Datos y variables

$y_i \in \{0,1\} \quad \forall i = 1\ldots,n$: variables de respuesta binarias. Usualmente representadas por el vector $\mathbf{y} = (y_1,\ldots,y_n)^t$

$\mathbf{x}_i \in \mathcal{X}^d \subseteq \mathbb{R}^d \quad \forall\ i = 1\ldots,n$: covariables o regresores. Si se usa por si sola x o $\mathbf{x}$ (vector), esta representa una variable arbitraria. Si se habla de toda la matriz de datos, se denota por $\mathbf{X} \in \mathcal{X}^{n\times d} \subseteq \mathbb{R}^{n\times d}$. Juntos con las y_i, se tienen los datos para el modelo: $\{(y_i,\mathbf{x}_i)\}_{i=1}^n$

$n \in \mathbb{N}$: número de observaciones en la muestra.

$d \in \mathbb{N}$: número de covariables, dimensionalidad de los regresores.

$\mathcal{X}^d$: subconjunto de $\mathbb{R}^d$, espacio de covariables. Formado por el producto punto de los rangos de cada variable: $\mathcal{X}^d = [a_1,b_1]\times[a_2,b_2]\times\ldots\times[a_d,b_d]$, donde $[a_j,b_j]\subset\mathbb{R}$ es un intervalo cerrado estándar en los reales.

Especificos del modelo

$z_i \sim N(\cdot)\forall i = 1\ldots,n$: variables latentes del modelo cuya distribución es normal . Usualmente se acomodan en un vector $\mathbf{z} = (z_1,\ldots,z_n)^t$

$f(\mathbf{x})$: función de proyección.

$f_j(x_j) \quad \forall j = 1,\ldots,d$: polinomio por partes anidada en la función de proyección. Sirve para hacer una transformación no lineal de la dimensión J. En ocasiones se acomodan en su forma vectorial $\mathbf{f}(\mathbf{x})$. Ver ecuación (**??**)

$\beta = (\beta_0,\beta_1,\ldots,\beta_j)^t$: vector de coeficientes para la regresión lineal.

$w_j = (w_{j,1},\ldots,w_{j,N^*})^t \quad \forall j = 1,\ldots d$: vector de pesos para las funciones base de la transformación lineal en j. Si se habla de todos los pesos en conjunto, esos se acomodan en una matriz $\mathbf{w} = [w_1,\ldots w_d]^t \in \mathbb{R}^{N^*\times d}$

$\Psi_j(\cdot) = (\Psi_{j,1}(\cdot),\ldots,\Psi_{j,N^*}(\cdot))^t$: vector de funciones base para los polinomios por parte flexibles. Ver ecuación (**??**) y (**??**)

N^*: número total de funciones base. Ver ecuación (**??**) para su expansión final.

M: número de bases para los polinomios por partes. M-1 indica el grado de los polinomios.

J: número de sub-intervalos en los que se parte cada $[a_j, b_j]$.

K: número de restricciones de continuidad impuestas.

$\mathcal{P}_j = \{\tau_1, \ldots, \tau_{J-1}\} \quad \forall j = 1, \ldots, d$: partición del espacio de la dimensión j.

τ: nodos, se omiten los índices para evitar confusión, pero se tienen un total de $d * (J-1)$ nodos acomodados en una matriz de igual tamaño.

Contadores e índices

i: contador, usado para denotar un conjunto de observaciones *hacia abajo*, i.e. $i = 1, \ldots, n$. En la sección **??** se usa para contar sobre el grado del polinomio i.e. $i = 1, \ldots, M$. (Ver ecuación **??**)

j: contador, usado para denotar el conjunto de variables *a lo largo*, i.e. $j = 1, \ldots, d$. Usualmente se hace referencia a la dimensión arbitraria j. En la sección**??** se usa para contar sobre los nodos entre intervalos i.e. $j = 1, \ldots, J-1$. (Ver ecuación **??**)

k: contador, usado para denotar el número de iteración en el algoritmo, i.e. $k = 1, 2, 3, \ldots$

l: contador adicional, asociado al número de funciones base N^* constante para cada dimensión j.

Probabilidad

$F(\cdot)$: Distribución arbitraria de la familia exponencial.

$N(\cdot|\mu, \sigma^2)$: distribución normal con su correspondiente parametrización de media y varianza. Se utiliza la misma notación para su forma vectorial añadiendo un subíndice indicando su dimensionalidad: $N_{\cdot}(\cdot|z, \sigma)$ con su correspondiente vector de medias μ y vector de varianza covariazna Σ.

$\Phi(\cdot) : \mathbb{R} \to (0,1)$: la función de distribución acumulada de una distribución normal estándar $N(\cdot|1, 0)$, con su correspondiente inversa Φ^{-1}.

$\text{Be}(\cdot|p)$: distribución bernoulli con probabilidad de éxito p.

$p \in [0, 1]$: probabilidad arbitraria.

$g(\cdot)$: función liga. Ver diagrama **??**

ϵ: errores aleatorios, usualmente distribuidos $N(\epsilon|\mu, \sigma^2)$.

$P(\cdot)$, $\mathbb{E}[\cdot]$, $\mathbb{V}[\cdot]$: medida de probabilidad, operadores de esperanza y varianza respectivamente.

$\theta \in \Theta$ parámetros canónicos de distribuciones exponenciales, con Θ su correspondiente espacio.

$\pi(\cdot)$: función de densidad.

α: símbolo de proporcionalidad.

$S(\cdot|\cdot)$: función de suavizamiento.

ρ: correlación.

Algoritmo

N_{sim}: número de simulaciones realizadas en el algoritmo.

k^*: número de observaciones por descartar, periodo de *burn-in*.

k_{thin}: parámetro de adelgazamiento.

r_{j^*}: residuales parciales para alguna j^* en particular.

Otros

$h(\cdot)$: función arbitraria.

$h^{(k)}$: (k)-ésima derivada de h.

$s : \mathbb{R} \to (0, 1)$: familia de funciones sigmoidales.

I: función indicadora.

$(\cdot)_+$: función parte positiva.

$\mathbf{1}$: vector de números uno.

ll: función *log-loss*.

El símbolo $\hat{\cdot}$ se usa para indicar que se trata de una variable estimada, i.e. $\hat{y}$ es la estimación de las variables correspondientes y.