

## CPE 721 4ª Série de Exercícios

### 1 - Absorção do escalamento:

1 - Na rede neural abaixo os neurônios 1 e 3 são lineares, e o neurônio 2 é tipo  $\text{tgh}(\cdot)$ . As variáveis de entrada e saída tem distribuição aproximadamente Gaussiana, com parâmetros conforme a tabela 1 abaixo. Após o treinamento com as entradas e saídas normalizadas para média zero e desvio padrão unitário as sinapses obtidas para a rede estão apresentadas na tabela 2 abaixo. Apresente os valores das sinapses a serem usadas com a rede com as entradas originais, não normalizadas.

Obs: Examine com cuidado o caso das sinapses conectadas ao neurônio 1, que ao mesmo tempo pertence à camada intermediária e fornece uma saída. Aplique a denormalização para as entradas  $x$ , em seguida para a saída  $y_a$  e finalmente como neurônio da camada intermediária para a denormalização de  $y_b$ .

Tabela 1

	$\mu$	$\sigma$
$x_a$	5	2
$x_b$	10	3
$y_a$	4	2
$y_b$	10	3

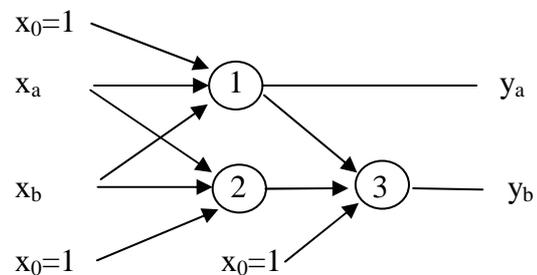


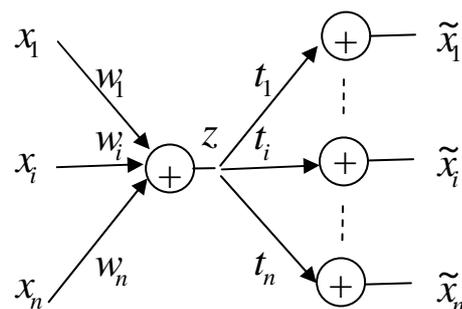
Tabela 2 - Sinapses  $W_{ij}$

$i \downarrow j \Rightarrow$	a	b	0	1	2
1	.5	1.1	.7	-	-
2	-.3	.2	0	-	-
3	-	-	.2	.3	-.3

### 2 - Rede Neural Auto-supervisionada - Compressão da informação.

2.1 - Considere um conjunto de vetores  $\underline{x}$  de  $n$  dimensões. A chamada primeira componente principal visa a representação de cada vetor  $\underline{x}$  por uma variável  $z = \underline{w}^t \underline{x}$  com uma única dimensão tal que permita a reconstituição de uma aproximação  $\tilde{\underline{x}}$  de  $\underline{x}$ ,  $\tilde{\underline{x}} = z \underline{t}$ , que minimiza o erro  $F$  da aproximação. As equações do processo usando os vetores  $\underline{w}$  e  $\underline{t}$  são:

$$z = \underline{w}^t \underline{x} \quad \tilde{\underline{x}} = z \underline{t} \quad F = E_{\underline{v}, \underline{x}} \left\{ \sum_{i=1}^n (x_i - \tilde{x}_i)^2 \right\}$$



A rede neural acima, chamada de rede neural autosupervisionada, realiza este processo. A primeira camada permite determinar a variável comprimida  $z$  e a segunda camada apresenta a aproximação obtida usando esta variável, para cada vetor  $\underline{x}$ . Apresente as fórmulas para o treinamento das sinapses  $t_i$  e  $w_i$ .

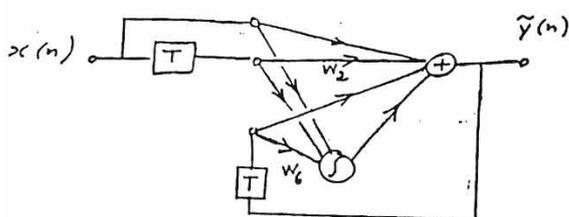
Obs: Considere que as variáveis  $x_i$  estão normalizadas para média zero e desvio padrão unitário, e observe que os neurônios não necessitam sinapses de polarização.

**2.2** - Para o exercício 2.1 acima considere que as variáveis normalizadas  $x_i$  foram geradas a partir das variáveis originais  $X_i$  com médias  $\mu_i$  e  $\sigma_i$ . Apresente as sinapses que deverão ser utilizadas para operação com entradas e saídas não normalizadas.

Obs: considere que o valor das sinapses de polarização na rede que utiliza entradas normalizadas é nulo, mas que os da rede que utiliza sinapses não normalizadas pode não sê-lo.

### 3 - Modelagem

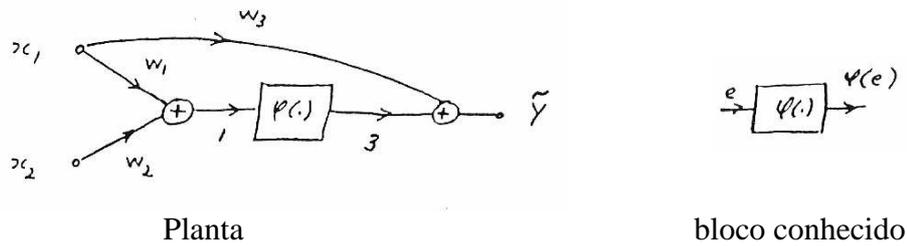
**3.1** - A figura abaixo mostra a estrutura da rede utilizada para simular um sistema dinâmico não-linear, quando em operação. A rede utiliza neurônios sem bias, um linear (+) e um tipo tgh (J) e atrasos unitários T. Parte da tabela usada no treinamento (série-paralelo, regra delta com  $\alpha = 0,1$ ) é apresentada a seguir.



n	x(n)	y(n)
⋮	⋮	⋮
48	.40	.35
49	.45	.50
50	.50	.60
51	.55	.65
52	.60	.75
⋮	⋮	⋮

Em um instante do treinamento, por acaso todas as sinapses são iguais a 0,25. É sorteado então  $n = 50$ , e o passo de treinamento é dado. Quais os novos valores das sinapses  $w_2$  e  $w_6$ ? Indique como realizou os cálculos!

**3.2** - O modelo de uma planta apresentado na figura abaixo inclui um bloco cuja função de transferência  $\varphi(\cdot)$  já é conhecida analiticamente e não necessita ser treinado.



deduza as fórmulas de treinamento de  $w_1$ ,  $w_2$  e  $w_3$  por backpropagation regra delta para minimizar o erro médio quadrático da saída  $y$  sobre um conjunto de pares de treinamento  $\{ \underline{x}, y \}$ .

Existem restrições sobre  $\varphi(\cdot)$  ? Caso sim, quais ?

#### 4 - Redes RBF

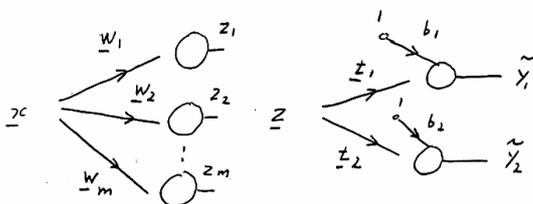
**4.1** - Na rede abaixo os neurônios da camada intermediária são do tipo RBF com função de transferência

$$z_i = \frac{1}{1 + u_i} \quad \text{onde} \quad u_i = (\underline{x} - \underline{w}_i)^t \underline{K} (\underline{x} - \underline{w}_i)$$

onde  $\underline{K}$  é uma matriz diagonal com elementos  $k_i$ . Os neurônios de saída são do tipo perceptron:

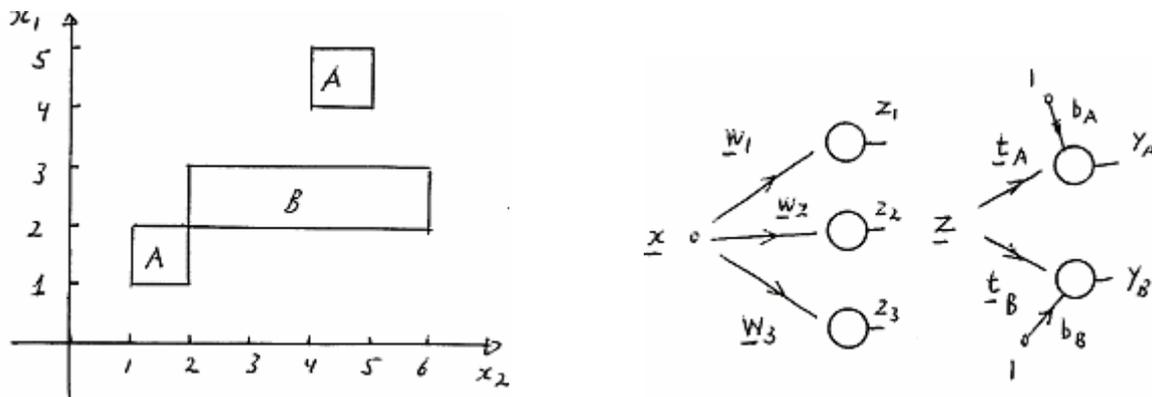
$$\tilde{y}_1 = u_1 \quad \text{onde} \quad u_1 = \underline{z}^t \underline{t}_1 + b_1 \quad e$$

$$\tilde{y}_2 = tgh(u_2) \quad \text{onde} \quad u_2 = \underline{z}^t \underline{t}_2 + b_2$$



A função objetivo a ser minimizada no treinamento é o erro médio quadrático na saída. Estabeleça as fórmulas para um treinamento tipo gradiente descendente (backpropagation).

4.2 – As duas classes abaixo devem ser separadas entre si e do restante do espaço de entrada pela rede RBF ao lado, da forma mais eficaz possível. Projete a rede.



Os neurônios da primeira camada são do tipo base radial:

$$u_i = k_i | \mathbf{x} - \mathbf{w}_i |^2$$

$$z_i = e^{-u_i}$$

E os da segunda camada do tipo tgh:

$$u_i = \mathbf{z}_i^t \mathbf{t}_i + b_i \quad y_i = \text{tgh}(u_i)$$

4.3 - Repita o exercício 4,2 considerando neurônios de “base elíptica”, isto é, com

$$u_i = (\mathbf{x} - \mathbf{w}_i)^t \mathbf{K}_i (\mathbf{x} - \mathbf{w}_i)$$

4.4 - Uma rede MLP com uma única camada com neurônios do tipo

$$u = \sum_{i=0}^{n+1} w_i x_i \quad \text{onde} \quad x_0 = 1 \quad e \quad x_{n+1} = \sum_{i=1}^n w_{n+1} x_i^2$$

$$v = \text{tgh}(u)$$

opera como um classificador com separadores esféricos.

a - Discuta a localização dos separadores em função do valor das sinapses

b - Apresente as fórmulas para o treinamento backpropagation.

c - Sugira o que usar como valores iniciais.

4.5 - Projete a rede de duas camadas para realizar o separador do exercício 4.2 em que os neurônios da primeira camada são do tipo definido no exercício 4.4 acima.

**4.6** - Uma rede MLP é treinada para fazer o mapeamento  $y(x)$  apresentado abaixo. A aproximação apresenta baixo erro, exceto no ponto indicado em vermelho (onde o erro é  $\varepsilon = y - \tilde{y} = -1,5$ ) e seus vizinhos imediatos. Isto ocorre devido ao mapeamento nesta região não conseguir ser adequadamente representado, porque o treinamento cai (quase sempre) em um mínimo local.

Para minorar este problema é acrescentado à rede treinada um neurônio RBF com função de ativação gaussiana  $v$  abaixo, e o treinamento é retomado do ponto onde havia parado. Indique os valores iniciais a atribuir a  $\sigma$ ,  $w_{n+1}$  e  $t_{n+1}$  para facilitar o treinamento. Explique o porque dos valores indicados.

$$y(x) \cong 2 \operatorname{sen} x \qquad v = e^{-\frac{1}{2\sigma^2}(x-w)^2}$$

