パーセプトロンの収束定理の証明

2010. 5. 11　石井

合計 n 個のパターン（特徴ベクトル）$\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ より成る学習パターンを用意する。各パターンはクラス ω_1, ω_2 のいずれかに属するものとし、かつこれらは線型分離可能とする。線型識別関数 $g(\mathbf{x})$ を

$$g(\mathbf{x}) = \mathbf{w}^t\mathbf{x} \tag{1}$$

とし、

$$\begin{cases} g(\mathbf{x}) = \mathbf{w}^t\mathbf{x} > 0 & \text{(for all } \mathbf{x} \in \omega_1) \\ g(\mathbf{x}) = \mathbf{w}^t\mathbf{x} < 0 & \text{(for all } \mathbf{x} \in \omega_2) \end{cases} \tag{2}$$

となるよう、重みベクトル $\mathbf{w}$ を決定したい。ここで $\mathbf{x} \in \omega_2$ の全パターンを $\mathbf{x} \to -\mathbf{x}$ と変更すると、式 (2) は

$$g(\mathbf{x}) = \mathbf{w}^t\mathbf{x} > 0 \qquad \text{(for all } \mathbf{x}) \tag{3}$$

となる。パーセプトロンの誤り訂正の過程で誤識別された学習パターンを順に

$$\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^k, ... \tag{4}$$

で表す。ただし

$$\mathbf{x}^k \in \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n\} \qquad (k = 1, 2, ...) \tag{5}$$

である。パーセプトロンの学習規則は以下に示す処理になる（定数 $\rho = 1$）。

重みベクトルの初期値 $\mathbf{w}^1$ を任意に設定し、k 番目の誤識別パターン $\mathbf{x}^k$ が発生したとき、重みベクトル $\mathbf{w}^k$ を次式により $\mathbf{w}^{k+1}$ に修正する。

$$\mathbf{w}^{k+1} = \mathbf{w}^k + \mathbf{x}^k \qquad (k \geq 1) \tag{6}$$

ここで、解となる重みベクトルの一つを $\hat{\mathbf{w}}$ とすると、

$$\hat{\mathbf{w}}^t\mathbf{x}_p > 0 \qquad (p = 1, 2, ..., n) \tag{7}$$

が成り立つ。ここで、定数 $\alpha\,(> 0)$ を用いると、式 (7) より

$$\alpha\hat{\mathbf{w}}^t\mathbf{x}_p > 0 \qquad (p = 1, 2, ..., n) \tag{8}$$

が成り立つから、重みベクトル $\alpha\hat{\mathbf{w}}$ も解である。そこで、

$$\|\mathbf{w}^{k+1} - \hat{\mathbf{w}}\| < \|\mathbf{w}^k - \hat{\mathbf{w}}\| \tag{9}$$

の代わりに

$$\|\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\| < \|\mathbf{w}^k - \alpha\hat{\mathbf{w}}\| \tag{10}$$

を示す。式 (6) の両辺より $\alpha\hat{\mathbf{w}}$ を引くと

$$\left(\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\right) = \left(\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right) + \mathbf{x}^{k} \qquad (k \geq 1) \tag{11}$$

両辺のノルムをとって自乗することにより、

$$\left\|\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\right\|^2 = \left\|\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right\|^2 + 2\left(\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right)^t\mathbf{x}^{k} + \left\|\mathbf{x}^{k}\right\|^2 \tag{12}$$

$$= \left\|\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right\|^2 + 2\left(\mathbf{w}^{k}\right)^t\mathbf{x}^{k} - 2\alpha\hat{\mathbf{w}}^t\mathbf{x}^{k} + \left\|\mathbf{x}^{k}\right\|^2 \tag{13}$$

が得られる。式 (4) で示したように、$\mathbf{x}^k$ は誤識別されたパターンであるので、式 (3) より

$$g(\mathbf{x}^k) = \left(\mathbf{w}^{k}\right)^t\mathbf{x}^{k} \leq 0 \tag{14}$$

が成り立つ。したがって式 (12)，(13) はさらに変形され

$$\left\|\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\right\|^2 \leq \left\|\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right\|^2 - 2\alpha\hat{\mathbf{w}}^t\mathbf{x}^{k} + \left\|\mathbf{x}^{k}\right\|^2 \tag{15}$$

が得られる。

$$\beta \stackrel{\text{def}}{=} \max_{p=1,\ldots,n} \left\|\mathbf{x}_p\right\| \tag{16}$$

$$\gamma \stackrel{\text{def}}{=} \min_{p=1,\ldots,n} \hat{\mathbf{w}}^t\mathbf{x}_p > 0 \tag{17}$$

とすると、式 (15) より、

$$\left\|\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\right\|^2 \leq \left\|\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right\|^2 - 2\alpha\gamma + \beta^2 \tag{18}$$

を得る。ここで α を

$$\alpha = \frac{\beta^2}{\gamma} \tag{19}$$

と設定すると、

$$\left\|\mathbf{w}^{k+1} - \alpha\hat{\mathbf{w}}\right\|^2 \leq \left\|\mathbf{w}^{k} - \alpha\hat{\mathbf{w}}\right\|^2 \quad - \beta^2 \tag{20}$$

$$\leq \left\|\mathbf{w}^{k-1} - \alpha\hat{\mathbf{w}}\right\|^2 - 2\beta^2 \tag{21}$$

$$\cdots$$

$$\leq \left\|\mathbf{w}^{1} - \alpha\hat{\mathbf{w}}\right\|^2 \quad - k\beta^2 \tag{22}$$

が成り立つ。式 (20) の左辺は負になり得ないので、

$$0 \leq \left\|\mathbf{w}^{1} - \alpha\hat{\mathbf{w}}\right\|^2 - k\beta^2 \tag{23}$$

これより

$$k \leq \frac{\left\|\mathbf{w}^{1} - \alpha\hat{\mathbf{w}}\right\|^2}{\beta^2} \tag{24}$$

（証明終）