



PageRank: el algoritmo de Google

En cuanto a la teoría de grafos: todos los puntos son básicos, pero este de aquí, de entrada, permite una ordenación un poco abstracta. Nosotros nos imaginamos que todas las páginas web están interconectadas, podemos pensar fácilmente que son una serie de puntos conectados con unas líneas, con unas aristas. Esto se extrapola a muchos problemas. Por ejemplo, podrían ser conexiones eléctricas entre diferentes ciudades, podrían ser oleoductos, diferentes conexiones de oleoductos, podrían ser perfectamente también conexiones de comunicaciones o de ciudades.

El típico problema en teoría de grafos es "cuál es el recorrido que tiene que hacer un vendedor para pasar por diferentes ciudades de manera que recorra la mínima distancia posible y pase por todas ellas solo una vez". Es una cosa muy abstracta pero que tiene una gran cantidad de aplicaciones hoy en día. La aplicación de las matrices en este algoritmo a pequeña escala es muy fácil y se entiende muy bien.

Obviamente, después vienen los problemas derivados del hecho que la matriz de Google es una matriz enorme, es una matriz muy grande. Estamos hablando de 2,7 billones por 2,7 billones, es una cosa gigantesca. Enormemente grande. Entonces, ¿qué pasa? Toda la teoría de matrices es una teoría que los estudiantes ven cuando empiezan el bachillerato, sobre todo cuando hacen resoluciones de sistemas de ecuaciones lineales. Pero es una herramienta que acompaña a las matemáticas desde el principio hasta el final porque permite acumular mucha información en un lugar, permite hacer cálculos muy rápido y fácilmente...

De hecho son una especie de armario que contiene información importante para tu problema. Esto que comentábamos de las matrices... Una cosa es manipular matrices pequeñas, y otra cosa es manipular matrices muy grandes. Todos los *focus* que se incorporan al *page rank* hacen manipular matrices muy grandes. ¿Y eso qué implica? Que tienes que diseñar algoritmos muy eficientes que te permitan calcular estos famosos vectores propios. Aquí entramos en una disciplina de las matemáticas que son el cálculo numérico o los métodos numéricos. Una cosa es la realidad y otra cosa es lo que tú calculas. Cuando calculamos cosas, realmente no es exactamente lo que querríamos calcular. Entonces, ¿cómo se hace de manera eficiente? El típico ejemplo



es el número Pi. Todo el mundo sabe que el número Pi tiene decimales infinitos y que nunca se repiten. Muy bien. Por lo tanto es imposible introducir Pi en un ordenador porque el ordenador siempre corta, tiene una capacidad máxima. Por lo tanto, no estás trabajando con Pi, sino que estás trabajando con una aproximación de Pi.

A partir de aquí, todo se va acumulando: los errores, las imprecisiones, las operaciones... Esencialmente, esto es el cálculo numérico.