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

$\mathcal{X}^d \subseteq \mathbb{R}^d$: 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 en los reales

Específicos del modelo

$z_i \sim N(\cdot) \quad \forall i = 1\ldots,n$: variables latentes del modelo cuya distribución es normal. En su forma vectorial: $\mathbf{z} = (z_1,\ldots,z_n)^t$

$\eta(\mathbf{x})$: función aditiva de predicción

$f_j(x_j) \quad \forall j = 1,\ldots,d$: polinomios por partes

$\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_l)^t$: vector de parámetros por estimar. Si se le añade tilde entonces el contador comienza en uno y el vector no contiene el parámetro independiente, es decir: $\tilde{\boldsymbol{\beta}} = (\beta_1, \ldots, \beta_l)^t$

$\Psi_l(\cdot) \quad \forall l = 1, \ldots, N^*$: funciones bases para la expansión en polinomios por partes de f_j. Ver (**??**) y (**??**)

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

M: tamaño de la base para los polinomios por partes, por lo tanto, 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 tienen un total de $d(J-1)$ nodos acomodados en una matriz de igual tamaño.

Contadores e índices

$i = 1, \ldots, n$: contador, usado para denotar un conjunto de observaciones

$j = 1, \ldots, d$: contador, usado para denotar el conjunto de covariables. Usualmente se hace referencia a la dimensión arbitraria j

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

$l = 1, \dots, N^*$: contador asociado al número de funciones base total en la expansión de polinomios por partes N^*.

$\hat{\imath} = 1, \dots, M - 1$: contador asociado al número de funciones base para cada subintervalo, M, en las expansione de polinomios truncados.

$\hat{\jmath} = 1, \dots, J - 1$ contador asociado al número de funciones base para cada subintervalo (parámetro M) en las expansione de polinomios truncados

Probabilidad

$F(\cdot)$: Función de distribución arbitraria de la familia exponencial

$\mathcal{N}(\,\cdot\,|\mu, \sigma^2)$: distribución normal con su correspondiente parametrización de media y varianza. Se utiliza la misma notación para su forma vectorial añadiendo un subíndice para denotar su dimensionalidad: $\mathcal{N}_{\cdot}(\cdot|\boldsymbol{\mu}, \sigma)$ 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 $\mathcal{N}(\,\cdot\,|1, 0)$, 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

$\propto$: operador de proporcionalidad

ρ: correlación lineal de Perarson

Algoritmo

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

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

k_{thin}: parámetro de adelgazamiento.

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

Abreviaciones

GAM : *Generalized aditive model*, modelo aditivo generalizado.

GLM : *Generalized linear model*, modelo lineal generalizado

MCMC : *Markov Chain Monte Carlo*, cadena de Markov Montecarlo

ML : *Machine Learning*, aprendizaje de máquina.

RSS : *Residual sum of squares*, suma de residuales cuadrados