

Homework 1 – Question 3

3.1

Quer se mostrar que,

$$\text{relu}(z) := \arg \min_{y \geq 0} \|y - z\|^2, \quad (1)$$

R:

O quadrado da norma é expandido na forma,

$$\|y - z\|^2 = \sum_{i=1}^K (y_i - z_i)^2.$$

Como a restrição $y \geq 0$ aplica-se independentemente a cada caso, sendo dado que a expressão é dada por um somatório, é suficiente mostrar para um dos casos i , onde $\min_{y_i \geq 0} (y_i - z_i)^2$.

Considerando,

$$\min_{y \geq 0} (y - z)^2.$$

Calcula-se primeiramente o gradiente,

$$\frac{d}{dy} (y - z)^2 = 0 \Leftrightarrow 2(y - z) = 0 \Leftrightarrow y = z.$$

Aplicando a restrição:

- Se $z \geq 0$, então $y = z$ satisfaz a restrição $y \geq 0$.
- Se $z < 0$, então o único valor que não viola a condição $y \geq 0$, é $y^* = 0$.

Como tal, sendo y^* o valor resultado da proposição inicial,

$$y^* = \begin{cases} z, & \text{if } z \geq 0, \\ 0, & \text{if } z < 0, \end{cases} = \text{relu}(z).$$

Assim, provou-se que,

$$\text{relu}(z) := \arg \min_{y \geq 0} \|y - z\|^2, \quad (2)$$

Podendo-se também concluir que *Sparsemax* é um *Relu* normalizado, sendo que adiciona a restrição $\sum_i y_i = 1$ à expressão a cima.