人工知能―AI の基礎から知的探索へ：演習問題解答例

第 6 章　機械学習の基礎

$$y = g(u) = \begin{cases} 1 & \mathit{if}\ u \geq 0 \\ 0 & \mathit{otherwise} \end{cases} \tag{5.10a}$$

$$u = \sum_{i=1}^{n} w_i x_i - \theta \tag{5.10b}$$

$$g^{+}(\mathbf{x}) = \sum_{j=1}^{n} x_j r_j^{+} - \frac{1}{2}\sum_{j=1}^{n} (r_j^{+})^2, \quad g^{-}(\mathbf{x}) = \sum_{j=1}^{n} x_j r_j^{-} - \frac{1}{2}\sum_{j=1}^{n} (r_j^{-})^2 \tag{6.9}$$

$$g(\mathbf{x}) = g^{+}(\mathbf{x}) - g^{-}(\mathbf{x}) = \sum_{j=1}^{n} w_j x_j - \theta \tag{6.10}$$

演習問題 6.1　式（6.10）の識別関数 $g(\mathbf{x})$のパラメータ w_j と θ を、平均パターンの要素を使って表せ（ヒント：式(6.9)を参照すること）。また、式(6.10)と式(5.10b)を比較し、線形識別関数と単一ニューロンの共通点について議論せよ。

解答　式(6.9)より、$g(\mathbf{x})$のパラメータ w_j と θは、以下のように求められる：

$$w_j = r_j^{+} - r_j^{-}$$

$$\theta = -\frac{1}{2}\sum_{j=1}^{n}[(r_j^{+})^2 - (r_j^{-})^2]$$

また、式(6.10)と式(5.10b)を比較するとわかるように、線形識別関数と単一ニューロンは、同じ式に基づいてパターン分類を行っている。

演習問題 6.2　全体集合 X は顔画像の集合、A={笑顔}は X に定義された概念であるとして、この概念を認識する NNC について議論せよ（ヒント：例題 6.1 を参照すること）。

解答　笑顔を識別する NNC を設計するために、まず、笑顔と笑顔でない画像を、例えば、それぞれ数千枚集め、それらにラベル＋１と－１を付けて、訓練集合Ωに入れる。このΩを NNC のプロトタイプ集合として、任意の顔画像 **x** に対して、それを式(6.2)と(6.3)にしたがって認識することができる。なお、ここの議論においては、顔画像自体はすでに別の方法で抽出されたと仮定している。もちろん、顔画像を抽出しながら笑顔を検出することも可能である。興味がある方はチャレンジしてみると良い。

$w_3 \cdot 1^3 + w_2 \cdot 1^2 + w_1 \cdot 1 + w_0 = 1.5$

$w_3 \cdot 2^3 + w_2 \cdot 2^2 + w_1 \cdot 2 + w_0 = 3$

$w_3 \cdot 3^3 + w_2 \cdot 3^2 + w_1 \cdot 3 + w_0 = 5.5$

$w_3 \cdot 4^3 + w_2 \cdot 4^2 + w_1 \cdot 4 + w_0 = 9$

演習問題 6.3　例題 6.4 の連立方程式を解いて、$h(x)$を求めよ。

解答　連立方程式は、以下のように書き直すことができる：

$$
\begin{aligned}
w_3 + w_2 + w_1 + w_0 &= 1.5 \\
8w_3 + 4w_2 + 2w_1 + w_0 &= 3 \\
27w_3 + 9w_2 + 3w_1 + w_0 &= 5.5 \\
64w_3 + 16w_2 + 4w_1 + w_0 &= 9
\end{aligned}
$$

ガウスの消去法を利用する。連立方程式を以下のように変形していく：

$$
\begin{aligned}
w_3 + w_2 + w_1 + w_0 &= 1.5 \\
4w_2 + 6w_1 + 7w_0 &= 9 \\
18w_2 + 24w_1 + 26w_0 &= 35 \\
48w_2 + 60w_1 + 63w_0 &= 87
\end{aligned}
$$

$$
\begin{aligned}
w_3 + w_2 + w_1 + w_0 &= 1.5 \\
w_2 + 1.5w_1 + 1.75w_0 &= 2.25 \\
3w_1 + 5.5w_0 &= 5.5 \\
12w_1 + 21w_0 &= 21
\end{aligned}
$$

$$
\begin{aligned}
w_3 + w_2 + w_1 + w_0 &= 1.5 \\
w_2 + 1.5w_1 + 1.75w_0 &= 2.25 \\
w_1 + 5.5/3w_0 &= 5.5/3 \\
w_0 &= 1
\end{aligned}
$$

$\therefore$　$w_0=1$;　$w_1=0$;　$w_2=0.5$;　$w_3=0$

$\therefore$　$h(x) = 1 + 0.5x^2$

演習問題 6.4　例題 6.5 において、$s(t)$をある地域の消費電力であるとする。その地域の発電所は、毎日どれくらい発電すればよいか、そのためにどれくらい燃料を用意すればよいかをわかっていれば、生産効率の向上に繋がる。直近 14 日（2 週間）の消費量をもとに、翌日の発電量を予測する問題を、機械学習で解決する方法について検討せよ。

解答　例題 6.5 と同じように考える。まず、訓練集合を作る。問題は、$s(t)$の現在値 $s(i)$と、直近に観測された 14 個の値 $s(i\text{-}1), \ldots, s(i\text{-}14)$をもとに、次の値 $s(i+1)$を予測することである。データは、14+1=15 次元のベクトル $\mathbf{x}_i=[s(i), s(i\text{-}1), \ldots, s(i\text{-}14)]^T$で、それに対応する期待出力は $y_i=f(\mathbf{x}_i)=s(i+1)$である。例えば、時刻 $i=1, 2, \ldots, N$ に対して、$<\mathbf{x}_i, y_i>$を記憶すれば、N 個の観測データを含む訓練集合Ωを作ることができる。ここで、一つのデータは一日に対応する。

次に、仮説集合を考える。この問題は、予測問題であるので、以下の AR モデルが使える：

$$s(m+1)=\sum_{k=0}^{n} w_k s(m-k)+c \tag{6.17}$$

ここで、w_kと c はパラメータである。式(6.15)の誤差を最小となるように、AR モデルのパラメータを求めればよい。そのために、Ωにあるデータをそれぞれ式(6.17)に代入し、N 個の方程式が得られる。データ数 N が n+1 であれば、連立方程式を解くことによって全てのパラメータを求めることができる。通常、N は n+1 よりもはるかに大きいので、パラメータが最小二乗法(least squares method)を利用して求められる。

以上の線形予測モデルが間に合わない場合には、ニューラルネットワーク、サポートベクトルマシンなどの非線形モデルを使い、それに対応する学習アルゴリズムを利用することができる。

表 6.2　例題 6.6 の訓練データ

クラス 0	クラス 1
-0.96, -0.27 -0.86, 0.51 0.86, -0.50 0.65,-0.76	-0.400000, 0.92 0.800000, 0.60 0.060000, 1.00 0.930000, 0.36

表 6.3　例題 6.6 の結果

初期プロトタイプ	学習で得られたプロトタイプ
0.882983 0.469404 0; -0.619131 -0.785288 0; 0.901795 -0.432164 1; -0.151238 -0.988497 1;	0.842607 -0.538529 0; -0.983217 -0.182438 0; 0.665301 0.746575 1; -0.151238 -0.988497 1;

$$\mathbf{r} = \mathbf{r} + \alpha(\mathbf{x} - \mathbf{r}) \tag{6.18}$$

$$\mathbf{r} = \mathbf{r} - \alpha(\mathbf{x} - \mathbf{r}) \tag{6.19}$$

演習問題 6.5　表 6.2 のデータが与えられ、プロトタイプが表 6.3 の一列目のように初期化されたとする。クラス 0 の一番目のデータを使い、その最近傍となるプロトタイプ（すなわち、勝者）を求め、それを式(6.18)か(6.19)で更新せよ。更新した後のプロトタイプを、図に描いて、図 6.8 の左図と比較し、その変化について議論せよ。

解答　データもプロトタイプも正規化されているので、勝者を求めるためには、内積を求めればよい。データ x1=(-0.96, -0.2)に対して、各プロトタイプとの内積を求めると、以下のようになる：

```
y1=<x1,p1>=-0.9744,
y2=<x1,p2>=0.8064,
y3=<x1,p3>=-0.7490,
y4=<x1,p4>=0.4121
```

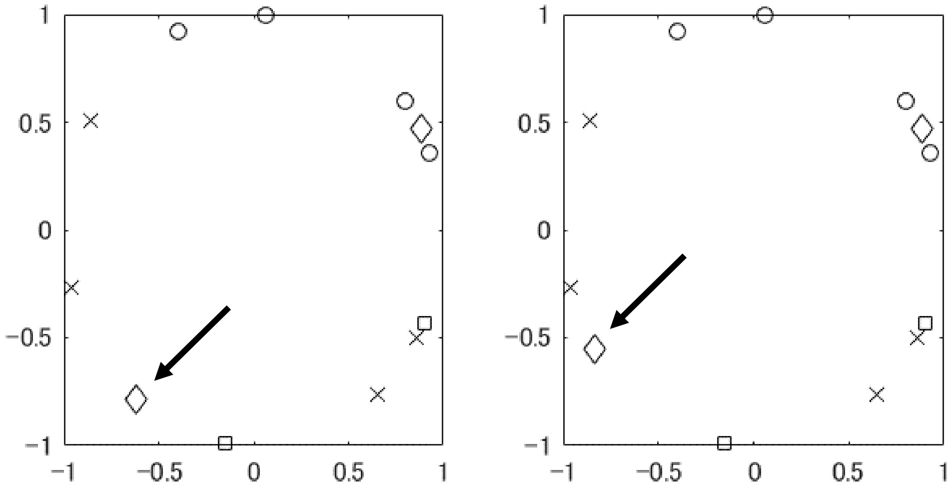
したがって、2 番のプロトタイプ p2 が勝者であり、p2 のクラスラベルが x1 と同じなので、p2 が式(6.18)を利用して更新される：

```
p2_new
=p2+0.5(x1-p2)
=(-0.619131,-0.785288)+0.5(-0.3409,0.5152)
=(-0.7896,-0.5276)
```

これを正規化すると、以下のようになる：

p2=p2_new/|p2_new|=(-0.8314,-0.5556)

これをプロットすると、下の図となる。図からわかるように、p2 が左側にある二つのデータにより近づいている。すなわち、よりクラスターセンターらしくなっている。



演習問題 6.6　例題 6.6 と同じパターンが与えられたとし、プロトタイプの初期値はそれぞれ (0,1)と (0,-1)であるとする。クラス０の３番目のパターンを利用して、プロトタイプを更新せよ。更新した後のプロトタイプを、図に描いて、それらがよりクラスターセンター (cluster centroid)らしくなっていることを確認せよ。

解答　データ x=(0.860000 -0.500000)と各プロトタイプとの類似度（内積）を求めると、以下のようになる：

y1=-0.5000
y2= 0.5000

したがって、２番目のプロトタイプ p2 が勝者である。それを更新し、正規化すると、以下のようになる：

p2=(0.4974 -0.8675)

結果をプロットすると、下の図となる。

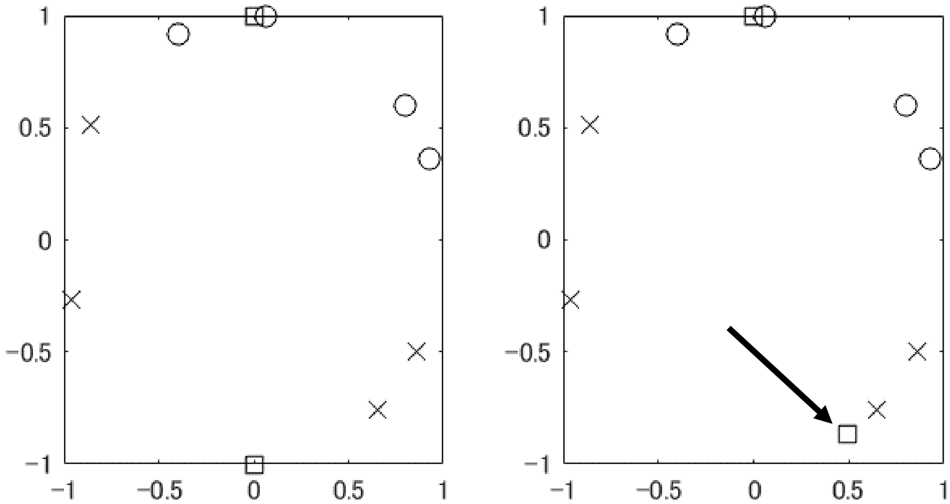


表 6.6　例題 6.9 の結果

クラス 0 のプロトタイプ	クラス 1 のプロトタイプ
(0.048296, 0.827414)	(0.847078, 0.356892)
(0.228646, 0.248717)	(0.551071, 0.427418)
(0.550853, 0.771827)	(0.551475, 0.827253)
(0.429797, 0.220361)	(0.844268, 0.778659)
(0.554753, 0.364483)	(0.430515, 0.183015)

演習問題 6.7　例題 6.8 の問題に対して、R^4 規則で得られたプロトタイプを図に描いて、対応する判別境界線を示せ。

解答　R^4 規則で求められたプロトタイプに対応するボロノイ図は、下の左図に示している。この図にある線のうち、同じクラスのプロトタイプの間の線を取り除けば、判別境界線が得られる。この境界線で領域$[0\ 1]^2$ にある点を分類すると、右図になる。図からわかるように、R4 規則で求めたプロトタイプは、理想の解と比べるとまだ大分差がある。しかし、訓練データを増やせば、理想の解をよりよく近似することができる。