦 10 × 横 13 の格子点から縦 3 × 横 4 の格子点への変換を示す。

Fig. 4 Dimension reduction of pressure data
Transformation from 10 × 13 to 3 × 4

とする．次元削減のパラメータとして n の値を変えることで新たに生成されるデータの格子点数を変化させ、分類結果の比較を行う．

IV SVM を用いた分類の評価実験

1. 実験データと手順

JRA-25 の 1981～1990 年のデータを学習データ，1991～2000 年のデータを検証データとして用いて分類実験を行った．学習データは 10 年間（1981～1990 年）の 3,652 日であり， α ， β ， γ に該当する日数はそれぞれ 464 日，266 日，2,922 日である．また，検証データは 10 年間（1991～2000 年）の 3,653 日であり， α ， β ， γ に該当する日数はそれぞれ 375 日，385 日，2,893 日である．

SVM のカーネルは線形と非線形の 2 通りで分類実験を行い，非線形カーネルには RBF カーネルを用いた．以下の非線形カーネルは RBF カーネルを指す．RBF カーネルのパラメータは， 1.0×10^{-N} として N を 1 から 7 まで変化させて分類を行い，誤分類の数（表 3 における $B+C$ ）が最も少なかったものを用いた．また，SVM の計算には LIBSVM (Chang and Lin 2000) を用いた．以上より，次の手順で SVM による分類実験を行った．

① 1981～1990 年の学習データ（日々の気圧データと正例と負例を示すデータ）から，SVM を用いて識別関数を生成する．

② 生成した識別関数を用いて，検証データ（1991～2000 年）を分類する．

③ 分類結果の正誤評価を行い，比較・検証する．

2. 分類結果の正誤評価

学習データ 1・2 を与えて分類した結果の正誤評価を表 3 に示す．表 2 と同様に，学習データ 1 では正例（西高東低冬型）を α と β ，負例（非西高東低冬型）を γ とし，学習データ 2 では正例（西高東低冬型）を α ，負例（非西高東低冬型）を β と γ とし

表 3 学習データを与えて分類した結果の正誤評価
Table 3 Evaluation of correction for results of classification using training data

		分類結果		
		表 2 の正例	表 2 の負例	計
正解	表 2 の正例	A	B（見逃し）	$N1 = (A+B)$
	表 2 の負例	C（空振り）	D	$N2 = (C+D)$
	計	$M1 = (A+C)$	$M2 = (B+D)$	$N = (A+B+C+D)$

てそれぞれ正誤評価を行う．

本実験における精度の比較・検証には，天気予報で用いられている適中率，見逃し率，空振り率，捕捉率，一致率を用いる（気象庁）．

適中率は，全事例数に対する正解数の割合を示す．また，見逃し率は，全事例数に対して，正解が西高東低冬型であるものを非西高東低冬型に分類した数（見逃し）の割合を示し，空振り率は，全事例数に対して，正解が非西高東低冬型であるものを西高東低冬型に分類した数（空振り）の割合を示す．さらに，捕捉率は，正解が西高東低冬型であるもののうち，正しく西高東低冬型に分類された数の割合を示し，一致率は，西高東低冬型に分類されたもののうち，正解が西高東低冬型であるものの割合を示す．

適中率，見逃し率，空振り率，捕捉率，一致率の計算式は，表 3 中の記号を用いて次の式で示される（気象庁）．

$$\text{適中率 (\%)} = \frac{A+D}{N} \times 100 \quad (2)$$

$$\text{見逃し率 (\%)} = \frac{B}{N} \times 100 \quad (3)$$

$$\text{空振り率 (\%)} = \frac{C}{N} \times 100 \quad (4)$$

$$\text{捕捉率 (\%)} = \frac{A}{N1} \times 100 \quad (5)$$

$$\text{一致率 (\%)} = \frac{A}{M1} \times 100 \quad (6)$$

ここで、適中率、捕捉率、一致率はそれぞれ 100 に近いほど精度がよく、一方で見逃し率と空振り率はそれぞれ 0 に近く、低いほど精度がよいことを示す。

3. 2 種類の学習データを用いた分類実験（実験 1）

III の 2 で述べた 2 種類の学習データを用いて識別関数を生成し、以下の分類実験を行う。

（実験 1-1）：学習データ 1 を用いて分類し、検証データの分類結果を表 3 で評価する。この実験の検証データ 3,653 日の中で正例となる「西高東低冬型」は 760 日である。

（実験 1-2）：学習データ 2 を用いて分類し、検証データの分類結果を表 3 で評価する。この実験の検証データ 3,653 日の中で正例となる「西高東低冬型」は 375 日である。

なお、実験 1-1、1-2 では非線形カーネル（RBF カーネル）のパラメータ δ をそれぞれ 1.0×10^{-5} 、 1.0×10^{-6} として実験を行った。

図 5 に実験 1-1 の分類の結果を示す。線形カーネルと比べて、非線形カーネルを用いた場合の方が、わずかではあるが、すべての評価尺度においてよい結果が得られた。また、適中率は、90% 以上、見逃し率は 5% 以下、空振り率は 3% 以下、捕捉率は 80% 弱、一致率は 85% 前後となった。

実験 1-1 の分類結果で見逃し（表 3 中の B）、空振り（表 3 中の C）であった事例を検討した結果、見逃しの多くは、西高東低冬型と他の型から成る移行型・複合型（ β ）の事例であり、空振りは、ほとんどが気圧の谷型と移動性高気圧型であった。これより、移行型・複合型（ β ）の事例は他の型が持つ特徴により、冬型に分類されず見逃されたと考えられる。

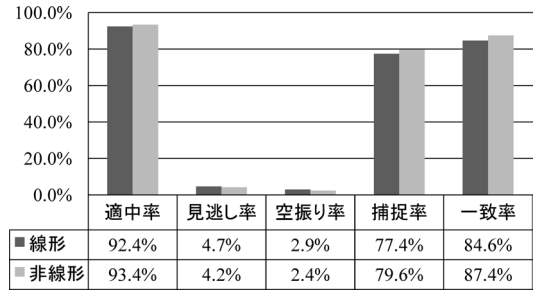


図 5 実験 1-1 の結果
Fig. 5 Results of experiment 1-1

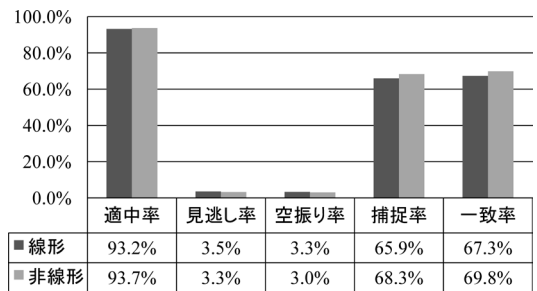


図 6 実験 1-2 の結果
Fig. 6 Results of experiment 1-2

図 6 に実験 1-2 の分類の結果を示す。実験 1-1 と同様に、すべての評価尺度において、非線形カーネルの方が線形カーネルよりよい結果となった。また、適中率は、実験 1-1 をわずかに上回った。これに対して、見逃し率と空振り率は共に 4% 以下であった。さらに、捕捉率と一致率は共に 70% 以下で実験 1-1 と比べて大幅に低い結果となった。

実験 1-2 の分類結果で空振り（表 3 中の C）であった事例を検討した結果、ほとんどの事例が西高東低冬型と他の型から成る移行型・複合型（ β ）であった。

なお、実験 1-1 では、捕捉率と一致率がそれぞれ 80% 弱と 85% 前後であったのに対して、実験 1-2 では、捕捉率と一致率が共に 70% 以下であった。ここで、実験 1-2 は実験 1-1 より検証データ中の正例（西高東低冬型）の数が少ないにもかかわらず、見

逃し率と空振り率の差がほとんどない。これより、実験1-1より表3中のAが少ないため、見逃し（表3中のB）と空振り（表3中のC）の数の影響を大きく受けたと考えられる。

以上より、本研究で正例と負例の設定に吉野（2002）による分類を用いたように、学習データに一定の基準に基づいた正例と負例を与えることによって、ある程度の精度で計算機による自動分類が可能であると考えられる。

4. 次元削減を行ったデータを用いた分類の評価 実験（実験2）

IIIの3の次元削減後の学習データ・検証データを用いて分類実験を行う。次元削減のパラメータ n を2~20（格子点数を $37 \times 57 \sim 1 \times 2$ ）に変化させ、分類結果を比較する。ここで、非線形カーネル、すなわちRBFカーネルのパラメータ δ は、 $n=2, 3$ のとき 1.0×10^{-5} 、 $n=4 \sim 8, 19, 20$ のとき 1.0×10^{-4} 、 $n=9 \sim 15, 18$ のとき 1.0×10^{-3} 、 $n=16, 17$ のとき 1.0×10^{-2} として実験を行った。

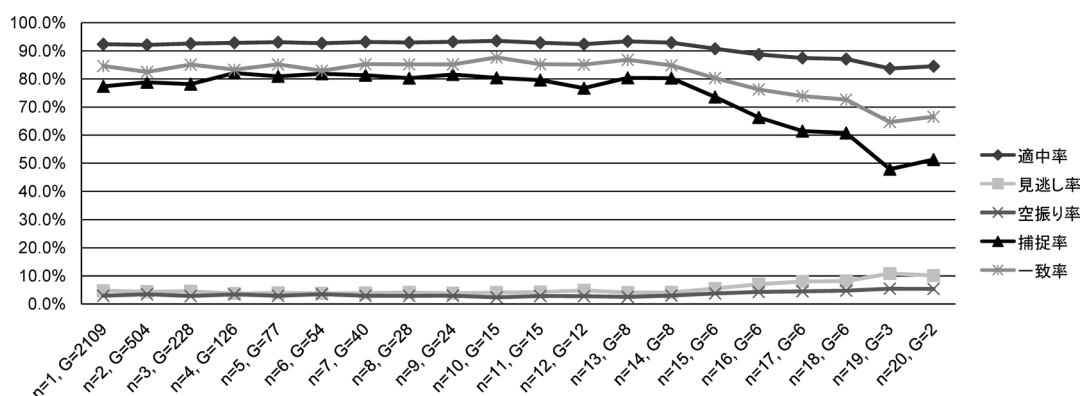


図7 実験2の結果（線形カーネル）

n ：次元削減のパラメータ G ：新たに生成したデータの格子点数（次元数）

Fig. 7 Results of experiment 2 (linear kernel)

n : parameter of dimension reduction G : number of grid points (dimensions) of data generated

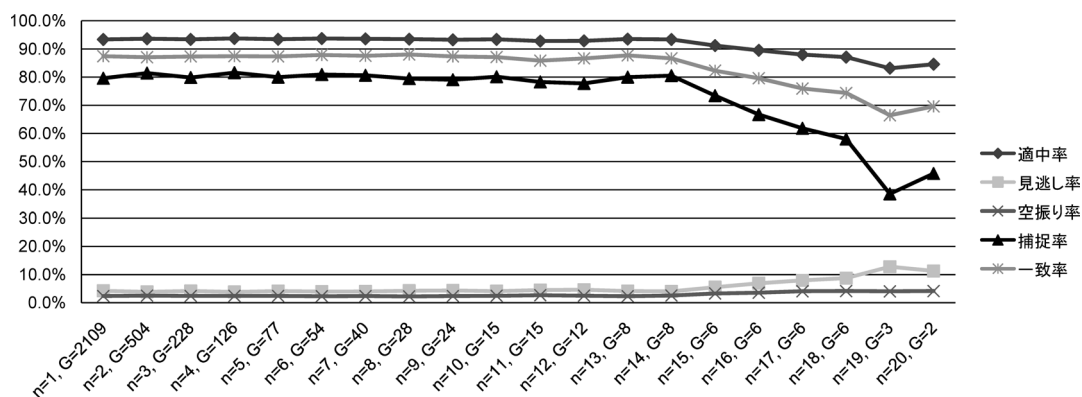


図8 実験2の結果（非線形カーネル）

n ：次元削減のパラメータ G ：新たに生成したデータの格子点数（次元数）

Fig. 8 Results of experiment 2 (nonlinear kernel)

n : parameter of dimension reduction G : number of grid points (dimensions) of data generated

なお、実験1-1と同様に、学習データには表2の学習データ1を用い、検証データの分類結果は表3を用いて評価する。

図7は線形カーネル、図8は非線形カーネルを用いた場合の実験結果である。図中の n は、次元削減のパラメータを、 G は新たに生成した格子点データの格子点数（次元数）を示す。また、比較対象として次元削減前のデータ（ $n=1$, $G=2, 109$ ）を示してある。次元削減前の結果と比べて、わずかではあるがよい精度を示したデータがあった。また、線形、非線形のどちらにおいても、6次元（ $n=15$ ）以降は精度が悪くなったが、8次元（ $n=14$ ）までは全体的にはほぼ近い値が得られ、精度を維持することができた。

以上のように、次元削減によって格子点数を減らしたデータを用いた場合、8次元まで削減した場合でも分類精度を維持することが可能であった。これより、気圧分布が広範囲にわたる西高東低冬型については、次元を削減した場合でも気圧配置の特徴が抽出されると考えられる。

V 結 論

本稿では、日々の気圧配置から西高東低冬型を検出することを目的として、SVMを用いた手法を提案した。この手法を用いて、日々の気圧配置を西高東低冬型と非西高東低冬型に分類し、手法の有用性を検討した。SVMの正例（西高東低冬型）と負例（非西高東低冬型）の与え方には、吉野（2002）の気圧配置分類を利用し、気圧データにはJRA-25を用いた。実験では正例と負例の与え方の違う2種類の学習データを用いて識別関数を生成し、それぞれにおいて分類結果の比較・検証を行った。さらに、気圧データの次元削減を行い、分類結果の比較・検証を行った。

本手法で実験を行った結果、西高東低冬型の気圧配置を、90%以上の適中率で検出できることがわ

かった。また、正例と負例を一定の基準に基づいて設定することで、計算機による自動分類が可能であることが確認された。さらに、次元を削減しても、ある程度の次元数までは精度を維持したままの分類が可能であった。

本手法の自動分類により、ある基準での気圧配置の分類が可能となる。この手法を用いた気圧配置の分類は、分類の際の負担を軽減でき、気候学研究に有用であると考えられる。また、本手法で長期間にわたるデータを分類することにより、気圧配置の出現頻度の変動について、より客観的な評価が可能になると期待される。

今後の課題は、分類精度を向上させることと、他の気圧配置に対する適用可能性の評価である。

本研究の一部は科学研究費補助金萌芽研究（#18650018）による。

なお本稿は、電子情報通信学会第19回データ工学ワークショップ（DEWS2008）において発表した内容をもとに加筆・改訂したものである。

（投稿 2008年6月30日）

（受理 2009年1月10日）

注

- 1) 15種類の分類とは、西高東低冬型、気圧の谷型a~d、移動性高気圧型a~d、前線型a~b、南高北低夏型、台風型a~cである。

文 献

- 石井健一郎・上田修功・前田英作・村瀬 洋 1998.『わかりやすいパターン認識』オーム社。
- 岡部孝弘・佐藤洋一 2003. 証明変化をとまなう物体認識へのサポートベクターマシンの適用. 情報処理学会論文誌 44, SIG 5(CVIM 6): 74-84.
- 甲斐啓子・吉野正敏 1975. 気圧配置型の組合わせ. 地理学評論 48: 497-503.
- 気象庁 検証方法の説明. <http://www.data.jma.go.jp/fcd/yoho/kensho/explanation.html>（最終閲覧日：2008年12月30日）
- 津田宏治 2000. サポートベクターマシンとは何か. 電子情報通信学会誌 83: 460-466.
- 宮 由可子・日下博幸 2009. 鉛直構造に着目した空っ風

- の気候学的研究. 地理学評論 82: 346–355.
- 山川修治 1988. 東アジアにおける卓越気圧配置型の季節推移からみた近年の気候変動. 地理学評論 61A: 381–403.
- 山川修治 1989. 20 世紀の寒冷期・温暖期・変動期における東アジア地上気圧配置の季節内変動と特異性. 地学雑誌 98: 833–852.
- 吉野正敏監修 2002. 『日本の気候 I 最新データでメカニズムを考える』二宮書店.
- 吉野正敏・甲斐啓子 1975. 気圧配置ごよみの気候学, 1941–1970. 天気 22: 204–205.
- Chang, C.-C. and Lin, C.-J. 2000. LIBSVM-A Library for Support Vector Machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/> (最終閲覧日: 2008 年 12 月 30 日)
- Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N. and Taira, R. 2007. The JRA-25 reanalysis. *J. Meteor. Soc. Japan* 85: 369–432.

Geographical Review of Japan Series A 82–4 323–331 2009

Detection of Pressure Patterns Using Support Vector Machine: Winter Type Pressure Pattern

KIMURA Hiroki*, KAWASHIMA Hideyuki**, KUSAKA Hiroyuki*** and KITAGAWA Hiroyuki**

(*Graduate student, University of Tsukuba, **Graduate School of Systems and Information Engineering/Center for Computational Sciences, University of Tsukuba, ***Graduate School of Life and Environmental Sciences/Center for Computational Sciences, University of Tsukuba)

In climate research, when climatologists need to know on which days a specific pressure pattern occurred, for example, “low in the west and high in the east (winter type)” or “high in the south and low in the north (summer type),” a huge number of surface weather charts must be scanned visually. To solve this problem, we propose an automatic detection method using pattern recognition developed in the computer science field. In this study, we used support vector machine (SVM), which is one pattern recognition method. Then, we classified pressure patterns into the “winter type” and “not winter type.” We confirmed the validity of the proposed method experimentally. In our experiments, we used the JRA-25 data from 1981 to 2000 for training data and test data.

The results showed that our SVM method achieves greater than 90% accuracy. Therefore, automatic detection of pressure pattern is possible by setting positive and negative examples based on a standard. Additionally, long-term data can be classified easily using our proposed method, and the results can be used to evaluate changes in the frequency of pressure patterns.

Key words: pressure pattern, winter type, support vector machine, surface weather charts