

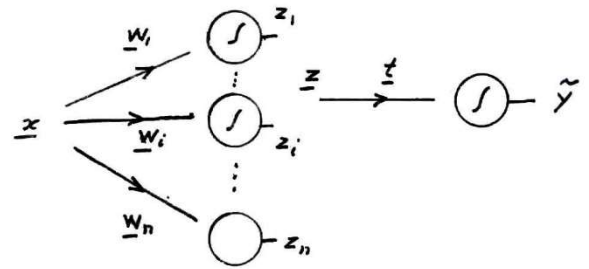
1 – Paralisia ao término do treinamento

Classificadores - Fim do treinamento

Valores ótimos das sinapses

$$\varepsilon \rightarrow 0 \quad |\tilde{y}_i| \rightarrow 1 \quad \left| \underline{z}^t \underline{t} \right| \rightarrow \infty$$

$$|\underline{t}| \rightarrow \infty$$



$$\varepsilon \rightarrow 0 \quad |z_i| \rightarrow 1 \text{ (max)} \quad |u_i| = |\underline{x}^t \underline{w}| \rightarrow \infty$$

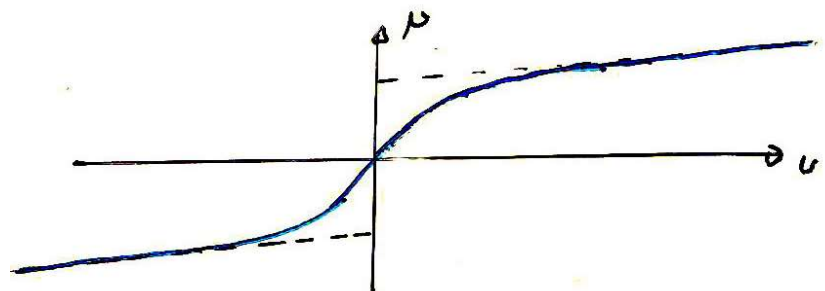
$$|\underline{w}| \rightarrow \infty$$

O treinamento paralisa ao término !

2 - Rede sem paralisia de treinamento

Camada intermediária

usar a função de ativação
não saturada



$$v = \beta u + \tanh(u) \quad 0 < \beta \ll 1$$

$$\frac{\partial v}{\partial u} = 1 + \beta - (v - \beta u)^2$$

Quickprop

Camada de saída em classificadores

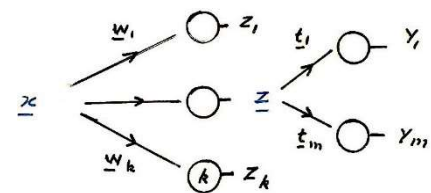
Usar entropia relativa como função objetivo (ver CPE 721).

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

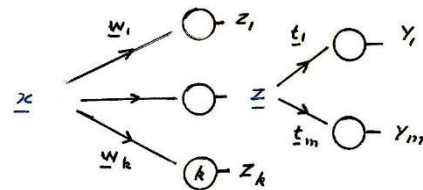
Se a distribuição do erro $p(\epsilon)$ for Gaussiana, os minimante da ER e do EMQ coincidem

Por BP regra delta

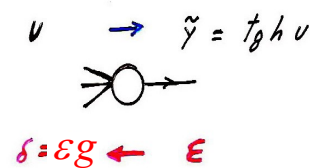
$$\Delta w = -\alpha \frac{\partial f}{\partial w} = -\alpha \frac{\partial f}{\partial \tilde{y}} \frac{\partial \tilde{y}}{\partial u} \frac{\partial u}{\partial w}$$



Camada de saída



EMQ $\Delta t_i = 2\alpha \epsilon (1 - \tilde{y}^2) z_i$



ER $\Delta t_i = 2\alpha \epsilon z_i$

