Estudio de la influencia a través de modelos de opinión

Ramírez Puertas, Pablo
Esteban Agenjo, Raúl

21 Mayo, 2022

Resumen

En este trabajo se estudia un modelo de opinión discreto y se presenta una propuesta para controlar la evolución del modelo mediante la modificación de los parámetros iniciales gracias a la introducción de individuos adicionales en la población objeto de estudio; así como la optimización de costes asociados mediante *software* de optimización no lineal.

1. Modelo de Opinión

1.1. ¿Qué es?

Hablamos de *modelos de opinión* cuando nos referimos a modelos matemáticos que tienen por objeto describir la evolución del conjunto de las opiniones de una población de personas.

Por ejemplo, supongamos que tomamos una muestra de 100 personas y a estas se les pide que evalúen con entre 1 y 10 puntos la calidad de un producto, su simpatía por un candidato político, etc. Cada uno de estos individuos tiene una opinión clara, pero ¿es inamovible? Si permitimos que esta población, tras ser encuestada, interactúe consigo misma, intercambiaría opiniones, críticas y puntos positivos del objeto, la marca o persona sobre la que fueron encuestados y esto hará que su opinión varíe. Si, por ejemplo, pasadas 2 horas volviésemos a encuestar a la población, veríamos que las estadísticas han cambiado. Predecir y entender la evolución de la opinión es lo que persiguen estos modelos.

1.2. Desarrollo del modelo

Empecemos tomando un conjunto de individuos, N. Cada uno de estos individuos tiene una opinión en el intervalo $[0, 1]$ y al interactuar con otro sujeto, la opinión de ambos sufre un modificación h constante en función de sus posiciones relativas, es decir, si denotamos por $\omega_i(t)$ la opinión de un individuo i a lo largo del tiempo, tras interactuar en el momento $t + dt$ con otro individuo tendremos

$$\omega_i(t+dt) = \begin{cases} \omega_i(t) + h & \text{, si } \omega_i(t) < \omega_j(t) \\ \omega_i(t) - h & \text{, si } \omega_i(t) > \omega_j(t) \\ \omega_i(t) & \text{, si } \omega_i(t) = \omega_j(t) \end{cases}$$

Para obtener las ecuaciones dividiremos el espacio de opiniones en una partición de varios subintervalos con la misma longitud : $I_i, \ldots, I_n \subset [0, 1]$. Ahora tomaremos la densidad de población en esos intervalos en un instante t: $S(i,t) = \frac{\# I_i}{N}$.

Si asumimos que la interacción entre individuos viene dada por una variable aleatoria Poisson de parámetro 1, podemos esperar que ocurra una interacción en promedio por unidad de tiempo. Pasado este tiempo, podemos ver la evolución del intervalo como la cantidad de individuos que había antes más la de nuevos individuos que entran menos la de individuos que sale, es decir:

$$S(i, t + dt) = S(i, t) + \frac{2}{N} (G(i, t) - L(i, t)) \tag{1}$$

, donde L y G representan los factores de pérdida y ganancia respectivamente. Como en el intervalo de tiempo dt solo interactúan dos individuos que cambian su opinión, $S(j, t)$ variará solo en $\frac{1}{N}$

Para cada caso la probabilidad de entrar a I_j es la probabilidad de encontrar un individuo en la vecindad, es decir en el intervalo I_{j+1} (respectivamente en el

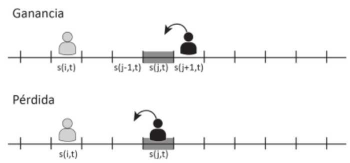

Figura 1: Perdida y ganancia de intervalos

intervalo I_{j-1}), y otro en un intervalo I_i con $i \leq j$ (respectivamente $i \geq j$) de manera que luego de la interacción el primer individuo se mueva h e ingrese al intervalo j. Para cada j y cada $t > 0$ tenemos

$$G(i,t) = S(i+1,t) \sum_{j \leq i} S(j,t) + S(i-1,t) \sum_{j \geq i} S(j,t) \tag{2}$$

De manera análoga obtenemos el término de pérdida

$$L(i,t) = S(i,t) \sum_{j \neq i} S(j,t) \tag{3}$$

$$L(i,t) = S(i,t) \left(\sum_{j \leq i} S(j,t) + \sum j \geq i S(j,t) - 2S(j,t) \right) \tag{4}$$

Sustituyendo en 1 los resultados 4 y 2 llegamos a la siguiente expresión:

$$\begin{aligned} S(i,t+dt) = & S(i,t) + \frac{2}{N} \bigg((S(i+1,t) - S(i,t)) \sum_{j \leq i} S(j,t) \\ & + (S(i-1,t) - S(i,t)) \sum_{j \geq i} S(j,t) + 2S(j,t)^2 \bigg) \end{aligned}$$

2. Formulación del problema

2.1. Estableciendo un objetivo

Aprovechando el modelo de opinión que hemos descrito en la sección anterior, nos planteamos ahora encontrar alguna forma de explotarlo. Para ello buscaremos, pasado un tiempo T, conseguir que una cierta cantidad de población se sitúe entre unos ciertos valores de opinión de interés. Para lograr este efecto en la población, introduciremos sujetos, con una opinión conocida, dentro del sistema, intentando minimizar los posibles costes asociados a estas decisiones.

En resumen:

- Tiempo límite: T.
- Valor de éxito deseado: l.
- Intervalos de interés: $I_1, \ldots, I_k$.
- Número de sujetos introducidos artificialmente en el intervalo I_j: $P(j)$.
- Costes asociados a introducir individuos en el intervalo I_j: c_j.

2.2. Restricciones

Para construir las restricciones consideramos: un número máximo de sujetos que se pueden introducir artificialmente en función del intervalo, la cantidad de población que queremos acumular en que intervalos, las condiciones iniciales de población y el modelo que describe la evolución del sistema.

Para la cota de sujetos introducidos artificialmente hemos optado por establecer como límite el mismo número de individuos del conjunto, es decir,

$$P(i) \leq \#I_i,\ 1 \leq i \leq n.$$

Para la condición de éxito, dado que los $S(i,t)$ dan la cantidad de población en términos de densidad, la condición será $l \in (0,1)$, de modo que

$$\sum_{q=1}^{k} S(J_q, t) \geq l.$$

A continuación planteamos las condiciones iniciales del modelo. Simplemente añadimos individuos al intervalo y dividimos por el total de sujetos dentro del sistema para obtener la densidad inicial:

$$S(i,0) = \frac{\#I_i + P(i)}{N + \sum_{j=1}^{n} P(j)}\ ,\ 1 \leq i \leq n.$$

Por último usaremos la versión discreta del modelo, teniendo presente que el total de individuos en el sistema ahora es distinto:

$$\begin{aligned} S(i,t+1) = &S(i,t) + \frac{2}{N + \sum_{j=1}^{n} P(j)} \Bigg(\big(S(i+1,t) - S(i,t)\big) \sum_{k=1}^{i} S(k,t) + \\ &\big(S(i-1,t) - S(i,t)\big) \sum_{k=i}^{n} S(k,t) \Bigg) + 2S(i,t)^2 \end{aligned}$$

2.3. Función objetivo

Como se indicó antes, nuestra función objetivo tendrá la forma

$$\text{mín} \sum_{j=1}^{n} c_j P(j). \tag{5}$$

Si los valores c_j son todos 1, el problema se traduce en minimizar la cantidad de sujetos que se introducen artificialmente en el sistema. Al añadir estos factores c_j podemos añadir distintas interpretaciones al objetivo del modelo, como por ejemplo, reducir los gastos que puedan estar asociados a buscar estos sujetos. Por esta razón, para el modelo se ha elegido de manera particular unos valores c_j de modo que los costes asociados a los extremos son mayores que los asociados a los intervalos centrales y todos se reparten de forma simétrica.

3. Resultados y Conclusiones

3.1. Sobre la población de estudio

Una particularidad de este modelo es la necesidad de partir de un conjunto de datos que describan una población inicial dividida en intervalos. Para poder realizar diversas experimentaciones de forma cómoda y razonable se han generado diversas poblaciones en base a una distribución aleatoria uniforme en el intervalo $[0, 1]$ y después, sobre esa población se ha contabilizado el número de individuos en cada uno de los subintervalos de estudios mediante un algoritmo en R («generador_muestra.R»[1]). Los datos después son pasados a un fichero «.dat» para poder usarlos en AMPL.

3.2. Estudio sobre una población

Para ejemplificar los resultados obtenidos nos centraremos en un caso particular. En este supuesto se ha generado una población de 344 individuos y 49 intervalos de opinión. La población ha sido distribuida conforme a una variable aleatoria uniforme.

En la figura 2 podemos ver en la gráfica negra la disposición original de la población, los individuos insertados artificialmente se representan por los puntos estrellados y la gráfica roja recoge la disposición final de toda la población (la incluida artificialmente inclusive).

En la figura 3 observamos la resolución del mismo problema cuando aumentamos el tiempo transcurrido.

Notar que el problema varía de una gráfica a otra, pues estamos exigiendo las mismas condiciones de éxito en instantes de tiempo distintos. Consecuencia de esto podemos ver un desplazamiento de los individuos introducidos. Este desplazamiento conlleva un aumento de la función objetivo, de donde se puede deducir que cuanto más largo es el instante de tiempo, más difícil es mover a la población de estudio a las zonas de interés.

Si dejásemos al modelo evolucionar libremente, como se puede observar en la gráfica 4[2], la población tiende a centralizarse entorno a la distancia media, en general, la población se amolda a la distribución de una variable normal $\mathcal{N}(0, 1)$. Este hecho nos permite inferir que el aumento de la función objetivo conforme aumenta el tiempo se debe a la *lucha* entre la tendencia natural de la población y la búsqueda de valores de interés.

3.3. Conclusiones

Como hemos visto, el modelo de opinión presentado describe una tendencia a la centralización de las opiniones en una población con el tiempo.

[1] El algoritmo precisa 2 datos. `N`, el número de participantes y `t` el número de subintervalos. Debido a errores en la programación que no se han podido resolver a la hora de clasificar los elementos, el conjunto de datos final consta de menos de `N` elementos y aparece `t-1` intervalos.

[2] La gráfica no se corresponde con la evolución de la población descrita anteriormente, aunque el resultado es válido para cualquier población.

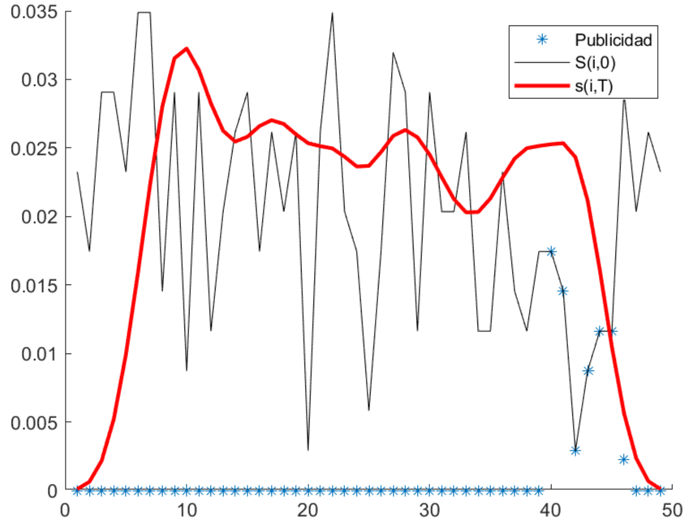

Figura 2: Evolución tras t=1000

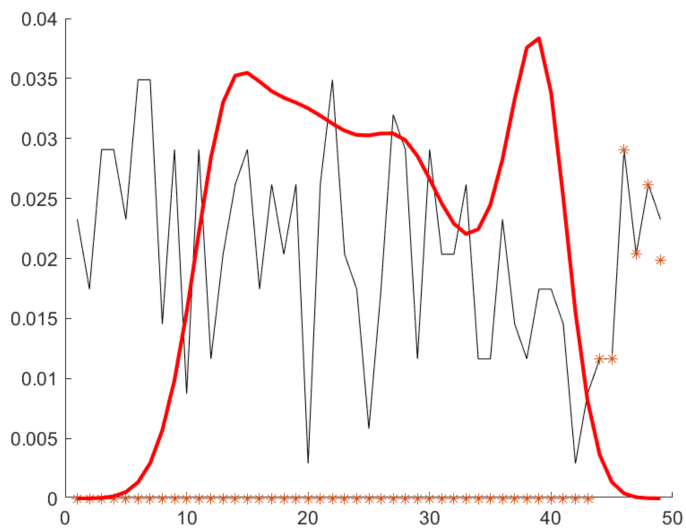

Figura 3: Evolución tras t=2000

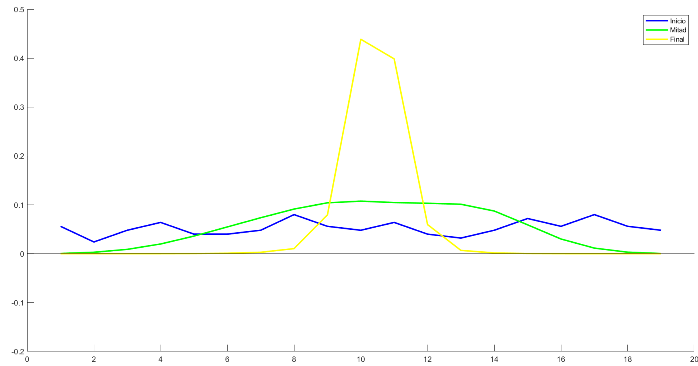

Figura 4: Evolución libre

La revisión del modelo propuesta podría representar situaciones en las que un marca se interesa en obtener una buena opinión del público general para una fecha particular, y como estrategia decide introducir individuos en la población objetivo con una opinión particular sobre su producto para modificar la opinión. Este modelo asume que los individuos son libres de cambiar su opinión. Otra forma de enfocar el acercamiento del modelo a la realidad estaría en el modo en el que las redes sociales nos presentan nuevos perfiles para poder seguir. La elección adecuada de estas personas con opiniones particulares que la red podría estimar, afectaría a como nuestra evolución respecto a una opinión puede cambiar con el tiempo según las interacciones que realicemos.

4. Bibliografía

- Pedraza, L. y Pinasco, J.P.(2018). *Modelos de formación de opinión.* Universidad de Buenos Aires.
- Grima, C. (2018). *¡Que las matemáticas te acompañen!.* Ariel