Perceptrón multicapa

Aproximador universal

Verónica E. Arriola-Rios

Facultad de Ciencias, UNAM

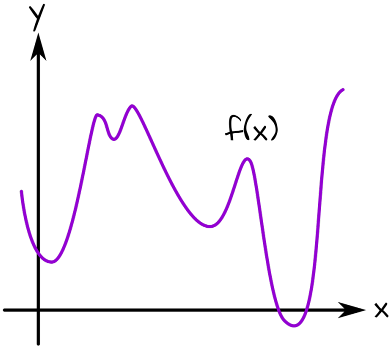
13 de febrero de 2023

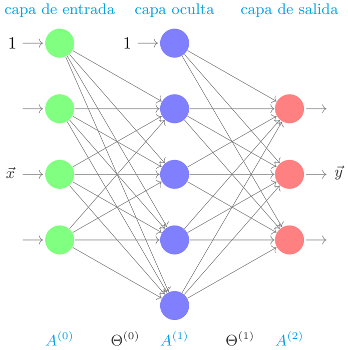
Teorema

Aproximador universal

Definición

Dada una función continua cualquiera $f(\vec{x})$, existe una red neuronal de tres capas tal que para cada entrada x la salida es $\hat{f}(\vec{x})$ con $|\hat{f}(\vec{x}) - f(x)| < \varepsilon$





Funciones $\mathbb{R}^{s_0} \to \mathbb{R}^{s_L}$

- El número de componentes s_0 en el vector de **entrada** $\vec{x}$ es la dimensión de los argumentos de la función.
- El número de componentes s_L en el vector de **salida** $\vec{y}$ es la dimensión de la salida de la función.

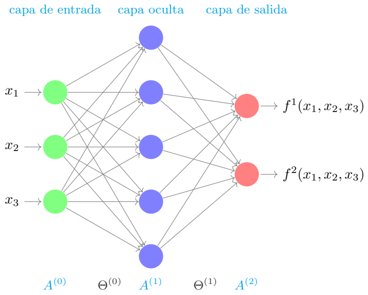


Figura: Ejemplo con $s_0 = 3$ y $s_L = 2$, sin sesgos

Observaciones

- Para incrementar la precisión de la aproximación se requiere incrementar el número de neuronas s_1 de la capa oculta.

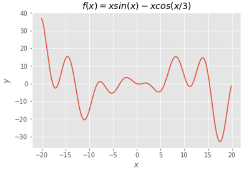


Figura: Función analítica

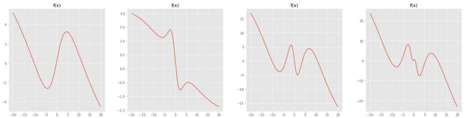
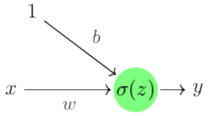


Figura: Aproximaciones en entrenamiento para 2, 5, 50 y 500 neuronas en la capa oculta.

Elementos

1 perceptrón

- Observemos el comportamiento de un solo perceptrón con las funciones:



$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$z = wx + b$$

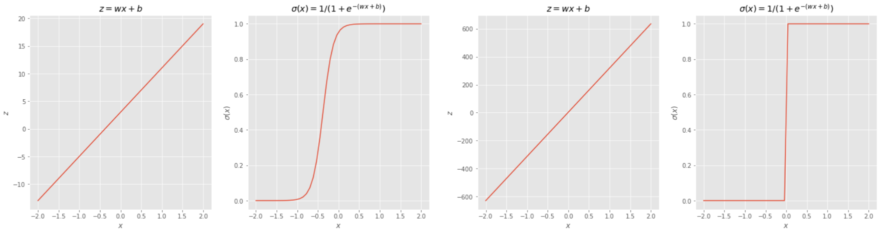


Figura: **Izquierda:** $w = 8, b = 3$ **Derecha:** $w = 316, b = 3$

- El valor del sesgo b indica la posición de la frontera entre $\mathrm{f}(\mathrm{x}) \sim 0$ y $\mathrm{f}(\mathrm{x}) \sim 1$.
- Conforme w se vuelve muy grande (> 500), la frontera tiende a ser vertical, para valores más pequeños la curva se vuelve más ancha.
- *La posición de la frontera es* ***directamente*** *proporcional a* b *e* ***inversamente*** *propocional a* w.Nielsen 2019

$$s = -\frac{\mathrm{b}}{w}$$

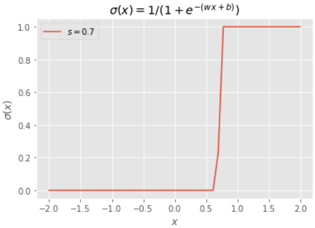


Figura: Frontera fijada con $w = 200, s = 0.7$

2 perceptrones

- Utilizando aproximaciones a funciones escalón en una red con dos perceptrones.

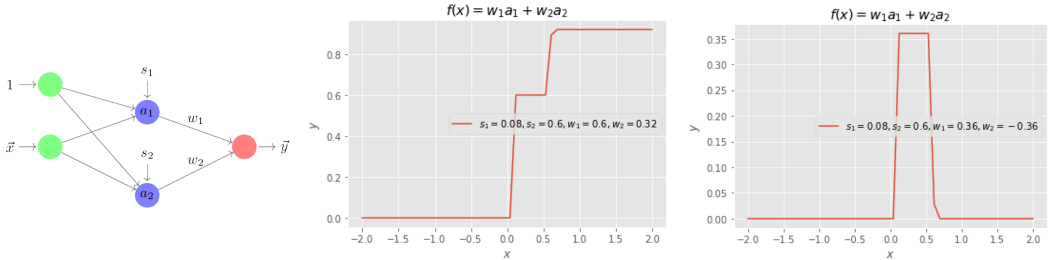


Figura: **Centro:** Los escalones de ambos perceptrones se suman. **Derecha:** Si las contribuciones de ambos son iguales y opuestas $w_1 = -w_2$, obtenemos un bloque cuyo ancho es $s_2 - s_1$ y su altura es $\mathrm{h} = w$.

Aproximando la función

- Se pueden utilizar entonces pares de perceptrones, diseñados intencionalmente para aproximar cualquier función utilizando secuencias de barras hasta la precisión deseada, dada por el ancho de las barras.
- Según la figura siguiente ¿cuántos pares más y en qué posiciones podrías ponerlos para aproximar mejor la función?

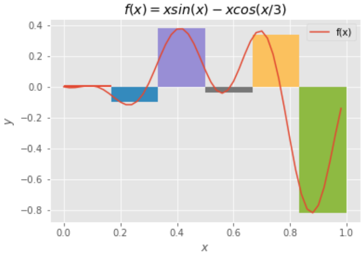


Figura: 6 pares de neuronas equidistantes.

- Un argumento semejante se puede utilizar para construir aproximaciones mediante bloques en más dimensiones.

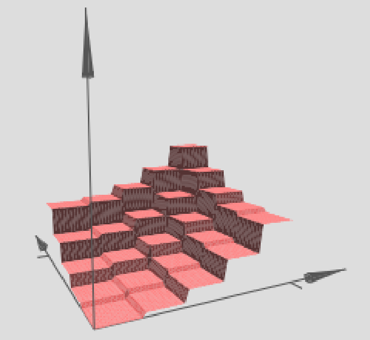


Figura: Construcción para funciones $\mathbb{R}^2 \to \mathbb{R}$. Tomado de Nielsen 2019

Referencias I

Nielsen, Michael (dic. de 2019). *A visual proof that neural nets can compute any function*. English. URL: http://neuralnetworksanddeeplearning.com/chap4.html.

Licencia