Perceptronul R își propune să maximizeze separarea între clase (o probl. de discriminare). Perceptronul P își propune să maximizeze asemănarea cu etichete presupuse a fi „generate" de funcția $y(x) = wx + b$ (deci o problemă de interpolare). Inadecvarea la problema de învățare se plătește cu eroarea de misclasare $\neq 0$ (uneori).

$$\Delta w^{k+1} = -\rho \nabla J(w)\Big|_{w = w^k}$$

$$\frac{1}{2} J(w) = \frac{1}{p} \sum^{p} \frac{1}{2}\left(y - w^T x\right)^2 + \frac{\lambda}{2} w^T w$$

$$\nabla J(w) = (y - w^T x)(-x) + \lambda w$$

$$\Delta w^{k+1} = \rho\left(y - (w^k)^T x\right)(x) - \lambda \rho w^k$$

Seminar 1A

14/04/06

- Fie $S^+ = \{(x, +1) / x \sim U(0,3)\}$; $S^- = \{(x,-1) / x \sim U(4,5)\}$ și $S = S^+ \cup S^-$.

Fie perceptronul P cu spațiul intrărilor $X = (0,3) \cup (4,5)$ și funcție de transfer liniară.

a) Figurați pictograma lui P. Particularizați definiția perceptronului formal la P.

b) Dacă spațiul intrărilor posibile este $Z = [(0,3) \times \{-1\}] \cup [(4,5) \times \{1\}]$ ~~$X \times \{+1, -1\} \supseteq S$~~ iar funcția de pierdere $\ell(u, y) = (y - u)^2$ scrieți riscul real a lui P.

b.1) În spațiul W al parametrilor lui P scrieți ecuația suprafeței riscului real a lui P. Ce reprezintă această suprafață?

b.2) Găsiți ipoteza h^* ce minimizează riscul real al perceptronului P.

c) În spațiul W a lui P scrieți riscul empiric corespunzător riscului real de la b).

c.1) Găsiți $\bar{h}$ ce minimizează riscul empiric a lui P pe mulțimea de antrenare $S = \{(x_i, y_i) \in Z\}_{i=1}^{m}$

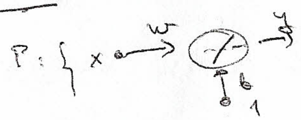
d) Fie funcționale
$$J(w, b) = 3 \int_0^3 (y(x) - 1)^2 dx + \int_4^5 (y(x) + 1)^2 dx$$
cu $y(x) = wx + b$.

Calculați $(w^*, b^*) = \arg\min_{(w,b)} J(w, b)$.

e) Calculați eroarea de misclasare a perceptronului de la punctele b), c) și d). Comparați rezultatele și justificați răspunsul.

Implementați la lab. + încercat cu $f(u) = \tanh(u)$.

Soluție

a)  cu $y(x) = wx + b$

$$P = (\underbrace{\{w, b\}}_{W}, \underbrace{\phi}_{\Im}, \underset{: \mathbb{R} \to \mathbb{R}}{G_W(x) = wx + b \equiv n}, \underset{: \mathbb{R} \to \mathbb{R}}{f_{\Im}(n) = n})$$

b) $$R_{real}(y) = \int_Z l(u, y(x))\, dF(x,u) =$$

$$= \underbrace{\tfrac{3}{4}}_{P(u=+1)} \int_0^3 (y(x)-1)^2 \underbrace{f_{U(0,3)}(x)}_{=1/3} dx + \underbrace{\tfrac{1}{4}}_{P(u=-1)} \int_4^5 (y(x)+1)^2 \underbrace{f_{U(4,5)}(x)}_{=1} dx$$

$\left(= \dfrac{\text{lg segment } (S_i)}{\text{lg segment } (S)} \right)$

b.1) $$J(w,b) = \frac{1}{4}\left[\int_0^3 (wx+b-1)^2 dx + \int_4^5 (wx+b+1)^2 dx\right]$$

$$= 1 - b + b^2 + \frac{9}{2} bw + \frac{22}{3} w^2 \quad \Big| \; 1 + b + b^2 + \frac{9}{2} bw + \frac{22 w^2}{3}$$

ce reprezintă un paraboloid $\left(\begin{matrix} \text{eliptic } \frac{x^2}{a^2} + \frac{y^2}{b^2} = z \\ \text{hiperbolic } \frac{x^2}{a^2} - \frac{y^2}{b^2} = z \end{matrix} \right.$

b.2) $$\begin{cases} \dfrac{\partial J}{\partial w} = 0 \\ \dfrac{\partial J}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \dfrac{9}{2} b + \dfrac{44}{3} w = 0 \\ -1 + 2b + \dfrac{9}{2} w = 0 \end{cases} \Rightarrow \begin{cases} w^* = -0.4854 13 \\ b^* = 1.61468 \end{cases}$$

$$\Rightarrow h^*(x) = w^* x + b^*$$

c) $$R_{emp}(y) = \frac{1}{n} \sum_{x_i \in (0,3)} (y(x_i)-1)^2 + \frac{1}{n} \sum_{x_i \in (4,5)} (y(x_i)+1)^2$$

cu $n_1 = \text{Card}\{x_j \in (0,3)\}$; $n_2 = \text{card}\{x_j \in (4,5)\}$.

c.1) $$\begin{cases} \dfrac{\partial R_{emp}}{\partial w} = 0 \\ \dfrac{\partial R_{emp}}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \dfrac{2}{n} \displaystyle\sum_{j=1}^{n_1} (wx_j + b - 1) x_j + \dfrac{2}{n} \displaystyle\sum_{j=1}^{n_2} (wx_j + b + 1) x_j = 0 \\ \dfrac{2}{n} \displaystyle\sum_{j=1}^{n_1} (wx_j + b - 1) + \dfrac{2}{n} \displaystyle\sum_{j=1}^{n_2} (wx_j + b + 1) = 0 \end{cases}$$

$$\Rightarrow \frac{n_1}{n} w \overline{x^2}_{(1)} + \frac{n_1}{n} b \overline{x}_{(1)} - \frac{n_1}{n} \overline{x}_{(1)} + \frac{n_2}{n} w \overline{x^2}_{(2)} + \frac{n_2}{n} b \overline{x}_{(2)} + \frac{n_2}{n} \overline{x}_{(2)} = 0$$

$$\frac{n_1}{n} w \overline{x}_{(1)} + b - 1 + \frac{n_2}{n} w \overline{x}_{(2)} + b + 1 = 0$$

$$\Rightarrow \begin{cases} (\overline{x^2}_{(1)} + \overline{x^2}_{(2)}) w + (\overline{x}_{(1)} + \overline{x}_{(2)}) b = \overline{x}_{(1)} - \overline{x}_{(2)} \\ (\overline{x}_{(1)} + \overline{x}_{(2)}) w + 2b = 0 \end{cases}$$

$$\overline{h}(x) = \overline{w} x + \overline{b}$$

d) $$E(w,b) = 10 + 10 b^2 + 4b(-4 + 9w) + \frac{2}{3} w (-27 + 71 w).$$

$$\begin{cases} \dfrac{\partial E}{\partial w} = 0 \\ \dfrac{\partial E}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \dfrac{284}{3} w + 36 b - 18 = 0 \\ 9w + 5b - 4 = 0 \end{cases} \Rightarrow \begin{matrix} w^{(E)} = -\dfrac{81}{224} \\ b^{(E)} = \dfrac{325}{224} \end{matrix}$$

$$h^{(E)}(x) = w^{(E)} x + b^{(E)}$$

e) În cazul b) $h^*(x) = 0 \Rightarrow x = 3.25526$

$\Rightarrow h^*((0,3)) > 0$ și $h^*((4,5)) < 0$

deci eroarea de misclasare este zero

în cazul c) dacă $\overline{h}((0,3)) > 0$ și $\overline{h}((4,5)) < 0$

atunci eroarea de misclasare este zero

în cazul d) $h^{(E)}(x) = 0 \Rightarrow x_0 = 4.01235$

$\Rightarrow x \in (4, x_0)$ sunt misclasate