

IMPLEMENTACIÓN NUMÉRICA DE ESPERANZAS CONDICIONALES USANDO REDES NEURONALES

Melanie Sánchez Pfeiffer

19 de diciembre de 2023

Contexto

Dadas X, Y variables aleatorias reales en $(\Omega, \mathcal{F}, \mathbb{P})$, se sabe que la esperanza condicional $\mathbb{E}(Y|X)$ es la proyección ortogonal de Y en $(\Omega, \sigma(X), \mathbb{P})$, por lo tanto, es la función medible de X más cercana a Y en L^2 . Es decir, $\hat{f} = \mathbb{E}(Y|X)$ es la única solución del problema

$$\min_{f \in K} \mathbb{E}[f(X) - Y]^2$$

donde K es el conjunto de todas las funciones medibles $f : \mathbb{R} \rightarrow \mathbb{R}$ tal que $f(X) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Sea $\mu = \text{Ley}(X)$ y definimos el conjunto $F_{N,\varphi}$ como todas las redes neuronales $f_\theta : \mathbb{R} \rightarrow \mathbb{R}$ formadas por 2 capas, N neuronas, con función de activación φ y parametrizadas por $\theta \in \mathbb{R}^{D(N)}$. Por el Teorema de Hornik sobre la universalidad de redes neuronales en $L^2(\mu)$ tenemos que $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ y $\theta \in \mathbb{R}^{D(N)}$ tal que:

$$\int |f_\theta(x) - \hat{f}(x)|^2 \mu(dx) < \epsilon$$

o análogamente:

$$\mathbb{E}([f_\theta(x) - \mathbb{E}(Y|X = x)]^2) < \epsilon.$$

Metodología

Observamos que:

$$\mathbb{E}([f_\theta(x) - \hat{f}(x)]^2) = \mathbb{E}([f_\theta(X) - Y]^2) + \mathbb{E}(\hat{f}(x)^2) - \mathbb{E}(Y^2)$$

de donde se obtiene

$$\arg \min_{\theta} \mathbb{E}([f_\theta(x) - \hat{f}(x)]^2) = \arg \min_{\theta} \mathbb{E}([f_\theta(X) - Y]^2)$$

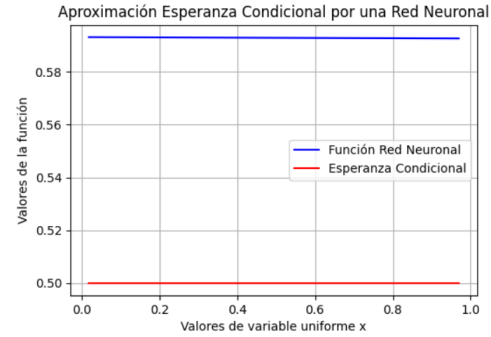
Por lo que se propone entrenar una red neuronal para que encuentre la función f_θ , es decir, una aproximación de la función $\hat{f}(x) = \mathbb{E}(Y|X = x)$, buscando

$$\hat{\theta} = \arg \min_{\theta} \mathbb{E}([f_\theta(X) - Y]^2).$$

Para esto y para un par de variables aleatorias (X, Y) dadas se implementará S.G.D, en base a una muestra $(X_i, Y_i) \text{ i.i.d} =^{\text{ley}} (X, Y)$ con f_θ una red neuronal de 2 capas y N neuronas.

Resultados

Para (X, Y) variables uniformes en $[0, 1]$ obtenemos:



Referencias

- 1 Kurt Hornik, Maxwell Stinchcombe, and Halber White. MultilayerFeedforward Networks are Universal Approximators. Neural Networks, Vol2, pp.359-366,1989.
- 2 Phillipp Grohs, and Gitta Kutyniok. Mathematical aspects of deep learning. Cambridge University Press, 2023.
- 3 Apuntes Curso MA5606-1: Tópicos Matemáticos en Aprendizaje de Máquinas, Redes Neuronales y Aprendizaje Profundo. Profesores Joaquín Fontbona y Claudio Muñoz, 2023.