

# Red Neuronal con Aprendizaje Supervisado

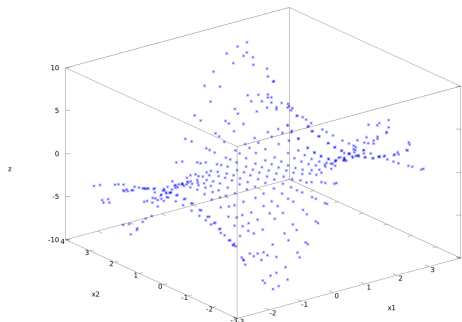
Grupo 5

3 de mayo de 2012

# El problema

## La función a aproximar

- Función escalar  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , de la cual se conocen 441 puntos.
- Intervalo  $[-3, 4]$  en los ejes  $x_1$  y  $x_2$ .  $[-10, 10]$  para el eje  $z$ .



Se utilizaran dos funciones distintas de activación

- **sigmoid-exp:**

$$g(x) = \frac{1}{1 + \exp^{-x}} \quad (1)$$

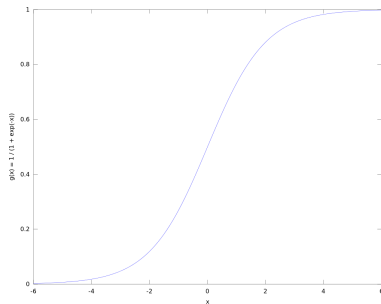
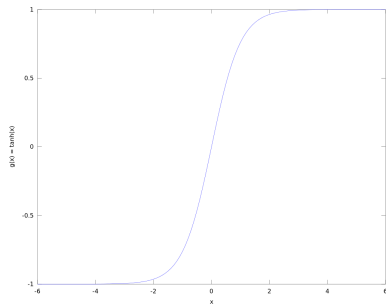
$$g'(x) = g(x)[1 - g(x)] \quad (2)$$

- **sigmoid-tanh:**

$$g(x) = \tanh(x) \quad (3)$$

$$g'(x) = \text{sech}(x)^2 \quad (4)$$

# Funciones de Activación



# Adaptación de la función al problema

## Normalización de la entrada

- sigmoid-exp: Se divide por 0.67 y el rango  $(-4, 4)$  se transforma a  $(-5.97, 5.97)$
- sigmoid-tanh: No hace falta ninguna modificación.

## Normalización de la salida

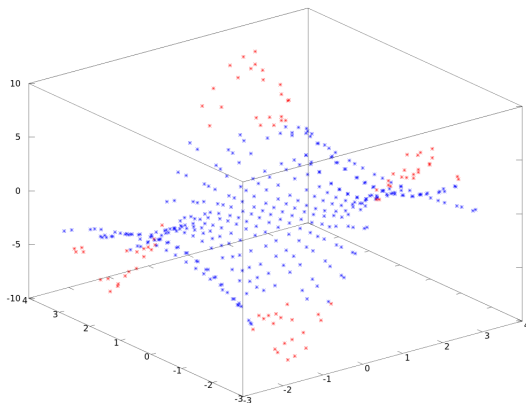
$$n_{exp}(x) = (x - 0,5) * 20 \quad (5)$$

$$n_{tanh}(x) = x * 10 \quad (6)$$

## ¿Cómo se construyeron los conjuntos de testeo y entrenamiento?

- Se dividió a los puntos en dos conjuntos: **puntos de interés** y **puntos de poco interés**.
- Si el valor absoluto del punto es mayor a un **umbral de interés**, el punto es colocado en el conjunto de **puntos de interés**, de lo contrario es colocado en el conjunto de **puntos de poco interés**.
- Se tomó mayor porcentaje de **puntos de interés** sobre **puntos de poco interés** para el conjunto de entrenamiento, y se dejó el resto para el conjunto de prueba.

# Conjuntos de prueba y entrenamiento



## ¿Qué mejoras se implementaron?

- $\eta$  adaptativo.
- *momentum*.



## ¿Cuándo dejamos de entrenar una red?

- Error cuadrático.
- Cantidad de épocas.

## Arquitecturas probadas

- Arquitectura pequeña: [2 6 2 1]
- Arquitectura mediana: [2 12 12 1]
- Arquitectura grande: [2 36 12 6 1]

## Conjuntos de puntos

- Conjunto de entrenamiento: formado por el 95 % de los puntos de interés y el 70 % de los puntos de poco interés, 338 puntos.
- Conjunto de prueba: el resto, 103 puntos.
- Se fijó la cantidad de épocas límite en 500.

## Tabla de resultados

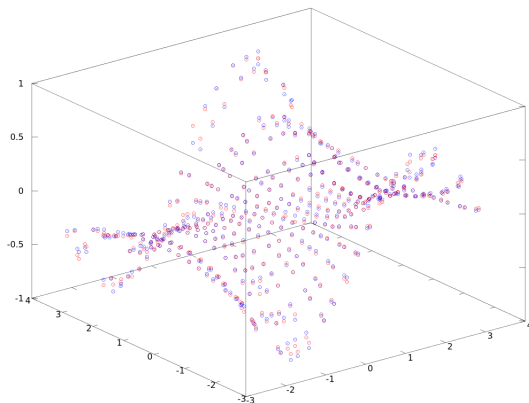
Arquitectura	$\eta$	$a$	$b$	momentum	Error
<b>pequeña</b>	0.5	0.1	0.01	0.2	0.170463
<b>mediana</b>	0.5	0.1	0.01	0.2	0.388060
<b>grande</b>	0.5	0.1	0.01	0.2	0.0867615
<b>pequeña</b>	0.5	0.1	0.01	0.5	0.396113
<b>mediana</b>	0.5	0.1	0.01	0.5	3.32919
<b>grande</b>	0.5	0.1	0.01	0.5	0.104419
<b>pequeña</b>	0.5	0.025	0.0025	0.7	13.1008
<b>mediana</b>	0.5	0.025	0.0025	0.7	110.037
<b>grande</b>	0.5	0.025	0.0025	0.7	45.6437

Se realizó un entrenamiento sobre la *arquitectura grande* con los siguientes parámetros:

- $a = 0.1$ ,  $b = 0.01$
- *factor momentum* = 0.2
- cantidad máxima de épocas = 5000
- error cuadrático medio mínimo = 0.017

Se obtuvo un error cuadrático medio de 0.0715993, cortando por la cantidad máxima de épocas.

# Resultados



# Resultados

