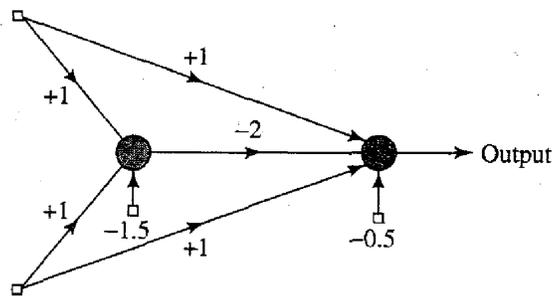


ICMC-USP
 Lista de Exercícios 4 - Perceptrons Multicamadas
 SCC-5809 - Redes Neurais
 2o. Semestre de 2011 - Prof. João Luís



1. A figura 1 abaixo mostra uma rede neural com um único neurônio escondido. Mostre que essa rede resolve o problema do XOR construindo (a) regiões de decisão e (b) uma tabela-verdade para a rede.

Figure 1: Rede Neural com único neurônio escondido.



2. À constante *momentum* α tem-se normalmente atribuído um valor positivo na faixa $0 \leq \alpha < 1$. Investigue a diferença que ocorreria no comportamento da equação 1 com respeito ao tempo t se à α for atribuído um valor negativo na faixa $-1 < \alpha \leq 0$.

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^n \alpha^{n-t} \frac{\partial \xi(t)}{\partial w_{ji}(t)} \quad (1)$$

3. Considere o exemplo simples de uma rede envolvendo um único peso, para a qual a função custo é

$$\xi(w) = k_1(w - w_0)^2 + k_2 \quad (2)$$

onde w_0 , k_1 e k_2 são constantes. Um algoritmo *back-propagation* com *momentum* é usado para minimizar $\xi(w)$. Explore a forma na qual a inclusão da constante *momentum* α influencia o processo de aprendizado, com referência ao número de épocas requeridas para a convergência versus α .

4. Quando se justifica utilizar num MLP duas camadas (uma escondida) ou três camadas (duas escondidas) e que tipos de funções podem ser implementadas em cada caso?
5. Que taxa de aprendizado η você usaria para treinar um MLP quando os padrões de uma mesma classe são semelhantes a alguns padrões de outra classe? E se forem bastante diferentes?

6. O que são mínimos locais? Como sair deles?
7. Por que o *back-propagation* é chamada de regra delta generalizada?
8. Por que se utiliza gradiente descendente para treinar MLP?
9. O que é *overfitting*?
10. Explique o compartilhamento de pesos, presente nas redes convolucionais.
11. [EXPERIMENTO COMPUTACIONAL] Investigue o emprego do aprendizado *back-propagation* usando uma não-linearidade sigmoideal para realizar mapeamentos um-a-um, como descritos abaixo:
 - (1) $f(x) = \frac{1}{x}$ $1 \leq x \leq 100$
 - (2) $f(x) = \log_{10}x$ $1 \leq x \leq 10$
 - (3) $f(x) = \exp(-x)$ $1 \leq x \leq 10$
 - (4) $f(x) = \text{sen } x$ $0 \leq x \leq \frac{\pi}{2}$

Para cada mapeamento, faça o seguinte:

- a) Estabeleça dois conjuntos de dados, um para o treinamento da rede e outro para teste.
- b) Use o conjunto de dados de treinamento para computar os pesos sinápticos da rede, assumindo haver uma única camada escondida.
- c) Avalie a acurácia computacional da rede usando os dados de teste.

Use apenas uma camada escondida mas com número variável de neurônios escondidos. Investigue como a performance da rede é afetada variando-se o tamanho da camada escondida.

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 2, 2o. semestre de 2010.