<u>**Lección n°1**</u>**:** Análisis de sensibilidad global : índices de Sobol y DGSM AMARUN 2024

1 Introducción

El análisis de sensibilidad global (GSA, por sus siglas en inglés Global Sensitivity Analysis) es un campo de la estadística que ha experimentado un notable crecimiento en los últimos tiempos. Su objetivo es cuantificar la influencia de variables de entrada, que se asumen aleatorias, en la salida de una función multivariada $f : \mathbb{R}^d \to \mathbb{R}$, cuya evaluación es costosa. Esta puede representar códigos de cálculo que modelan fenómenos complejos o algoritmos de inteligencia artificial cuyo funcionamiento no se comprende bien.

En [1] se estableció una conexión significativa entre el campo del análisis de sensibilidad global y el de las desigualdades funcionales, un área matemática previamente no relacionada. En este trabajo, se demuestra que la conexión entre dos índices de sensibilidad proviene directamente de una desigualdad de Poincaré. Esto lleva a reconsiderar a las desigualdades de Poincaré desde la perspectiva de la explicabilidad de modelos.

Este curso tiene como objetivo dar una muy breve introducción al análisis de sensibilidad global, utilizando las desigualdades de Poincaré unidimensionales para medidas de probabilidad como herramienta. Todo el material está cubierto en [2] y para esta lección también hemos consultado [3].

El curso requiere que se conozcan los conceptos básicos de la teoría de probabilidades (variables aléatorias y esperanza condicional), los cuales pueden ser consultados a detalle aquí 🔗.

Denotaremos la esperanza de una variable aleatoria real Y por $\mathbb{E}[Y]$ y su varianza por $\mathrm{Var}(Y) = \mathbb{E}\left[(Y-\mathbb{E}[Y])^2\right]$. Dado un vector aleatorio $X = (X_1, \ldots, X_d)$ en $\mathbb{R}^d$ y un conjunto $I \subseteq \{1, \ldots, d\}$, denotamos por X_I al vector formado por las variables X_j, con $j \in I$, y adicionalmente denotamos $X_{-i} := X_{\{1,\ldots,d\}\setminus\{i\}}$.

2 Descomposición de Hoeffding-Sobol e índices de Sobol

Los iń‌dices de Sobol son índices de sensibilidad que permiten medir la influencia de variables aleatorias dentro de un modelo. Los especialistas a menudo los prefieren debido a su fácil interpretabilidad. Estos índices se definen a partir de la descomposición ortogonal de Hoeffding-Sobol (también conocida como descomposición ANOVA, por ANalysis Of Variance), descrita en el teorema que se presenta a continuación.

Teorema 2.1 (Descomposición ortogonal de Hoeffding-Sobol) *Sea $X = (X_1, \ldots, X_d)$ un vector de variables aleatorias independientes y $f\colon \mathbb{R}^d \longrightarrow \mathbb{R}$ una función que satisface $\mathbb{E}\left[|f(X)|^2\right] < \infty$. Entonces existe una única descomposición de $f(X)$, escrita como*

$$f(X) = \sum_{I\subset\{1,\ldots,d\}} f_I(X_I) = \underbrace{\mathbb{E}[f(X)]}_{f_\emptyset} + \sum_{i=1}^{d} f_i(X_i) + \sum_{i<j}^{d} f_{i,j}(X_i, X_j) + \sum_{i<j<k}^{d} f_{i,j,k}(X_i, X_j, X_k) + \ldots,$$

que satisface

$$\mathbb{E}[f_I(X_I) \mid X_J] = 0, \quad \textit{para cualquier } J \subsetneq I \tag{1}$$

Por convención denotamos $\mathbb{E}[\,\cdot\,|X_\emptyset] = \mathbb{E}[\,\cdot\,]$.

Ejercicio 2.1 *Muestre la formula recursiva*

$$f_I(X_I) = \mathbb{E}\left[f(X)|X_I\right] - \sum_{J\subsetneq I} f_J(X_J),$$

que caracteriza de manera única la descomposición de Hoeffding-Sobol.

La descomposición se dice ortogonal pues la condición (1) implica $\mathbb{E}\left[f_I(X_I)f_J(X_J)\right] = 0$, para todo $I \neq J$ (**Ejercicio 2.2** Pruébelo). En particular, tomando $J = \emptyset$, tenemos que cada $f_I(X_I)$ $(I \neq \emptyset)$ es una variable aleatoria centrada (i.e. $\mathbb{E}[f(X_I)] = 0$) y que podemos descomponer la varianza total de $f(X)$ como una suma de varianzas:

$$\operatorname{Var}\left(f(X)\right) = \sum_{I,J \subseteq \{1,\dots,d\}} \operatorname{Cov}\left(f_I(X_I), f_J(X_J)\right) = \sum_{\substack{I \subseteq \{1,\dots,d\} \\ I \neq \emptyset}} \operatorname{Var}\left(f_I(X_I)\right).$$

Tenemos entonces que cada $f_I(X_I)$ contiene la información de X_I dentro del modelo $f(X)$, en términos de varianza. En particular para cada $i \in \{1,\dots,d\}$, el *efecto principal* $f_i(X_i)$ explica la influencia de X_i sin tomar en cuenta su interacción con el resto de variables, mientras que el *efecto total* $f_i^{\text{tot}}(X) = \sum_{I \ni i} f_I(X_I)$ representa su influencia total. Los índices de Sobol surgen naturalmente como proporciones de varianza.

Definición 1 (Indices de Sobol) *Sea $X = (X_1,\dots,X_d)$ un vector de variables aleatorias independientes y $f\colon \mathbb{R}^d \longrightarrow \mathbb{R}$ una función que satisface $\mathbb{E}\left[|f(X)|^2\right] < \infty$. El* índice de Sobol de primer orden *y el* índice de Sobol total *de $i \in \{1,\dots,d\}$ se definen como*

$$S_i = \frac{\operatorname{Var}(f_i(X_i))}{\operatorname{Var}(f(X))} \in [0,1], \qquad S_i^{\text{tot}} = \frac{\operatorname{Var}(f_i^{\text{tot}}(X))}{\operatorname{Var}(f(X))} = \frac{\sum_{I \ni i} \operatorname{Var}(f_I(X_I))}{\operatorname{Var}(f(X))} \in [0,1],$$

respectivamente.

Los índices S_i^{tot} tienen más relevancia que los índices S_i, pues se interpretan como porcentajes de influencia total de cada variable X_i. En particular, los índices totales pueden ser empleados con fines de "screening", identificando variables con una influencia mínima en el modelo si nuestros datos indican que S_i^{tot} está por debajo de cierto umbral, por ejemplo, del 5%.

Ejercicio 2.3 *1. Pruebe que los efectos totales también se pueden escribir como*

$$f_i^{\text{tot}}(X) = f(X) - \mathbb{E}\left[f(X) \,|\, X_{-i}\right]$$

para todo i $in\{1,\dots,d\}$.

2. Muestre que

$$S_i^{\text{tot}} = 1 - \frac{\operatorname{Var}(\mathbb{E}\left[f(X) \,|\, X_{-i}\right])}{\operatorname{Var}(f(X))}.$$

Ejercicio 2.4 *Sean $X_1 \sim \mathcal{U}([-1,1])$ y $X_2 \sim \mathcal{N}(0,1)$, dos variables aleatorias independientes, y considere el modelo $f(X_1,X_2) = X_1^4 + \cos(2X_2)$.*

1. *Pruebe que $\mathbb{E}\left[\cos(tX_2)\right] = e^{-\frac{t^2}{2}}$ para todo $t \in \mathbb{R}$.*
2. *Determine la descomposición de Hoeffding-Sobol de $f(X)$.*
3. *Calcule los índices de Sobol de primer orden.*
4. *¿Qué variable presenta mayor influencia en el modelo?*

Ejercicio 2.5 (Función de Sobol) *Sea $X = (X_1,\dots,X_d)$ un vector de variables aleatorias independientes uniformemente distribuidas en $[0,1]$ y sean $a_1,\dots,a_d \in \mathbb{R}$. Definimos*

$$f(X) = \prod_{k=1}^{d} \frac{|4X_k - 2| + a_k}{1 + a_k}.$$

Calcule S_i^{tot} para $i \in \{1,\dots,d\}$. ¿Qué sucede con S_i^{tot} cuando a_i aumenta?

2.1 Estimación de los índices de Sobol totales

Presentamos un posible estimador del índice de Sobol total S_i^{tot}. Este se basa en la siguiente proposición.

Proposición 2.1 *Sea $X = (X_1, \dots, X_d)$ un vector de variables aleatorias independientes y $f \colon \mathbb{R}^d \longrightarrow \mathbb{R}$ una función que satisface $\mathbb{E}\left[|f(X)|^2\right] < \infty$. Consideremos el vector $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}'_{i-1}, X_i, \tilde{X}_{i+1} \dots, \tilde{X}_d)$, donde cada $\tilde{X}_j$ $(j \neq i)$ es una variable aleatoria independiente de las demás y de misma distribución que X_j. Entonces,*

$$\operatorname{Cov}(f(X), f(\tilde{X})) = \operatorname{Var}\left(\mathbb{E}\left[f(X) \mid X_{-i}\right]\right).$$

Demostración. Supongamos sin pérdida de generalidad que $f(X)$ (y por lo tanto $f(\tilde{X})$) es centrada. Entonces

$$\operatorname{Cov}(f(X), f(\tilde{X})) = \mathbb{E}\left[f(X) f(\tilde{X})\right] = \mathbb{E}\left[\mathbb{E}\left[f(X) f(\tilde{X}) \mid X_i\right]\right].$$

Al fijar X_i con la esperanza condicional y por la independencia entre todas las variables, $f(X)$ y $f(\tilde{X})$ son variables aleatorias independientes. Por lo tanto,

$$\operatorname{Cov}(f(X), f(\tilde{X})) = \mathbb{E}\left[\mathbb{E}\left[f(X) \mid X_i\right] \mathbb{E}\left[f(\tilde{X}) \mid X_i\right]\right] = \mathbb{E}\left[\mathbb{E}\left[f(X) \mid X_i\right]^2\right] = \operatorname{Var}\left(\mathbb{E}\left[f(X) \mid X_i\right]\right).$$

La segunda igualdad se debe a que las variables X_j y $\tilde{X}_j$ tienen la misma distribución para todo $j \neq i$. ∎

Considerando la expresión alternativa de S_i^{tot} dada en el Ejercicio 2.3, podemos crear un primer estimador de S_i^{tot} usando dos grupos de muestras independientes $X^1, \dots, X^N$ y $\tilde{X}^1, \dots, \tilde{X}^N$ (con cada $\tilde{X}^k$ tal como se menciona en la Proposición 2.1). Este estimador, llamado estimador "*pick and freeze*" (elige tu variable y congélala), se define por:

$$\widehat{S_i^{\text{tot}}} = 1 - \frac{\frac{1}{N}\sum_{k=1}^N f(X^k) f(\tilde{X}^k) - \left(\frac{1}{N}\sum_{k=1}^N f(X^k)\right)\left(\frac{1}{N}\sum_{k=1}^N f(\tilde{X}^k)\right)}{\frac{1}{N}\sum_{k=1}^N f(X^k)^2 - \left(\frac{1}{N}\sum_{k=1}^N f(X^k)\right)^2}$$

Note entonces que, a pesar de su alta interpretabilidad, la estimación de los índices de Sobol a partir del médoto *pick and freeze* requiere una muestra de gran tamaño, convirtiéndolos en una herramienta costosa en términos de recursos computacionales. Por esta razón, se buscan alternativas más económicas para estimar o acotar S_i^{tot}, involucrando otros índices de sensibilidad, como por ejemplo los índices DGSM que revisamos en la siguiente sección.

3 Indices DGSM

Cuando disponemos de las derivadas (o de las derivadas débiles, que definiremos en la Lección 2) de la función $f : \mathbb{R}^d \to \mathbb{R}$, podemos calcular los índices DGSM (por sus siglas en inglés, Derivative Global Sensitivity Measure).

Definición 2 (Indices DGSM) *Sea $X = (X_1, \dots, X_d)$ un vector de variables aleatorias independientes y $f \colon \mathbb{R}^d \to \mathbb{R}$ una función (débilmente) diferenciable que satisface $\mathbb{E}\left[|f(X)|^2\right] < \infty$. Entonces, el índice DGSM asociado a la variable X_i se define como*

$$\nu_i = \mathbb{E}\left[\left|\frac{\partial f}{\partial x_i}(X)\right|^2\right],$$

siempre que esta cantidad exista.

Cada índice DGSM mide en promedio la variabilidad de f con respecto a su variable objetivo y por definición ellos son computacionalmente económicos (se los puede estimar por ejemplo, mediante Monte-Carlo). Sin embargo, esta medición no siempre refleja la influencia real de las variables, especialmente cuando f varía muy rápidamente. Considere por ejemplo el modelo siguiente, para el cual el índice DGSM de una variable que no tiene una influencia significativa (según se mide con S_i^{tot}) resulta ser muy grande.

Ejemplo 3.1 *Sean $X_1 \sim \mathcal{U}([0,1])$ y X_2 dos variables aleatorias independientes, con X_2 tal que $\mathrm{Var}(X_2) > 1/2$, y considere el modelo $f(X_1, X_2) = \cos(\pi n X_1) + X_2$ ($n \geq 1$ entero). Entonces se tiene que*

$$S_1^{\mathrm{tot}} = \frac{\frac{1}{2}}{\frac{1}{2} + \mathrm{Var}(X_2)} < \frac{\mathrm{Var}(X_2)}{\frac{1}{2} + \mathrm{Var}(X_2)} = S_2^{\mathrm{tot}}.$$

Pero

$$\nu_1 = \frac{1}{2}\pi^2 n^2 > 1 = \nu_2$$

Ejercicio 3.1 *Verifique los valores de los índices del ejemplo anterior.*

En lugar de emplear ν_i para evaluar la influencia de X_i, lo utilizaremos para obtener información sobre S_i^{tot}, a través de la cota superior

$$S_i^{\mathrm{tot}} \leq \frac{C_P(\mu_i)}{D} \nu_i. \tag{2}$$

La medida μ_i denota la distribución de X_i y $C_P(\mu_i)$ es la constante óptima de su *desigualdad de Poincaré*, sobre la cual se discutirá en la siguiente lección. En particular, si se dispone de $C_P(\mu_i)$ o si se la aproxima, la parte derecha de (2) proporciona una alternativa económica para realizar *screening*.

Referencias

[1] M. Lamboni, B. Iooss, A.-L. Popelin, and F. Gamboa. Derivative-based global sensitivity measures: General links with sobol' indices and numerical tests. *Mathematics and Computers in Simulation*, 87:45–54, 2013.

[2] O. Roustant, F. Barthe, and B. Iooss. Poincaré inequalities on intervals – application to sensitivity analysis. *Electronic Journal of Statistics*, 11(2):3081 – 3119, 2017.

[3] A. Lagnoux. An introduction to sensitivity analysis. Notas de curso, Master M2R, Universidad Jean Jaurès. Toulouse, Francia.