RSME 2019 SANTANDER

El problema del transporte óptimo y algunas de sus aplicaciones estadísticas

Juan A. Cuesta-Albertos
Universidad de Cantabria

El problema del transporte óptimo fue formulado inicialmente por Monge en 1781 y consiste en minimizar el coste de transportar un montón de arena a una localización previamente fijada. Suponiendo que la arena está compuesta por granos, el objetivo consiste en determinar el punto exacto al que debe ser trasladado cada grano, teniendo en cuenta que

1. se conoce tanto la situación inicial de cada grano, pudiendo haber varios apilados en el mismo punto, como el número de granos que tiene que haber finalmente en cada uno de los puntos de destino

2. el coste del transporte de cada grano es una función conocida $c : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^+$ que sólo depende de las posiciones inicial y final del grano y el coste total del transporte es la suma de los costes de transportar los granos individuales

Aparte de la generalización obvia sobre el espacio en el que se encuentra la arena, es interesante la introducida por Kantorovich en 1942 según la cual las condiciones anteriores siguen vigentes, pero los granos pueden ser divididos de modo que cada grano puede ser repartido en un área (elegida por el transportista) de la localización final.

Ambos problemas comparten el hecho de que las cantidades inicial y final de arena coinciden. Si suponemos que la masa es finita, el problema viene a ser la búsqueda de un emparejamiento óptimo entre dos distribuciones de probabilidad, y, bajo condiciones adecuadas sobre c, sucede que el coste del transporte induce una métrica entre distribuciones de probabilidad. En el caso del espacio euclídeo y $c(x, y) = \|x - y\|^2$, la *distancia de Wasserstein* entre dos distribuciones de probabilidad, P, Q, en la σ-álgebra de Borel en $\mathbb{R}^d$ vale

$$\mathcal{W}_2(P, Q) := \inf_{X,Y} \left\{ \left(\int_\Omega \|X - Y\|^2 d\mu \right)^{1/2} \right\},$$

donde el ínfimo se toma en todas las posibles parejas de variables aleatorias X e Y que pueden coexistir en algœn espacio probabilístico (Ω, σ, μ) satisfaciendo que sus distribuciones son, respectivamente, P y Q.

En esta charla prestaremos atención al interés de este problema desde un punto de vista estadístico y al estado del conocimiento en el campo.