



## PageRank: l'algoritme de Google

---

Respecte de la teoria de grafs: tots els punts són bàsics, però aquest d'aquí, primer permet una ordenació una mica abstracta. Nosaltres ens imaginem totes les pàgines web interconnectades, podem pensar fàcilment que són una sèrie de punts connectats amb unes línies, amb unes arestes. Això s'extrapola a molts problemes. Per exemple, podrien ser connexions elèctriques entre diferents ciutats, podrien ser oleoductes, diferents connexions d'oleoductes, podrien ser perfectament també connexions de comunicacions o de ciutats.

El típic problema en teoria de grafs és "quin és el recorregut que ha de fer un venedor per passar per diferents ciutats de manera que recorri la mínima distància i passi per totes elles només una vegada". És una cosa molt abstracta però que té una gran quantitat d'aplicacions avui dia. L'aplicació de les matrius en aquest algoritme a petita escala és molt fàcil i s'entén molt bé.

Òbviament, després vénen els problemes derivats del fet que la matriu de Google és una matriu enorme, és una matriu molt gran. Estem parlant de 2,7 bilions per 2,7 bilions, és una cosa gegantesca. Enormement gran. Què passa doncs? Tota la teoria de matrius és una teoria que els estudiants veuen quan comencen el batxillerat, sobre tot quan fan resolucions de sistemes d'equacions lineals. Però és una eina que acompanya les matemàtiques des del començament fins al final perquè permet acumular molta informació en un lloc, permet fer càlculs molt ràpid i fàcilment...

De fet són una mena d'armari que conté informació important per al teu problema. Això que comentàvem de les matrius... Una cosa és manipular matrius petites, i una altra cosa és manipular matrius molt grans. Tots els *focus* que s'incorporen al *page rank* fan manipular matrius molt grans. I això què implica? Que has de dissenyar algoritmes molt eficients que et permetin calcular aquests famosos vectors propis. Aquí entrem en una disciplina de les matemàtiques que és el càlcul numèric o els mètodes numèrics. Una cosa és la realitat i una altra cosa és el que tu calcules. Quan calculem coses, realment no és ben bé el que voldríem calcular. Aleshores, com fer-ho de manera eficient? El típic exemple és el número Pi. Tothom sap que el número Pi té



una infinitat de decimals i que mai no es repeteixen. Molt bé. Per tant és impossible introduir Pi en un ordinador perquè l'ordinador sempre talla, té una capacitat màxima. Per tant, no estàs treballant amb Pi sinó que estàs treballant amb una aproximació de Pi.

A partir d'aquí, tot es va acumulant: els errors, les imprecisions, les operacions... Això és el càlcul numèric essencialment.