Lección n°3: Conexión entre el análisis de sensibilidad global y las desigualdades de Poincaré. AMARUN 2024

1 Cota superior para los índices de Sobol totales

Una vez introducidos los índices de Sobol y los índices DGSM en la lección 1, establecemos su conexión a partir de una desigualdad de Poincaré. Detallamos la demostración propuesta en [1].

Recordemos brevemente las definiciones de los índices. Considere $X = (X_1, \dots, X_d)$, un vector de variables aleatorias independientes y una función $f\colon \mathbb{R}^d \to \mathbb{R}$ (débilmente) diferenciable en cada variable, que satisface $\mathbb{E}[|f(X)|^2] < \infty$. Dado $I \subseteq \{1, \dots, d\}$, denote por X_I el vector compuesto de las variables X_i, con $i \in I$, y $X_{-i} = X_{\{1,\dots,d\}\setminus\{i\}}$.

- La descomposición de Hoeffding-Sobol es la única descomposición de $f(X)$ de la forma

$$f(X) = \sum_{I \subseteq \{1,\dots,d\}} f(X_I), \tag{1}$$

a condición de que

$$\mathbb{E}\left[f_I(X_I) \,|\, X_J\right], \qquad \text{para todo } J \subsetneq I. \tag{2}$$

Entonces el índice de Sobol total de X_i se define como

$$S_i^{\text{tot}} = \frac{\operatorname{Var}(f_i^{\text{tot}}(X))}{\operatorname{Var}(f(X))} = \frac{\operatorname{Var}\left(\sum_{I \ni i} f_I(X_I)\right)}{\operatorname{Var}(f(X))}.$$

- El índice DGSM asociado a X_i se define como $\nu_i = \mathbb{E}\left[\left|\frac{\partial f}{\partial x_i}(X)\right|^2\right]$.

Necesitamos además dos propiedades importantes con respecto a la esperanza condicional de variables aleatorias, cuyas demostraciones dejamos como ejercicio:

Ejercicio 1.1 *Sean Y y Z dos variables aleatorias reales tales que $\mathbb{E}\left[|Y|^2\right] < \infty$, $\mathbb{E}\left[|Z|^2\right] < \infty$.*

- *Pruebe que*

$$\mathbb{E}[\mathbb{E}[Y \mid Z]] = \mathbb{E}[Y].$$

- *Pruebe que si Y y Z son independientes y f es una función que satisface $\mathbb{E}\left[|f(X,Y)|\right]^2 < \infty$, entonces se tiene que*

$$\mathbb{E}\left[f(Y,Z) \,|\, Z\right] = g(Z), \qquad \text{donde } g(z) = \mathbb{E}\left[f(Y,z)\right].$$

En pocas palabras, la segunda propiedad nos indica que para calcular $\mathbb{E}\left[f(Y,Z) \,|\, Z\right]$, cuando las variables son independientes, basta con considerar Z como constante (como si no fuera aléatorio) y tomar la esperanza con respecto a Y.

Sin más preámbulos, presentamos la cota superior junto con su demostración.

Teorema 1.1 *Sea $X = (X_1, \dots, X_d)$ un vector de variables aleatorias independientes. Sea $f\colon \mathbb{R}^d \longrightarrow \mathbb{R}$ una función que satisface $\mathbb{E}\left[|f(X)|^2\right] < \infty$. Supongamos que la distribución de X_i, denotada por μ_i, está definida en un intervalo (a_i, b_i) y satisface la desigualdad de Poincaré:*

$$\mathbb{E}\left[|g(X_i)|^2\right] \le C_P(\mu_i)\mathbb{E}\left[|g'(X_i)|\right]^2, \tag{3}$$

para todo $g \in H^1_{\mu_i}(a_i, b_i)$. Entonces

$$S_i^{\text{tot}} \le \frac{C_P(\mu_i)}{\operatorname{Var}(f(X))}\nu_i. \tag{4}$$

Demostración. Debemos demostrar que

$$\operatorname{Var}\left(\sum_{I\ni i} f_I(X_I)\right) \leq C_P(\mu_i)\,\mathbb{E}\left[\left|\frac{\partial f}{\partial x_i}(X)\right|^2\right].$$

Dado que $\sum_{I\ni i} f_I(X_I)$ es una variable aleatoria centrada, podemos escribir

$$\operatorname{Var}\left(\sum_{I\ni i} f_I(X_I)\right) = \mathbb{E}\left[\left(\sum_{I\ni i} f_I(X_I)\right)^2\right] = \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{I\ni i} f_I(X_I)\right)^2 \middle| X_{-i}\right]\right].$$

Gracias a la independencia entre X_i y X_{-i} y a la segunda propiedad del Ejercicio 1.1, la cantidad aleatoria dentro de la esperanza condicional solo depende de X_i. Además, $\sum_{I\ni i} f_I(X_I)$ es centrada con la esperanza condicional respecto a X_{-i}, ya que

$$\mathbb{E}\left[\sum_{I\ni i} f_I(X_I)\middle| X_{-i}\right] = \sum_{I\ni i}\mathbb{E}\left[f_I(X_I)\middle| X_{-i}\right] = \sum_{I\ni i}\mathbb{E}\left[f_I(X_I)\middle| X_{I\setminus\{i\}}\right] = 0.$$

En la penúltima igualdad usamos el hecho de que cada X_I no depende de $\{1,\ldots,d\}\setminus I$ y en la última igualdad aplicamos la condición (2). Esto nos permite meter en colación la desigualdad de Poincaré (3) con la función $\sum_{I\ni i} f_J$ (vista como una función de x_i, con x_{-i} fijo). Se sigue que

$$\begin{aligned}\operatorname{Var}\left(\sum_{I\ni i} f_I(X_I)\right) &\leq \mathbb{E}\left[C_P(\mu_i)\mathbb{E}\left[\left(\frac{\partial}{\partial x_i}\left(\sum_{I\ni i} f_I(X_I)\right)\right)^2\middle| X_{-i}\right]\right]\\ &= C_P(\mu_i)\mathbb{E}\left[\left(\sum_{I\ni i}\frac{\partial f_I}{\partial x_i}(X_J)\right)^2\right].\end{aligned} \tag{5}$$

Pero gracias a la descomposición de f en (1), para $i\in I$ fijo tenemos que

$$\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\sum_{J\subseteq\{1,\ldots,d\}} f_J\right) = \sum_{I\ni i}\frac{\partial f_I}{\partial x_i},$$

que es la misma expresión que aparece en la esperanza en el lado derecho de (5), con lo cual la prueba queda completa. ■

Dejamos como ejercicio un simple ejemplo de aplicación de la cota superior.

Ejercicio 1.2 *Considere el siguiente modelo $f(X_1,X_2)=\cos(\pi X_1)+\max(0,X_2)$, donde $X_1\sim\mathcal{U}([0,1])$ y $\mathcal{N}(0,1)$ son dos variables aleatorias independientes. Pruebe que los índices de Sobol del modelo son*

$$S_1^{\text{tot}} = \frac{\frac{1}{2}}{1-\frac{1}{2\pi}} \approx 0.5946, \qquad S_2^{\text{tot}} = \frac{\frac{1}{2}-\frac{1}{2\pi}}{1-\frac{1}{2\pi}} \approx 0.4054$$

y sus cotas superiores están dadas por

$$\frac{C_P(\mu_1)}{\operatorname{Var}(f(X_1,X_2))}\nu_1 = \frac{\frac{1}{2}}{1-\frac{1}{2\pi}} \approx 0.5946, \qquad \frac{C_P(\mu_2)}{\operatorname{Var}(f(X_1,X_2))}\nu_2 = \frac{\frac{1}{2}}{1-\frac{1}{2\pi}} \approx 0.5946,$$

respectivamente.

Es pertinente recordar que, en general, es complicado obtener el valor teórico de la constante $C_P(\mu_i)$. Uno podría tentar evitar el cálculo de la constante reduciendo el problema al caso uniforme, en el cual se sabe que la constante óptima de Poincaré es igual a $1/\pi^2$. En efecto, es posible transformar el modelo $f(X_1,\ldots,X_d)$ en

$$g(U_1,\ldots,U_d) = f(F_1^{-1}(U_1),\ldots,F_d^{-1}(U_d)),$$

donde $U_1, \ldots, U_d \sim \mathcal{U}([0,1])$ son variables aleatorias independientes y cada F_i representa la función de distribución de X_i (vea el método de la transformación inversa en el **Ejercicio** 1.2). De esta manera, los índices de Sobol en ambos modelos son iguales (**Ejercicio 1.3** Pruébelo). Sin embargo, las derivadas de las funciones F_i^{-1} podrían ser muy grandes en el cálculo de los nuevos índices DGSM. Esto se ilustra con el siguiente ejemplo, propuesto en [2].

Ejemplo 1.1 *Sean $X_1, X_2 \sim \mathcal{N}(0,1)$, $U_1, U_2 \sim \mathcal{U}([0,1])$, cuatro variables aleatorias independientes. Denote por Φ la función de repartición de la distribución normal $\mathcal{N}(0,1)$, definida por $x \in \mathbb{R} \mapsto \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$,y considere los modelos $f(X_1, X_2) = X_1 + X_2$, $g(U_1, U_2) = \Phi^{-1}(U_1) + \Phi^{-1}(U_2)$. Entonces:*

- *Todos los índices de Sobol totales de $f(X_1, X_2)$ y $g(U_1, U_2)$ valen $1/2$.*
- *Los índices DGSM de $f(X_1, X_2)$ valen 1, mientras que los de $g(U_1, U_2)$ son infinitos.*

Como cada variable tiene el mismo protagonismo en sus respectivos modelos, es claro que cada índice de Sobol es igual a $1/2$. Denotamos por $\nu_1(f)$, $\nu_1(g)$ a los índices DGSM de X_1 en cada modelo. Por un lado, es evidente que $\nu_1(f) = 1$. Por otro lado, calculamos

$$\nu_1(g) = \mathbb{E}\left[\left(\frac{\partial g}{\partial u_1}(U_1, U_2)\right)^2\right] = \mathbb{E}\left[\left(\left(\Phi^{-1}\right)'(U_1)\right)^2\right] = \mathbb{E}\left[\frac{1}{\left(\Phi'(\Phi^{-1}(U_1))\right)^2}\right].$$

Pero, puesto que $\Phi^{-1}(U_1)$ sigue la misma distribución de X_1, podemos escribir

$$\nu_1(g) = \mathbb{E}\left[\frac{1}{(\Phi'(X_1))^2}\right] = \mathbb{E}\left[\left(\frac{1}{\sqrt{2\pi}} e^{-X_1^2/2}\right)^{-2}\right] = 2\pi\mathbb{E}\left[e^{X_1^2}\right] = \sqrt{2\pi}\int_{\mathbb{R}} e^{\frac{1}{2}x^2} dx = \infty.$$

Ejercicio 1.2 (Método de la transformada inversa) *Sea X una variable aleatoria real con función de repartición $F_X : \mathbb{R} \to [0,1]$. Sea $U \sim \mathcal{U}([0,1])$. Definimos la inversa generalizada de F como*

$$u \in [0,1] \mapsto F_X^{-1}(u) = \inf\{x \in \mathbb{R} \,|\, F_X(x) \geq u\}.$$

Pruebe que $F_X^{-1}(U)$ sigue la misma distribución que X.

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.