P1001. El de ML

Anal Turing

Es bien conocido que los tiburaçaos, en particular los mates-datos, tienen que inventarse nombres fancies para poder colocar sus estafas al tonto de turno. Tenemos varios ejemplos de sus fechorías: *salario emocional* para no decirte que echarás más horas que un reloj, *somos una start-up* para que trabajes en Nvidia gratis o *grado universitario* para hablar de la fp que hacen además de mates. Esta vez han llamado

$$\{f|f(x) = a_0 + \sum_{i=1}^{n} a_i \max(b_i x + c_i, 0)\}$$

al conjunto de funciones lineales a trozos , qué retorcidos son estos *data scientits*.
Veamos que realmente es así:

- $\subseteq$) Sea $f \in B$, para cada $i = 1, 2, ..., n$, sea x_i el punto de $[a, b]$, si existe, tal que la función $\max(b_i x + c_i, 0)$ cambia en ese punto, es decir, $b_i x + c_i < 0$ si $x < x_i$ y $b_i x + c_i > 0$ si $x > x_i$ o al revés. $b_i x + c_i$ es una función lineal, por lo que si este punto existe, es único. Introducimos ahora una nueva ordenación de manera que, si en total hay k x_i's, $x_1 \leq x_2 \leq ... \leq x_k$. Entonces, para cada $j = 1, ..., k-1$, $f(x) = a_0 + \sum_{i=1}^{n} a_i \max(b_i x + c_i, 0)$ con $\max(b_i x + c_i, 0) = b_i x + c_i$ o $\max(b_i x + c_i, 0) = 0$ para cada i y cada x en $[x_j, x_{j+1}]$, por lo que f es lineal a trozos.
- $\supseteq$) Sea f lineal a trozos. Sea $\{x_1 = a < x_2 < ... < x_n = b\}$ una partición de $[a, b]$ tal que f es lineal en $[x_i, x_{i+1}]$ para todo i. Sea m_i la pendiente de f en $[x_i, x_{i+1}]$. Definimos $b_1 = m_1$, $b_i = |m_i - \sum_{j=1}^{i-1} b_j| > 0$, para $i = 2, ..., n-1$, $c_i = -b_i x_i$, $a_0 = f(a)$, y $a_i = \text{sign}(m_i - \sum_{j=1}^{i-1} b_j)$. Con estas definiciones,
  $\max(b_i x + c_i, 0) = \begin{cases} b_i x + c_i & \text{si } x > x_i \\ 0 & \text{si } x \leq x_i \end{cases}$ por lo que si $g(x) = a_0 + \sum_{i=1}^{n} a_i \max(b_i x + c_i, 0)$, y $x \in [x_i, x_i{+}1]$, $g(x) = a_0 + \sum_{j=1}^{n} a_j \max(b_j x + c_j, 0) = a_0 + \sum_{j=1}^{i} a_j (b_j x + c_j) = m_i x + a_0 + \sum_{j=1}^{i} a_j c_j$. Es una función lineal con la misma pendiente que f en ese punto. Ahora, en cada tramo la pendiente de g es la misma que la de f, por lo que si los puntos iniciales de cada tramo coinciden, lo harán en todo el tramo. Tenemos que $g(a) = f(a) \implies g(x_1) = f(x_1) \implies ... \implies g(x_n) = f(x_n)$, luego $f = g$.

Una vez hemos desenmascarado a los estadísticos con traje[1], usamos matemáticas de verdad para resolver el problema.

Ahora solo queda demostrar que este conjunto es denso en $\mathcal{C}([a, b], \mathbb{R})$. Para ello usaremos el teorema de Stone, que nos dice que si B es retículo, es decir, cerrado respecto del máximo y el mínimo de dos funciones de B, y se cumple que $\forall x, y \in [a, b], \alpha, \beta \in \mathbb{R}, \exists f \in B$ tq $f(x) = \alpha, f(y) = \beta$, entonces B es denso.

La segunda propiedad es trivial de comprobar, ya que las funciones lineales pertenecen a B. Para la primera, observamos que dadas dos funciones lineales f, g definidas en $[c, d]$ tales que $\exists x_0 \in [c, d]$ con $f(x) - g(x) < 0$ en $[c, x_0]$ y $f(x) - g(x) > 0$ en $[x_0, d]$, entonces tanto $\max(f, g)$ como $\min(f, g)$ son lineales a trozos en $[c, d]$. Ahora, si $f, g \in B$ y $\{x_0 = a < x_1 < ... < x_n = b\}$, $\{y_0 = a < y_1 < ... < y_m = b\}$ son particiones de $[a, b]$ tales que f es lineal en $[x_i, x_{i+1}]$ para todo i y g es lineal en $[y_j, y_{j+1}]$ para todo j, si nos restringimos a cada uno de los tramos de la partición definida por la unión de ambas, las dos funciones son lineales, por lo que la función máximo y la función mínimo serán lineales a trozos en cada uno de estos tramos, y por tanto serán lineales a trozos en B. Por tanto, B es retículo y por el teorema de Stone, denso en $\mathcal{C}([a, b], \mathbb{R})$.

[1]No estamos 100% seguros de que la demostración esté bien, como no lo esté el ridículo es histórico.