

## CPE 721 – RNs Feedforward

### 1ª Série de Exercícios – Otimização

**Obs: O objetivo da série de exercícios é a fixação do aprendizado. A série pode ser feita em grupo, mas é importante que cada um tente achar as soluções individualmente antes do trabalho em grupo.**

1 - Considere a função

$$F(\vec{w}) = w_1^3 w_2^2 - 2w_1^3 w_2 - 3w_1 w_2^2 + 6w_1 w_2$$

1.1 – Calcule a expressão analítica do gradiente e da Hessiana.

1.2 – Calcule o valor da função, do gradiente e da Hessiana nos quatro pontos indicados à seguir e interprete os resultados em termos de singularidades e extremos da função. Pontos  $(w_1, w_2)$ : (0,0), (1,1), (-1, 1), (1, -1)

1.3 – Escreva a expressão analítica da curva de nível que passa pelo ponto (1, -1).

1.4 – Calcule a expressão da tangente a curva de nível no ponto (1, -1). Sugestão: lembre das propriedades do gradiente.

1.5 – Supondo que a aproximação quadrática é válida e partindo do ponto (1, -1), calcule o extremo usando Newton Raphson

$$\vec{w}_{extremo} = \vec{w}_0 - \vec{H}^{-1}(\vec{w}_0) \nabla(\vec{w}_0)$$

Pela análise do gradiente verifique se o ponto encontrado é mesmo um extremo. Comente.

1.6 - Repita o item 1.5 a partir da origem (0, 0). Comente.

1.7 – Novamente a partir do ponto (1, -1) calcule e aplique o passo ótimo para minimização (sem usar a Hessiana, usando otimização em linha) via gradiente descendente. Analise o ponto encontrado.

2 – Considerando que a aproximação quadrática é válida mostre que para que uma função objetivo diminua seu valor é necessário que o passo seja menor que o dobro do passo ótimo,  $\alpha < 2\alpha_{ótimo}$ .

3 - Considere a função  $F(\vec{w}) = \left[ (w_1^2 + w_2^2 - 1) w_1 w_2 \right]^2$ . Calcule o valor exato e o da aproximação da Hessiana proposta por Levenberg-Marquardt para os pontos  $\vec{w}_a = [0.7000 \ 0.7000]^t$  e  $\vec{w}_b = [1.000 \ 1.000]^t$ . Compare os erros e comente a diferença.

Observe que  $F(\vec{w})$  tem calhas (mínimos) nos eixos  $w_1$ ,  $w_2$  e em uma circunferência de raio unitário centrada na origem.

4 – Considerando que a aproximação de segunda ordem para  $F(w_i)$  é válida crie uma fórmula que permita calcular o passo  $\alpha_{ótimo}$  para uma sinapse do algoritmo BP resiliente que leve em um passo ao ponto de gradiente nulo.