

パーセプトロン収束定理について

東京理科大学工学部情報科学科 児玉賢史 (Satoshi Kodama)
東京理科大学大学院工学研究科情報科学専攻 水谷友哉 (Tomoya Mizutani)
東京理科大学工学部情報科学科 明石重男 (Shigeo Akashi)
Department of Information Sciences, Faculty of Science and Technology,
Tokyo University of Science

1 パーセプトロン収束定理とハーン・バナッハの定理の関係

ヒルベルト空間に含まれる、共通部分を持たない 2 個の凸集合を分離する超平面が存在することは、ハーン・バナッハの定理によって保証されている。先ほど述べた分離ベクトルの存在にハーン・バナッハの定理を用いた場合、具体的に w の構成法を示すことは不可能である。これは、ハーン・バナッハの定理の証明の中で、ツォルンの補題が用いられているためである。この証明法に関する考え方は、ヒルベルトとブラウアーにまでさかのぼることができる。「存在することさえ保証されていれば、それをみつける方法は問わない」とするヒルベルトの提唱に対して、「存在の保証だけでなく、それをみつけるための具体的手順が明示されていなくてはならない」とするブラウアーの提唱は、その後の数学に大きな影響を与えた。当時の数学は、証明という考え方に対して、ヒルベルトの提唱を採用し、現代の数学はその流れを汲んでいる。当時の数学者たちは、「もしブラウアーの提唱を採用したならば、数学のかなり多くの重要な結果が得られなくなる」ことに気づき、数学を再構成する必要があること、そして再構成に伴う難問が新たに多数生じてくることを予測したと言われている。しかし、現在、数学に対する自然科学の他の分野からの需要が高まっている今日では、このヒルベルトの提唱が受け入れられているとは言い難い。なぜなら、数学を含む理学が「why」を追及するのに対して、工学は「how」を追及するものであり、そのためには、研究対象となるものの構造が明確に示されていることが必要とされるからである。

パーセプトロン収束定理は、有限次元ユークリッド空間の共通部分を持たない相異なる 2 個の凸多面体を分離するための超平面を具体的に構成する方法を与える。しかし、無限次元ベクトル空間では、この定理の証明法は成立しない。これは、パーセプトロン収束定理に基づいて分離可能性が保証される 2 個の凸集合に対して、有限個のベクトルから構成される凸包に限られることが条件づけられているからである。一方ハーン・バナッハの定理は、無限次元空間でも成立するだけでなく、共通部分を持たない 2 個の凸集合に対しても適用可能である。しかし、2 個の集合を分離する超平面を具体的に構成することはできない。そこで本原稿では、分離対象となる共通部分を持たない 2 個の凸集合が、それぞれ有限個のベクトルの凸包として構成されているという条件を満たす場合に適用可能なアルゴリズム論的証明法を紹介し、非線形解析学的考察を加える。

2 パーセプトロン収束定理の非線形解析学的側面

\mathcal{H} をヒルベルト空間とする。 m および n を自然数とし、 $\{x_i; 1 \leq i \leq m\}$ および $\{y_j; 1 \leq j \leq n\}$ を、それぞれ有限個のベクトルから構成される集合とし、それぞれ X および Y と書くことにする。この時、あるベクトル w およびある実数 r が存在して、

$$\langle w, x_i \rangle < r, \quad 1 \leq i \leq m$$

および

$$\langle w, y_j \rangle > r, \quad 1 \leq j \leq n$$

という条件を満たすとき、ベクトル w は、これら2個のベクトルの集合 X および Y を分離すると言われる。以下では、このような分離ベクトルの存在を仮定する。一般性を失うことなく、 $r = 0$ と仮定することができる。さらに、 X および Y を構成する全ての要素のノルムが1であることを仮定しても一般性を失うことはない。ここで、分離ベクトル w の仮定により、 X および Y を用いて定義される以下の定数：

$$\rho = \min \left\{ \min_{1 \leq i \leq m} | \langle w, x_i \rangle |, \min_{1 \leq j \leq n} | \langle w, y_j \rangle | \right\}$$

は、正の値をとる。このとき、次の定理が成り立つ。

命題 1 ヒルベルト空間 \mathcal{H} の要素 v が、 $\|v - \frac{w}{\rho}\| < 1$ を満たすとき、 v も分離ベクトルとなる。

証明 1 任意の X の要素 x_i に対して、

$$\begin{aligned} \langle p, x_i \rangle &= \left\langle p - \frac{w}{\rho}, x_i \right\rangle + \left\langle \frac{w}{\rho}, x_i \right\rangle \\ &< \left\| p - \frac{w}{\rho} \right\| \cdot \|x_i\| + \left\langle \frac{w}{\rho}, x_i \right\rangle \\ &= 1 + \frac{\langle w, x_i \rangle}{\rho} \end{aligned}$$

が成り立つ。ここで、 ρ の定義式より、右辺第2項は-1以下であることが分かる。従って、 p は、 x_i に関して正しい判定を行っている。 $\langle p, y_j \rangle > 0$ の成立も同様にして証明できる為、 p が分離ベクトルであることが示せた。

アルゴリズム 1

手順1：最初に x_1 から x_m を並べ、次に y_1 から y_n を並べて構成される周期的ベクトル列を $\{z_k\}$ とする。続いて、 w を近似するベクトル列 $\{w_k\}$ を次のように構成する。

手順2：以下の手順を繰り返す。

(2-0) 初期ベクトル w_0 を0ベクトルとする。カウンター変数 k を0と設定して、以下に進む。

(2-1) もし z_k が X の要素であるとき、 $\langle w_k, z_k \rangle < 0$ が成立するか否かを調べる。また、もし z_k が Y の要素であるとき、 $\langle w_k, z_k \rangle > 0$ が成立するか否かを調べる、成立している場合、 $w_{k+1} = w_k$ と定義し、カウンター変数 k を1増加した後、(2-2)に進む。もし成立

していない場合は、 z_k が X の要素であるとき、 $w_{k+1} = w_k + z_k$ と定義する。また、 z_k が Y の要素であるとき、 $w_{k+1} = w_k - z_k$ と定義する。その後、カウンター変数 k を 1 増加した後、(2-2) に進む。

(2-2) k が $m+n$ 以下であれば、(2-1) に戻る。 k が $m+n+1$ 以上の場合、 $w_{k-(m+n)} = \dots = w_{k-1}$ が成立していれば、(2-3) に進む。成立していなければ、(2-1) に戻る。

(2-3) w_{k-1} は、 X と Y の分離ベクトルとなっている。

命題 2 上記アルゴリズムは、有限回の操作で収束する。

証明 2 自然数 k に対して、 z_k が X の要素である場合、 $\langle w_k, z_k \rangle < 0$ が成立すれば、正しく判定されたことになるため、 $\|w_{k+1}\| = \|w_k\|$ が成立する。もし正しく判定されなかった場合は、 $\langle w_k, x_k \rangle > 0$ が仮定されることになり、そのため、以下の不等式が成立することになる。

$$\begin{aligned} \|w_{k+1} - \frac{w}{\rho}\|^2 &= \|(w_k - x_k) - \frac{w}{\rho}\|^2 \\ &= \|(w_k - \frac{w}{\rho}) - x_k\|^2 \\ &= \|w_k - \frac{w}{\rho}\|^2 - 2 \langle w_k - \frac{w}{\rho}, x_k \rangle + \|x_k\|^2 \\ &= \|w_k - \frac{w}{\rho}\|^2 - 2 \langle w_k, x_k \rangle + \frac{2 \langle w, x_k \rangle}{\rho} + 1 \\ &< \|w_k - \frac{w}{\rho}\|^2 - 1 \end{aligned}$$

これは、数列 $\{\|w_k - \frac{w}{\rho}\|\}$ が単調減少列であることを示しており、補正が行われるたびに 1 ずつ減少していく。今、ある自然数 k_0 に対して、 $\|w_{k_0} - \frac{w}{\rho}\| < 1$ が成立した場合、命題 1 により w_{k_0} が分離ベクトルとなっていることが保証されたベクトルに到達している。従って、このアルゴリズムは有限実行時間で必ず停止する。

3 実行例と実行時間評価

X を $\{(1/\sqrt{2}, 1/\sqrt{2}), (-1/\sqrt{2}, 1/\sqrt{2})\}$ から構成される 2 点集合、 Y を $\{(1/\sqrt{2}, -1/\sqrt{2}), (-1/\sqrt{2}, -1/\sqrt{2})\}$ から構成される 2 点集合とする。このとき、前節のアルゴリズムを適用した場合、求める分離ベクトルは、以下のようにして決定される。

- (1). 分離ベクトルの初期状態 w_0 を $(0, 0)$ とする。
- (2). $\langle (0, 0), (1/\sqrt{2}, 1/\sqrt{2}) \rangle \geq 0$ となるため、 $w_1 = (0, 0) + (1/\sqrt{2}, 1/\sqrt{2}) = (1/\sqrt{2}, 1/\sqrt{2})$ と定める。
- (3). $\langle (1/\sqrt{2}, 1/\sqrt{2}), (-1/\sqrt{2}, 1/\sqrt{2}) \rangle \geq 0$ となるため、 $w_2 = (1/\sqrt{2}, 1/\sqrt{2}) + (-1/\sqrt{2}, 1/\sqrt{2}) = (0, \sqrt{2})$ と定める。
- (4). $\langle (0, \sqrt{2}), (1/\sqrt{2}, -1/\sqrt{2}) \rangle < 0$ となるため、 $w_3 = w_2$ と定める。
- (5). $\langle (0, \sqrt{2}), (-1/\sqrt{2}, -1/\sqrt{2}) \rangle < 0$ となるため、 $w_4 = w_3$ と定める。

- (6). $\langle (0, \sqrt{2}), (1/\sqrt{2}, 1/\sqrt{2}) \rangle > 0$ となるため、 $w_5 = w_4$ と定める。
 (7). $\langle (0, \sqrt{2}), (-1/\sqrt{2}, 1/\sqrt{2}) \rangle > 0$ となるため、 $w_6 = w_5$ と定める。
 (8). X および Y の合計要素数が 4 個であり、 $w_2 = w_3 = w_4 = w_5 = w_6$ の成立が確認されたため、この時点でアルゴリズムが終了する。

実行時間評価に関しては、「分離ベクトルを見出すのに必要とされる実行時間の上限」を求められることが理想である。今回のアルゴリズムでは、「分離ベクトルの候補が分からなければ実行時間を評価することができない」という問題点を含んでいる。 $\|w/\rho\|^2$ の値によってその実行時間を評価することが可能であるが、残念なことに、 w はもともと存在の仮定された分離ベクトルであるため、このベクトルの存在は、保障されているが具体的構造はわからないという状況である。しかし、 w/ρ は、 $\|w\|$ の値に関係なく w と X および Y に含まれる各要素との相対的位置関係によってのみ決定される。従って、 w が、 $x_1, \dots, x_m, y_1, \dots, y_n$ のいずれかであると仮定した場合、 $m + n + \{1/\cos^2((\cos^{-1} \rho)/2)\}$ は、上限回数を下から評価する値と見なすことが可能である。この値を上記例に適用するならば、 $2 + 2 + \{\cos^2(\cos^{-1}(0))\} = 6$ ステップとして求められる。実際に、求める分離ベクトルを得られるまでに 7 ステップを必要としているので、この評価方法は、上実行時間上限に関するある程度の目安を与えているとみなすことができる。

参考文献

- [1]. 高橋渉著, 凸解析と不動点近似, 数理解析シリーズ, 横浜図書, 2000 年.
 [2]. PHILIP D. WASSERMAN 著, 石井尚宏/塚田稔共訳, ニューラルコンピューティング
 -理論と実績-, 森北出版株式会社, 1993 年.