**Proyecto final de Álgebra Lineal**

**Daniel Charua A01017419**

**Luis Barajas A01370934**

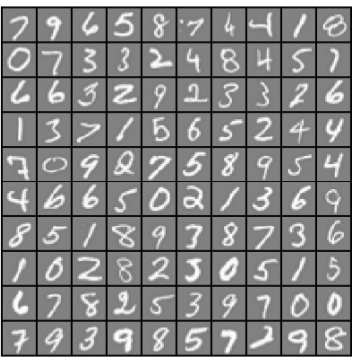
Operación de matrices y vectores en la implementación de una red neuronal para reconocer dígitos

Introducción

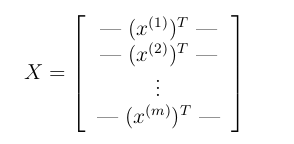
En el siguiente documento explicaremos la implementación de una red neuronal que se entrena para reconocer dígitos mediante un serie de operaciones entre matrices y vectores.

Desarrollo

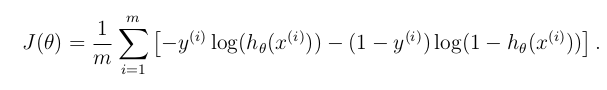
Para entrenar a la red neuronal, se usaron 5000 ejemplos de entrenamiento, donde cada elemento es una imagen de 20 por 20 pixeles en escala de grises.



Cada píxel es representado por un número flotante que indica la intensidad en la escala de grises de esa ubicación. Estos dígitos se transaldan de una matriz 20n x 20M a una fila, generando un vector (x) de 400 unidades, los cuales se agrupan en una sola matriz(X) de 5000 por 400 donde cada fila es un ejemplo de la imagen de un dígito escrito a mano.



Para esta red neuronal se utilizó regresión logística para lograr obtener la función que aproxima los 10 dígitos posibles.



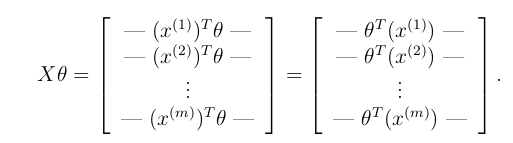
Donde la hipótesis para cada vector xi denotada como hθ(x(i)) es:



y g(z) es la función sigmoid:

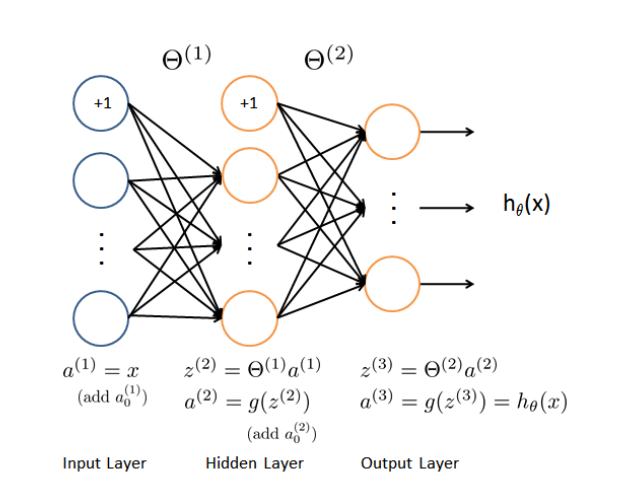


Se acelera y simplifica mucho la computación si se utiliza vectorizacion en lugar de un for loop, de manera que cada elemento del vector multiplicado por su theta (la función de aproximación) se puede escribir como:



De esta manera la red neuronal logra encontrar la función de aproximación que mejor le permita dividir cada imagen procesada en 10 categorías, donde cada categoría es un dígito, se calculan todas las posibilidades (digitos 0-9) para cada imagen y se selecciona la que resultó con mayor probabilidad, es decir, la que mejor acomoda a ese dígito en las opciones posibles (0-9).

La red neuronal tiene la siguiente estructura:



En la imagen anterior cada círculo azul representa un vector con los pixeles de un dígito de entrenamiento, la segunda capa es la que busca una aproximación o hipótesis y la tercera capa es la que deduce la posibilidad para cada uno de los dígitos posibles (0-9), de esta manera se escoge la posibilidad más alta para dar la respuesta

**Referencias**

* **Coursera. (2011). *Machine Learning: Stanford University.* Obtenido de:** [**https://www.coursera.org/learn/machine-learning/home/week/4**](https://www.coursera.org/learn/machine-learning/home/week/4)