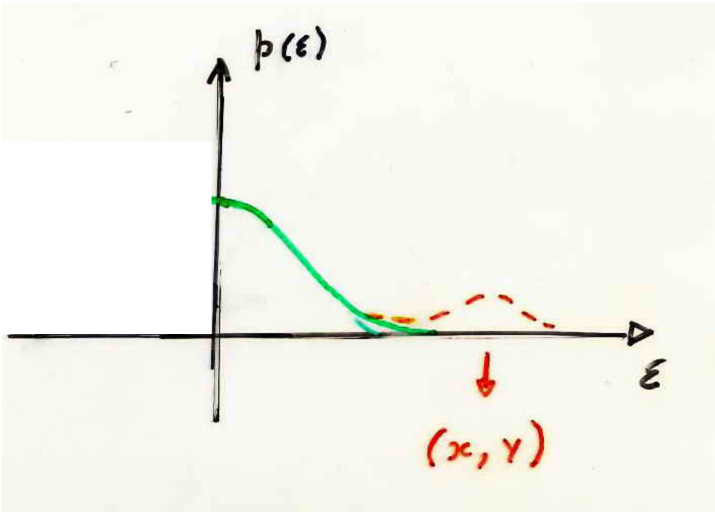
13 - Histograma dos erros

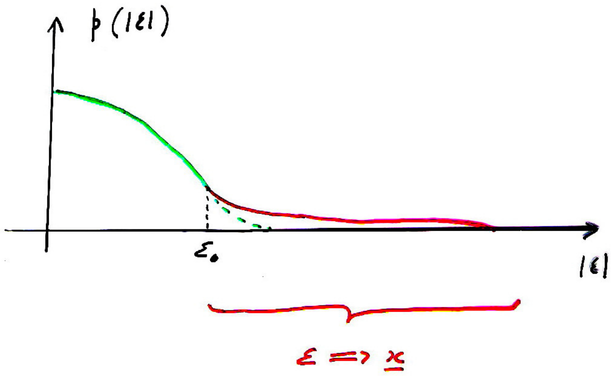
Anomalias, Localização, Correções



Intrusos ou Regiões de baixa população

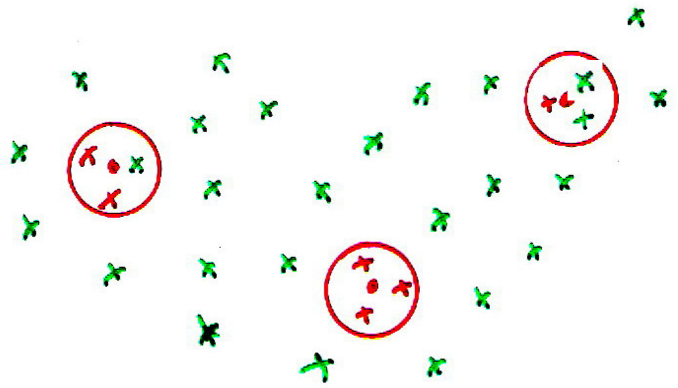
Localização das anomalias

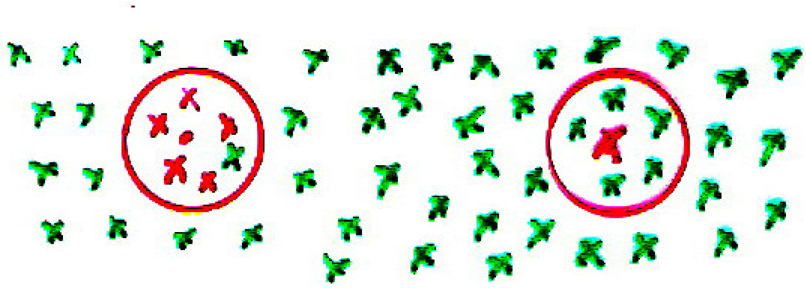
$$(\underline{x}, y) \Rightarrow (\underline{x}, y, \bar{y}, \varepsilon) \Rightarrow (\underline{x}, \varepsilon)$$



$$|\varepsilon| > \varepsilon_0 \qquad \underline{x} \in X_0$$

Clusterizar $\left.\begin{matrix} \underline{x} \\ \underline{y} \end{matrix}\right] \quad \forall\, \underline{x} \in X_0$





Correção das anomalias

Análise das anomalias

Intrusos - eliminar

Baixa população

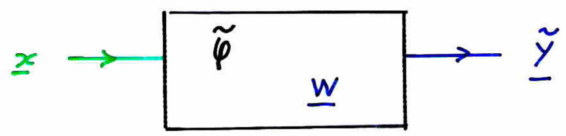
Balancear as populações

Mapeamentos complexos

Inserção de mais neurônios na região

neurônios tipo RBF

14 – Outros critérios de erro



Composição de dF/dw

a partir

do critério de erro

e

da estrutura da rede

$F(\underline{y}, \tilde{\underline{y}}) \qquad \tilde{\underline{y}}(\underline{w})$

$$\frac{\partial F}{\partial w_i} = \frac{\partial F}{\partial \tilde{y}} \frac{\partial \tilde{y}}{\partial w_i}$$

depende da função erro

depende da estrutura do modelo

saídas múltiplas

$$\frac{\partial F}{\partial w_i} = \sum_l \frac{\partial F}{\partial \tilde{y}_l} \frac{\partial \tilde{y}_l}{\partial w_i}$$

14.1 Caso geral: Erro médio quadrático ponderado

$$F = \underset{k}{E}\left\{\sum_l a(k,l)\ \varepsilon_l^2\right\}$$

onde a(l) peso variável por saída l

a(k) peso variável por par k

como $\varepsilon_l = y_l - \tilde{y}_l$

$$\frac{\partial F}{\partial \tilde{y}} = -2\underset{k}{E}\left\{\sum_l a(k,l)\ \varepsilon_l\right\}$$

No treinamento tipo regra delta o acréscimo na sinapse w_{ij} sera dado por

$$\Delta w_{ij} = 2\alpha v_j \delta_i \quad onde \quad \delta_i = \sum_{l=1}^{m} a(k,l)\,\varepsilon_l g_{li}$$

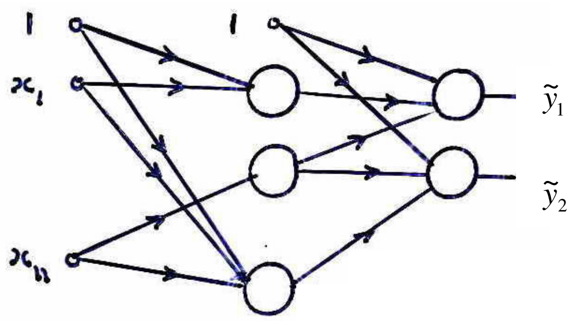
onde

v_j **é o sinal na entrada da sinapse na fase de propagação do sinal para a frente e**

$g_{li} = \dfrac{\partial \tilde{y}_l}{\partial u_i}$ **é o ganho linearizado da saída da sinapse w_{ij} à saída *l* da rede**

Ex 1:

A rede tem duas saídas, mas
o erro em y_1 é 3 vezes mais importante que o erro em y_2



a(k,1) = 3

a(k,2) = 1

$$F = \underset{k}{E}\left(3\varepsilon_1^2 + \varepsilon_2^2\right)$$

$$\Delta w_{ij} = 2\alpha v_j \delta_i \quad onde \quad \delta_i = 3\varepsilon_1 g_{1i} + \varepsilon_2 g_{2i}$$

Ex 2:

Erro <u>relativo</u> médio quadrático nas variáveis originais, <u>não escaladas</u>

$$F(\underline{w}) = E\left\{\sum_l \left[\frac{Y_l - \tilde{Y}_l}{Y_l}\right]^2\right\}$$

$$y_l = \frac{1}{\sigma_{Y_l}}\left(Y_l - \mu_{Y_l}\right) \qquad Y_l = \sigma_{Y_l} y_l + \mu_{Y_l}$$

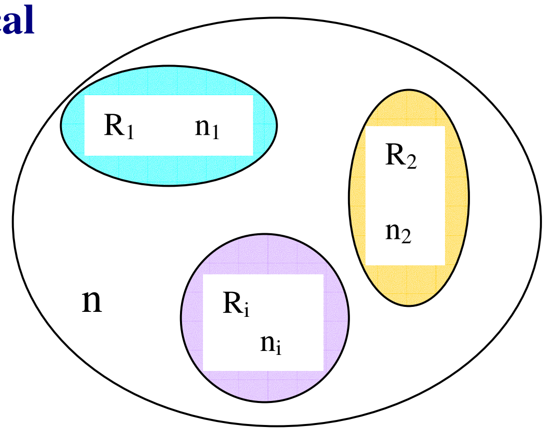
$$\tilde{Y}_l = \sigma_{Y_l} \tilde{y}_l + \mu_{Y_l}$$

$$F(\underline{w}) = E\left\{\sum_l \left[\frac{Y_l - \tilde{Y}_l}{Y_l}\right]^2\right\} = E\left\{\sum_l \frac{1}{\left(y_l + \frac{\mu_{Y_l}}{\sigma_{Y_l}}\right)^2}(y_l - \tilde{y}_l)^2\right\} \qquad a(k,l) = \frac{1}{\left(y_l + \frac{\mu_{Y_l}}{\sigma_{Y_l}}\right)^2}$$

No treinamento tipo BP o acréscimo na sinapse w_{ij} é dado por

$$\Delta w_{ij} = 2\alpha v_j \delta_i \quad onde \quad \delta_i = \sum_{l=1}^{m} \frac{1}{\left(y_l + \frac{\mu_{Y_l}}{\sigma_{Y_l}}\right)^2} \varepsilon_l g_{li}$$

Ex 3: Correção do efeito da população local



Cada par pertençe a uma região R_i

Cada região R_i tem uma população n_i

A população total é $n = \sum_{\forall i} n_i$

$$F = \sum_{\forall\, região\, R_i} \boxed{\sum_{\forall\, par \in R_i} f(y, \tilde{y})}$$

peso: n_i pares

peso: 1 par, independe de n_i

$$F = \sum_{\forall\, região\, R_i} \boxed{\frac{1}{n_i} \sum_{\forall\, par \in R_i} f(y, \tilde{y})}$$

$$a(i,l) = a(i) = 1/n_i$$

n_i – população da região onde o par pertence.

14.2 Erros não quadráticos

14.2.1 Erros em potências maiores

$$F(\underline{w}) = E\left\{ \sum_l (y_l - \tilde{y}_l)^{2n+2} \right\} \qquad n = 1,2,\ldots$$

prioriza a redução dos maiores erros, mas F(w) é mais abrupta.

14.2.2 Erro absoluto

$$F(\underline{w}) = E\left\{ |\, \vec{y} - \tilde{\vec{y}}\, | \right\} = E\left\{ \sqrt{\sum_l (y_l - \tilde{y}_l)^2} \right\}$$

não prioriza a redução dos maiores erros como o emq

- não derivável na origem, necessita procedimentos especiais

14.2.3 Erro MAPE – Mean Absolute Percentual Error

$$F(\underline{w}) = E\left\{ \frac{\left| \vec{Y} - \tilde{\vec{Y}} \right|}{\left| \vec{Y} \right|} \right\} = E\left\{ \sqrt{\frac{\sum_l \left(Y_l - \tilde{Y}_l\right)^2}{\sum_l \left(Y_l\right)^2}} \right\}$$

variáveis não escaladas

não prioriza a redução dos maiores erros percentuais, como o epmq.

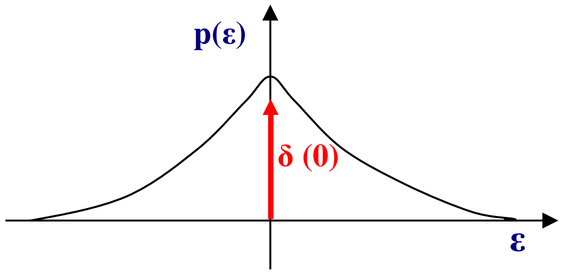
- não derivável na origem, necessita procedimentos especiais

14.3 Entropia

O objetivo na minimização do erro ε é

obter o histograma de p(ε) com média nula (fácil) e o mais estreito possível

ideal: p(ε) = δ (0) (delta de Dirac, erro nulo)



Se a distribuição do erro p(ε) for Gaussiana, o objetivo é alcançado minimizando o erro médio quadrático (a variância de p(ε))

Se a distribuição do erro p(ε) não for Gaussiana, o objetivo é alcançado minimizando a entropia (ver os trabalhos de J.C. Príncipe)