CPE 721 – RNs Feedforward
6ª Série de Exercícios

1 – Apresente a fórmula de $\frac{\partial F}{\partial \tilde{y}}$ a ser utilizada para minimizar o erro MAPE em um processo backpropagation. Obs: combine os processos de erro absoluto e erro relativo sem escalamento.

$$F(\underline{w}) = \underset{k}{E} \sum_{l} \frac{\left|Y_{kl} - \tilde{Y}_{kl}\right|}{\left|Y_{kl}\right|} \qquad \text{(redefinido escalar)}$$

Para cada par *k* para cada saída *l*

$$F_{kl} = \frac{\left|Y_{kl} - \tilde{Y}_{kl}\right|}{\left|Y_{kl}\right|}$$

Definindo em função das variáveis escaladas

$$y_{kl} = \frac{1}{\sigma_{Y_l}}\left(Y_{kl} - \mu_{Y_l}\right) \qquad Y_{kl} = \sigma_{Y_l} y_{kl} + \mu_{Y_l} \qquad \tilde{Y}_{kl} = \sigma_{Y_l} \tilde{y}_{kl} + \mu_{Y_l}$$

$$F_{kl} = \frac{1}{\left|y_{kl} + \frac{\mu_{Y_l}}{\sigma_{Y_l}}\right|}\left|y_{kl} - \tilde{y}_{kl}\right| = a_{kl}\left|\varepsilon_{kl}\right| \qquad \varepsilon_{kl} = y_{kl} - \tilde{y}_{kl} \qquad a_{kl} = \frac{1}{\left|y_{kl} + \frac{\mu_{Y_l}}{\sigma_{Y_l}}\right|}$$

$$\frac{\partial F_{kl}}{\partial \tilde{y}_{kl}} = a_{kl} \frac{\partial\left|\varepsilon_{kl}\right|}{\partial \tilde{y}_{kl}}$$

$$\frac{\partial F_{kl}}{\partial \tilde{y}_{kl}} = a_{kl}\,\frac{\partial\left|\varepsilon_{kl}\right|}{\partial \tilde{y}_{kl}}$$

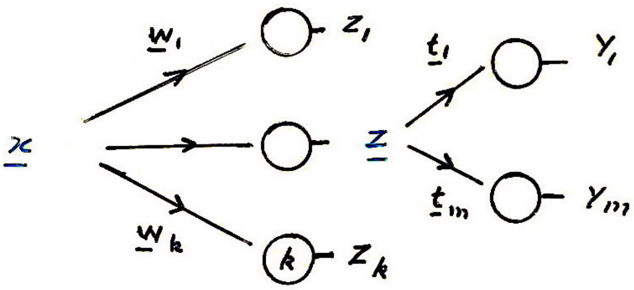
$$|\,\varepsilon_{kl}\,| = \begin{cases} \varepsilon_{kl} = y_{kl} - \tilde{y}_{kl} & se \quad \varepsilon_{kl} > 0 \\ -\varepsilon_{kl} = -(y_{kl} - \tilde{y}_{kl}) & se \quad \varepsilon_{kl} < 0 \end{cases}$$

$$\frac{\partial\,|\,\varepsilon_{kl}\,|}{\partial \tilde{y}_{kl}} = \left\{ \begin{array}{lll} -1 & se & \varepsilon_{kl} > 0 \\ +1 & se & \varepsilon_{kl} < 0 \end{array} \right\} = -sign(\varepsilon_{lk})$$

$$\frac{\partial F_{kl}}{\partial \tilde{y}_{kl}} = a_{kl}\,\frac{\partial\left|\varepsilon_{kl}\right|}{\partial \tilde{y}_{kl}} = \frac{-sign(\varepsilon_{lk})}{\left|y_{kl} + \dfrac{\mu_{Y_l}}{\sigma_{Y_l}}\right|}$$

$$\Delta w = -\alpha\,\frac{\partial F}{\partial w} = \alpha \underset{k}{E} \sum_{l} \frac{sign(y_{kl} - \tilde{y}_{kl})}{\left|y_{kl} + \dfrac{\mu_{Y_l}}{\sigma_{Y_l}}\right|}\,\frac{\partial \tilde{y}_{kl}}{\partial w}$$

2 – Considere um classificador com duas camadas de neurônios tipo tgh que deve poder ser retreinado a qualquer instante, i.e., cujo treinamento não paralise. Escreva o algoritmo de treinamento usando o método proposto no ítem 2 do cap 13.

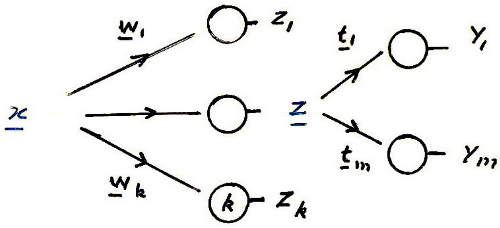


$$\tilde{y}_l = tgh(u_l) \qquad u_l = \sum_i t_{li} z_i \qquad z_i = \beta v_i + tgh(v_i) \qquad v_i = \sum_i w_{ij} x_i$$

$$F_{ER} = \underset{\forall\, pares}{E} \sum_l \left[(1+y_l)\ln\frac{1+y_l}{1+\tilde{y}_l} + (1-y_l)\ln\frac{1-y_l}{1-\tilde{y}_l} \right]$$

Para um par $\quad \Delta p = -\alpha \dfrac{\partial F}{\partial p}$

Camada de saída



$$\Delta t_{li} = -\alpha \frac{\partial F}{\partial t_{li}} = -\alpha \frac{\partial F}{\partial \tilde{y}_l} \frac{\partial \tilde{y}_l}{\partial u_l} \frac{\partial u_l}{\partial t_{li}}$$

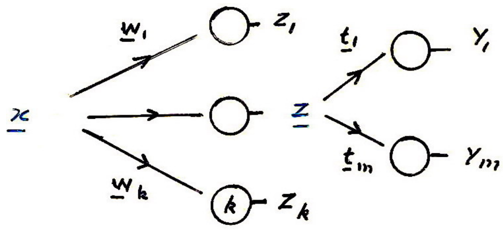
$$F_{ER} = \sum_l \left[(1+y_l)\ln\frac{1+y_l}{1+\tilde{y}_l} + (1-y_l)\ln\frac{1-y_l}{1-\tilde{y}_l} \right] \qquad \frac{\partial F_{ER}}{\partial \tilde{y}_l} = \frac{1}{1-\tilde{y}_l^2}$$

$$\tilde{y}_l = tgh(u_l) \qquad \frac{\partial \tilde{y}_l}{\partial u_l} = 1 - \tilde{y}_l^2$$

$$u_l = \sum_i z_i t_{li} \qquad \frac{\partial u_l}{\partial t_{li}} = z_i$$

$$\Delta t_{li} = -\alpha \frac{\partial F}{\partial \tilde{y}_l} \frac{\partial \tilde{y}_l}{\partial u_l} \frac{\partial u_l}{\partial t_{li}} = -\alpha \frac{1}{1-\tilde{y}_l^2}(1-\tilde{y}_l^2) z_i = -\alpha z_i$$

Camada intermediária



$$\Delta w_{ij} = -\alpha \frac{\partial F}{\partial w_{ij}} = -\alpha \sum_l \frac{\partial F}{\partial \tilde{y}_l} \frac{\partial \tilde{y}_l}{\partial u_l} \frac{\partial u_l}{\partial z_i} \frac{\partial z_i}{\partial v_i} \frac{\partial v_i}{\partial w_i}$$

$$F_{ER} = \sum_l \left[(1+y_l)\ln\frac{1+y_l}{1+\tilde{y}_l} + (1-y_l)\ln\frac{1-y_l}{1-\tilde{y}_l} \right] \qquad \frac{\partial F_{ER}}{\partial \tilde{y}_l} = \frac{1}{1-\tilde{y}_l^2}$$

$$\tilde{y}_l = tgh(u_l) \qquad \frac{\partial \tilde{y}_l}{\partial u_l} = 1-\tilde{y}_l^2$$

$$u_l = \sum_i z_i t_{li} \qquad \frac{\partial u_l}{\partial z_i} = t_{li}$$

$$z_i = \beta v_i + tgh(v_i) \qquad \frac{\partial z_i}{\partial v_i} = \beta + 1 - (z_i - \beta v_i)^2$$

$$v_i = \sum_i w_{ij} x_j \qquad \frac{\partial v_i}{\partial w_{ij}} = x_j$$

$$\Delta w_{ij} = -\alpha \frac{\partial F}{\partial w_{ij}} = -\alpha \sum_l \frac{\partial F}{\partial \tilde{y}_l} \frac{\partial \tilde{y}_l}{\partial u_l} \frac{\partial u_l}{\partial z_i} \frac{\partial z_i}{\partial v_i} \frac{\partial v_i}{\partial w_i} =$$

$$= -\alpha \sum_l \frac{1}{1-\tilde{y}_l^2} (1-\tilde{y}_l^2) t_{li} \left[\beta + 1 - (z_i - \beta v_i)^2 \right] x_j =$$

$$= -\alpha \sum_l t_{li} \left[\beta + 1 - (z_i - \beta v_i)^2 \right] x_j$$