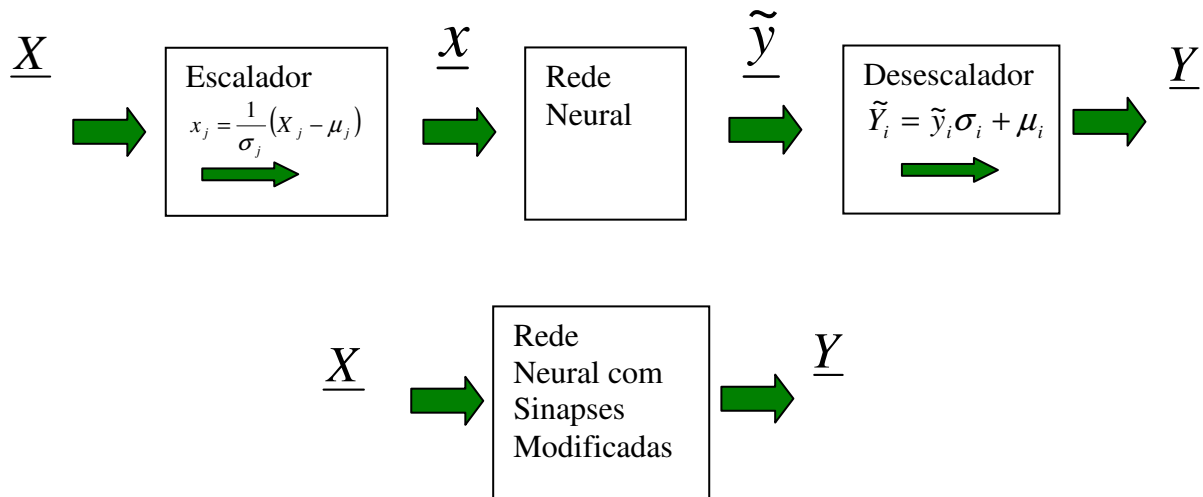


## Pós treinamento - Absorção do escalamento pelas sinapses

### Rede em Operação



O escalamento/desescalamento podem ser vistos como a ação de uma camada linear entre as variáveis reais e as escaladas. Assim, na fase de operação da rede o escalamento/desescalamento podem ser absorvido pela primeira/última camada da rede.

### Entrada $X_j$

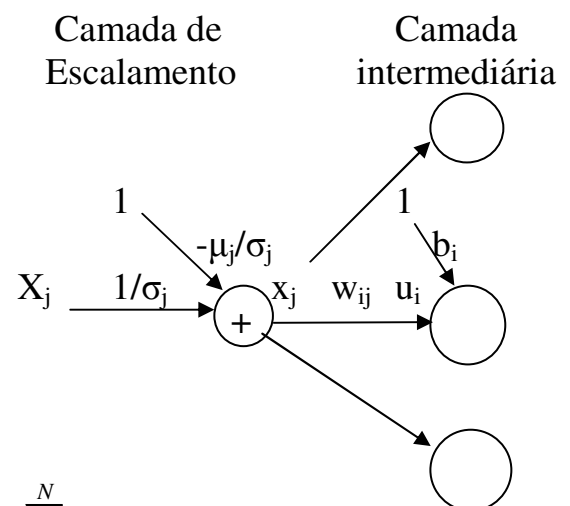
**Critério: manter os  $u_i$ 's inalterados**

**Para  $X_j$  contínuo**

$$x_j = \frac{1}{\sigma_j} (X_j - \mu_j)$$

$$u_i = \sum_{j=1}^N w_{ij} x_j + b_i = \sum_{j=1}^N \frac{w_{ij}}{\sigma_j} X_j - \sum_{j=1}^N \frac{\mu_j}{\sigma_j} + b_i = \sum_{j=1}^N W_{ij} X_j + B_i$$

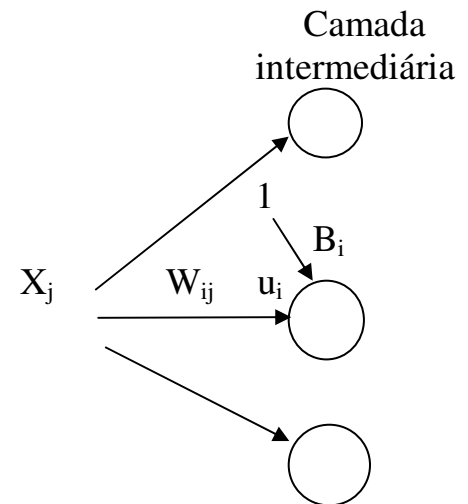
**onde**  $W_{ij} = \frac{w_{ij}}{\sigma_j}$  e  $B_i = -\sum_{j=1}^N \frac{\mu_j}{\sigma_j} + b_i$



**Na nova rede,**

$$u_i = \sum_{j=1}^N W_{ij} X_j + B_i$$

**onde**  $W_{ij} = \frac{w_{ij}}{\sigma_j}$     e     $B_i = -\sum_{j=1}^N \frac{\mu_j}{\sigma_j} + b_i$



**se  $X_j$  discreto,**     $X_j \in \{0,1\}$

**para que os  $u_i$  não se alterem aplique a mesma transformação considerando**

$$\sigma_j = \mu_j = 0,5$$

**Saída  $Y_j$**

**Critério: manter os  $z_i$ 's inalterados**

**Para  $Y_j$  contínuo**

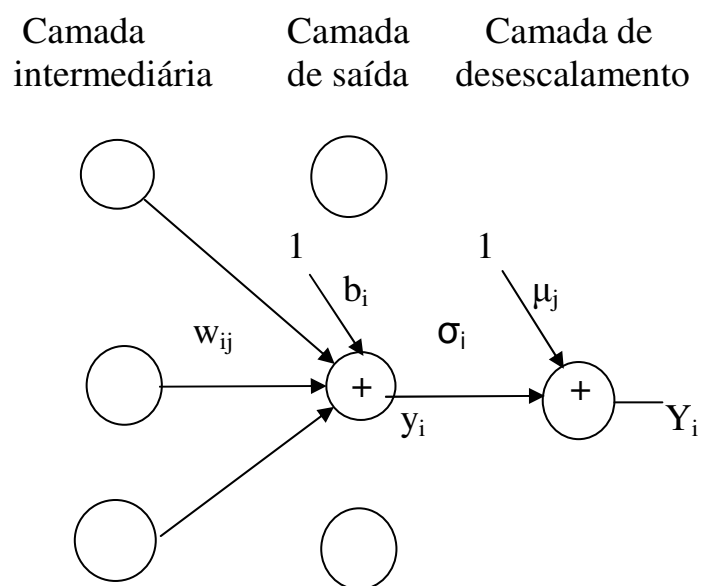
$$y_i = \frac{1}{\sigma_i} (Y_i - \mu_i) \quad Y_i = y_i \sigma_i + \mu_i$$

$$y_i = \sum_{j=1}^N w_{ij} z_j + b_i$$

$$Y_i = \sigma_i y_i + \mu_i = \sum_{j=1}^N \sigma_i w_{ij} z_j + \sigma_i b_i + \mu_i$$

$$= \sum_{j=1}^N W_{ij} z_j + B_i$$

**onde**     $W_{ij} = w_{ij} \sigma_i$     e     $B_i = \sigma_i b_i + \mu_i$

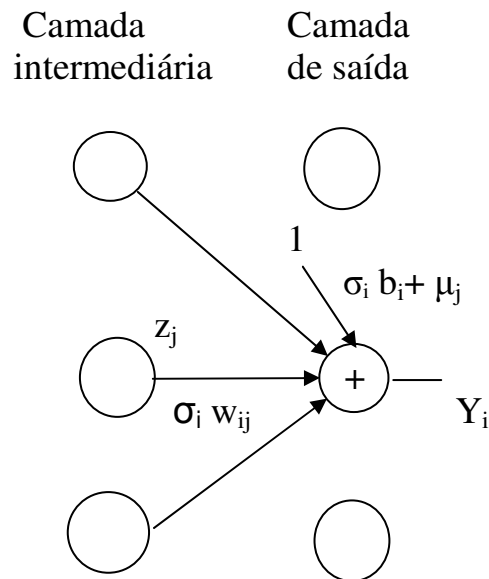


**Na nova rede:**

$$Y_i = \sum_{j=1}^N W_{ij} z_j + B_i$$

**onde**

$$W_{ij} = w_{ij} \sigma_i \quad e \quad B_i = \sigma_i b_i + \mu_i$$



se  $Y_j$  discreto,  $Y_i \in \{0,1\}$ , apenas substitua na camada de saída o neurônio tipo  $\text{tgh}(u)$  por um neurônio tipo  $\text{degrau}(u)$ , mantendo as sinapses do neurônio de saída inalteradas

**Resumo:**

**Variáveis contínuas:**

| Camada de entrada                        |                                                    | Camada de Saída                                          |                              |
|------------------------------------------|----------------------------------------------------|----------------------------------------------------------|------------------------------|
| Rede escalada, original                  | Rede modificada                                    | Rede escalada, original                                  | Rede modificada              |
| $x_j = \frac{1}{\sigma_j} (X_j - \mu_j)$ | $X_j$                                              | $\tilde{y}_i = \frac{1}{\sigma_i} (\tilde{Y}_i - \mu_i)$ | $\tilde{Y}_i$                |
| $w_{ij}$                                 | $W_{ij} = \frac{w_{ij}}{\sigma_j}$                 | $w_{ij}$                                                 | $W_{ij} = w_{ij} \sigma_i$   |
| $b_i$                                    | $B_i = -\sum_{j=1}^N \frac{\mu_j}{\sigma_j} + b_i$ | $b_i$                                                    | $B_i = \sigma_i b_i + \mu_i$ |

**Variáveis discretas**

se  $X_j$  discreto,  $x_j \in \{-1,1\}$  e  $X_j \in \{0,1\}$  considere  $\sigma_j = \mu_j = 0,5$

se  $Y_j$  discreto,  $Y_i \in \{0,1\}$ , neurônio  $\text{tgh}(u) \Rightarrow \text{degrau}(u)$ , sinapses inalteradas