V1.0

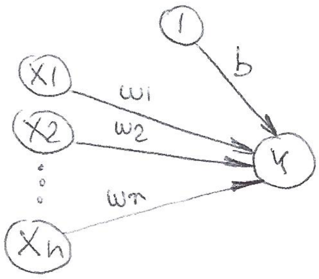
Rede ADALINE

ADALINE = ADAptive LInear NEuron (Widrow & Hoff em 1960).

IMPORTANCIA ⇒ pode-se combinar várias ADALINE e construir uma rede multicamadas – MADALINE.

⇒ regra DELTA ← Treinamento.

Arquitetura:



$$y\_in = b + \sum_{i=1}^{n} x_i \cdot w_i$$

Onde

$$\boxed{y = y\_in}$$

A FUNÇÃO DE ATIVAÇÃO É A FUNÇÃO IDENTIDADE

ALGORITMO DE APRENDIZADO

REGRA DELTA:

- Basea-se na minimização do Erro quadratico entre a saída e o valor "alvo".

  $$E = (t - y\_in)^2$$

- Pode ser utilizada em arquiteturas de varias saídas.

ALGORITMO DE TREINAMENTO:

⟹ Segue.

Algoritmo de Treinamento

PASSO Φ: - INICIALIZE OS PESOS (valores randomicos)

- ADOTE UMA taxa de aprendizado α
  (P/ UM ÚNICO neuronio na saída:
  $0,1 \leq n \cdot \alpha \leq 1,0$ onde
  n = número de entradas)

PASSO 1: Enquanto (condição de parada for falsa) faça

PASSO 2: Para cada par de treinamento $s:t$

PASSO 3: $x_i = s_i \quad (i = 1, \dots n)$

PASSO 4: $y\_in = b + \sum_{i=1}^{n} x_i \cdot w_i$

PASSO 5: $b(novo) = b(velho) + \alpha(t - y\_in)$

$$\boxed{w_i(novo) = w_i(velho) + \underbrace{\alpha(t - y\_in) \cdot x_i}_{\Delta w_i}}$$

Regra delta

PASSO 6: Se o maior valor "delta" do peso alterado é menor que a tolerancia especificada PARE; do contrario continue.

Nota: No caso da ADALINE ser utilizada como um classificador deve-se Aplicar uma função Degrau como se segue (depois de treinada).

ADALINE – CLASSIFICADOR:

PASSO ϕ: INICIALIZE OS PESOS (utilize os pesos da rede treinada)

PASSO 1: Para cada vetor bipolar faça:

PASSO 2: $x_i = s_i$;

PASSO 3: $y\_in = b + \sum_{i=1}^{n} x_i w_i$

PASSO 4: Aplique a função de Ativação

$$y = \begin{cases} 1 & \text{se } y\_in \geq 0 \\ -1 & \text{se } y\_in < 0 \end{cases}$$

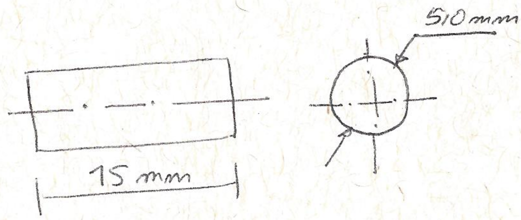
—//—

Subsidios:

① Erro Quadratico Médio

→ Suponha que uma grandeza assume os seguintes valores:

$$X = (5{,}0 \quad 5{,}1 \quad 5{,}0 \quad 4{,}9 \quad 5{,}3)$$



Média aritmética:

$$\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

$$\bar{X} = \frac{1}{5}(5{,}0 + 5{,}1 + 5{,}0 + 4{,}9 + 5{,}3) \Rightarrow \boxed{\bar{X} = 5{,}06}$$

"tabela de erro"

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	Obs.
5,0	5,0 − 5,06 = −0,06	0,0036	—
5,1	5,1 − 5,06 = +0,04	0,0016	—
5,0	5,0 − 5,06 = −0,06	0,0036	—
4,9	4,9 − 5,06 = −0,16	0,0256	—
5,3	5,3 − 5,06 = +0,24	0,0576	—

Erro Quadratico.

→ Erro Quadratico Médio $= \dfrac{(x_i - \bar{x})^2}{n} = \dfrac{(0{,}0036 + 0{,}0016 + 0{,}0036 + 0{,}0256 + 0{,}0576)}{5}$

$$= 0{,}0184$$

→ Desvio Padrão $= \sqrt{\text{Erro quadratico Médio}} \Rightarrow \sqrt{0{,}0184} = 0{,}1356$

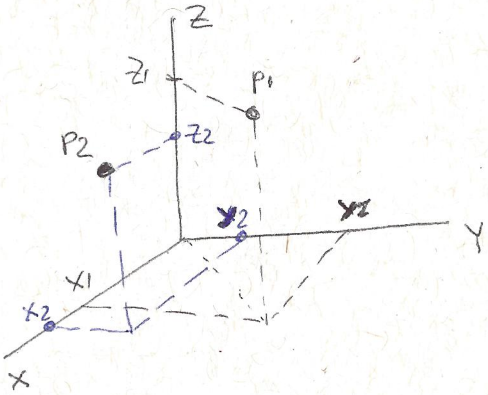
Tem-se: $(5{,}06 \pm 0{,}14)$

R_1

4

(2) OPERADOR GRADIENTE

CAMPO Escalar = CAMPO que pode ser caracterizado em cada ponto por um ÚNICO valor. Ex: TEMPERATURA.



EM $P_1 \Rightarrow t_1 = T(x_1, y_1, z_1)$

EM $P_2 \Rightarrow t_2 = T(x_2, y_2, z_2)$

Problema: como calcular $\Delta T = t_1 - t_2$?

DADO:

$P_1 = (x_1, y_1, z_1)$

$P_2 = (x_2, y_2, z_2)$

$$\Longrightarrow \vec{\Delta r} = \Delta x\,\vec{i} + \Delta y\,\vec{j} + \Delta z\,\vec{k}$$

Para conhecer a variação da grandeza em função da posição basta derivá-la em relação à posição, isto é:

$$\underbrace{\frac{\partial T}{\partial x}}_{\text{variação infinitesimal em } x}\vec{i} + \frac{\partial T}{\partial y}\vec{j} + \frac{\partial T}{\partial z}\vec{k} = \underbrace{\text{grad}}_{\text{Operador}}\,T = \nabla T$$

Obs.: O gradiente de um escalar é um vetor!!!

$$\boxed{\text{grad}\,T = \left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right)}$$

$$\underset{\nabla}{\text{grad}} \equiv \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} + \frac{\partial}{\partial z}\vec{k} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

Se desejamos conhecer como uma grandeza varia no tempo basta derivá-la c/ relação ao tempo. Isto é:

$$\boxed{\vec{v} = \frac{dx}{dt}\vec{i} + \frac{dy}{dt}\vec{j} + \frac{dz}{dt}\vec{k}}$$

também conhecida como velocidade.

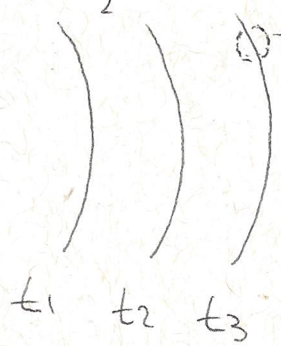
Retomando:

Produto Escalar

$$\Delta t = \nabla T \cdot \Delta R = \frac{\partial T}{\partial x}\cdot \Delta x + \frac{\partial T}{\partial y}\Delta y + \frac{\partial T}{\partial z}\cdot \Delta z =$$

$$\boxed{t_2 - t_1 = \frac{\partial T}{\partial x}(x_2 - x_1) + \frac{\partial T}{\partial y}(y_2 - y_1) + \frac{\partial T}{\partial z}(z_2 - z_1)}$$

Interpretação Geométrica:



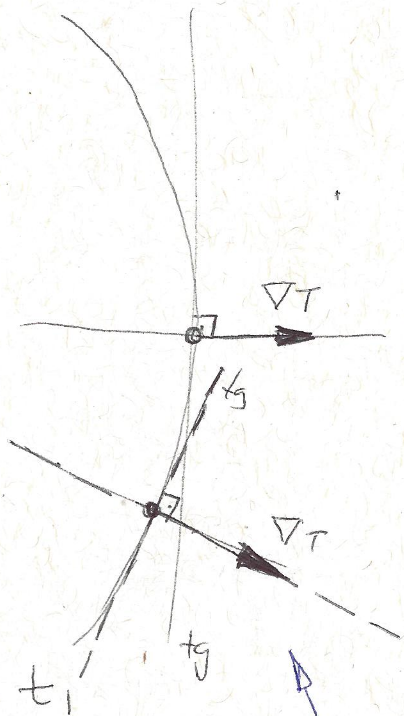
Isoterma
(L.G. = lugar Geométrico que possui a mesma temperatura isto é $\Delta t = 0$)

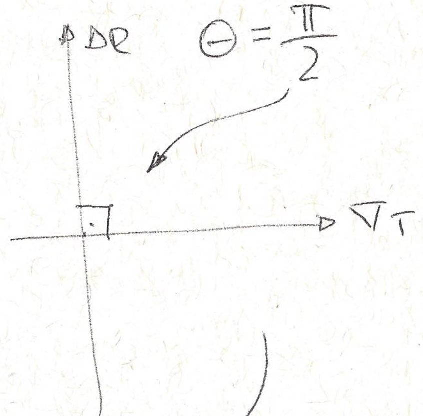
$$\boxed{\Delta t = \nabla T \cdot \Delta R}$$

$$0 = \nabla T \cdot \Delta R$$

$$= |\nabla T| \times |\Delta R| \cdot \cos\theta$$

$\therefore$ $\theta = \frac{\pi}{2}$





reta tangente.

O vetor gradiente indica a <u>direção</u> e o <u>sentido</u> da max variação da Grandeza.

* 6

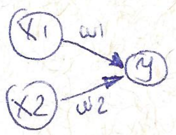
Obs.: O valor max é quando $\theta = 0°$ $\Rightarrow$ caminhar na direção do Gradiente

Regra DELTA p/ uma única saída.

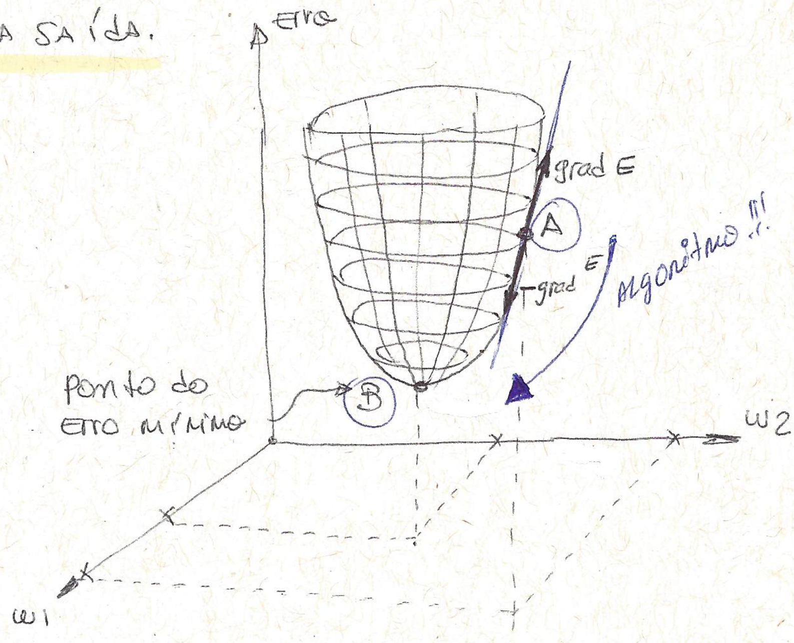
$$E = (t - y\_in)^2 \quad (1)$$

Erro Quadrático

$$y\_in = \sum_{i=1}^{n} w_i \cdot x_i$$



$$E = (t - x_1 \cdot w_1 - x_2 \cdot w_2)^2$$



Sabe-se que o operador gradiente indica a direção e o sentido da maior variação da função, portanto para chegar-se ao pto mínimo deve-se caminhar em sentido contrário, isto é: $-\text{grad } E$.

tem-se de (1):

$$\frac{\partial E}{\partial w_i} = -2\,(t - y\_in) \cdot \frac{\partial y\_in}{\partial w_i} \quad (2)$$

$$\frac{\partial y\_in}{\partial w_i} = \frac{\partial}{\partial w_i}\left(\sum_{i=1}^{n} w_i \cdot x_i\right) \Rightarrow \frac{\partial}{\partial w_1} = \frac{\partial}{\partial w_1}(w_1 \cdot x_1 + w_2 \cdot x_2 + \dots w_n \cdot x_n) = x_1$$

$$\frac{\partial}{\partial w_2} = \frac{\partial}{\partial w_2}(w_1 \cdot x_1 + w_2 \cdot x_2 + \dots) = x_2$$

$$\vdots$$

$$\boxed{\frac{\partial}{\partial w_i} = x_i} \quad (3)$$

Substituindo (3) em (2) e retomando

$$\boxed{\frac{\partial E}{\partial w_i} = -2\,(t - y\text{-}in) \cdot x_i}$$ ← Gradiente de $\vec{wi}$

$$\Delta w_i = -\alpha \cdot \frac{\partial E}{\partial w_i}$$

sentido contrário ao GRADIENTE

$$\boxed{\Delta w_i = \alpha\,(t - y\text{-}in) \cdot x_i}$$

Regra Delta

"O Erro local será reduzido mais rapidamente (p/ um dado α) Ajustando-se o peso com a "regra" Acima"

CH