

ICMC-USP
Lista de Exercícios 3
SCC-5809 - Redes Neurais
2o. Semestre de 2012 - Prof. João Luís



UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO
Departamento de Ciências de Computação

1. O perceptron é um classificador de padrões linear. Justifique esta afirmação. Verifique que as equações envolvidas no algoritmo de convergência do perceptron são consistentes com as desigualdades da prova do teorema de convergência.
2. O perceptron pode ser usado para executar numerosas funções lógicas. Demonstrar a implementação das funções lógicas binárias AND, OR e NOT. Entretanto, uma limitação básica do perceptron é que ele não pode implementar o OU-EXCLUSIVO. Explique a razão para esta limitação.
3. O algoritmo LMS ou Regra Delta é dito ser *convergente na média* se o valor médio do vetor peso $\mathbf{w}(n)$ aproxima-se da solução ótima \mathbf{w}_0 quando o número de iterações n aproxima-se do infinito:

$$E[\mathbf{w}(n)] \rightarrow \mathbf{w}_0 \text{ quando } n \rightarrow \infty \quad (1)$$

O algoritmo LMS é dito ser *convergente na média dos quadrados* se o valor do quadrado médio do sinal erro $e(n)$ aproxima-se de um valor constante quando o número de iterações n aproxima-se do infinito, isto é,

$$E[e^2(n)] \rightarrow \text{constante quando } n \rightarrow \infty \quad (2)$$

No 1o. caso, para garantia de convergência, o parâmetro de velocidade de aprendizado deve ser escolhido como:

$$0 < h < 2/\lambda_{max} \quad (3)$$

onde λ_{max} é o maior auto-valor da matriz de autocorrelação $R = E[\mathbf{x}(n)\mathbf{x}^T(n)]$. No 2o. caso, este parâmetro deve ser escolhido como

$$0 < h < 2/tr[R] \quad (4)$$

onde $tr[R]$ denota o traço da matriz R ¹. Desta forma sabendo que a matriz de correlação é definida por:

$$R = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

Pede-se determinar o valor da taxa de aprendizado η para

¹Traço de uma matriz é a soma dos elementos da sua diagonal principal

- (a) o algoritmo LMS ser convergente na média.
(b) o algoritmo LMS ser convergente na média dos quadrados.
4. Considere um conjunto de pontos referentes a uma classe \mathcal{C}_1 está distribuído próximo ao ponto (8,5) no plano XY . Os pontos não pertencentes à classe \mathcal{C}_1 são externos à região em que os pontos de \mathcal{C}_1 são amostrados e correspondem à classe \mathcal{C}_2 . Mostre a solução de uma rede perceptron para fazer a separação das duas classes \mathcal{C}_1 e \mathcal{C}_2 , sem utilizar algoritmos de treinamento convencionais. Obtenha a solução fazendo apenas a interpretação geométrica do problema. Discuta também como a variância da distribuição dos elementos da classe \mathcal{C}_1 afeta o problema da classificação.
5. Verifique que as equações 5, 6, 7 e 8 abaixo

$$\text{sgn}(v) = \begin{cases} +1, & \text{se } v > 0 \\ -1, & \text{se } v < 0 \end{cases} \quad (5)$$

$$y(n) = \text{sgn}[\mathbf{w}^T(n)\mathbf{x}(n)] \quad (6)$$

$$d(n) = \begin{cases} +1, & \text{se } \mathbf{x}(n) \text{ pertence a classe } \mathcal{C}_1 \\ -1, & \text{se } \mathbf{x}(n) \text{ pertence a classe } \mathcal{C}_2 \end{cases} \quad (7)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta[d(n) - y(n)]\mathbf{x}(n) \quad (8)$$

que resumem o algoritmo de convergência do perceptron, são consistentes com:

1. “Se o n -ésimo termo do conjunto de treinamento $\mathbf{x}(n)$ é corretamente classificado por $\mathbf{w}(n)$ computado na n -ésima iteração do algoritmo, nenhuma correção é feita ao vetor de pesos de acordo com a regra
 - (a) $\mathbf{w}(n+1) = \mathbf{w}(n)$, se $\mathbf{w}^T\mathbf{x}(n) > 0$ e $\mathbf{x}(n)$ pertence à classe \mathcal{C}_1 .
 - (b) $\mathbf{w}(n+1) = \mathbf{w}(n)$, se $\mathbf{w}^T\mathbf{x}(n) \leq 0$ e $\mathbf{x}(n)$ pertence à classe \mathcal{C}_2 .
2. Caso contrário, o vetor de pesos é atualizado de acordo com a regra
 - (a) $\mathbf{w}(n+1) = \mathbf{w}(n) - \eta(n)\mathbf{x}(n)$, se $\mathbf{w}^T(n)\mathbf{x}(n) > 0$ e $\mathbf{x}(n)$ pertence à classe \mathcal{C}_2 .
 - (b) $\mathbf{w}(n+1) = \mathbf{w}(n) + \eta(n)\mathbf{x}(n)$, se $\mathbf{w}^T(n)\mathbf{x}(n) \leq 0$ e $\mathbf{x}(n)$ pertence à classe \mathcal{C}_1 .onde $\eta(n)$ controla o ajuste aplicado ao vetor de pesos na iteração n .”

References

- [1] A. P. Braga, A. P. L. F. Carvalho, T. B. Ludermir, *Redes Neurais Artificiais - Teoria e Aplicações*, 2a. edição. LTC, 2007.
- [2] S. Haykin, *Neural networks - a comprehensive foundation*, 2nd. ed. Prentice Hall, 1999.
- [3] R. A. F. Romero, “SCC 5809 - Redes Neurais,” Lista de Exercícios 1, 2o. semestre de 2010.