

Seminário em Engenharia Matemática

Isabel Mendes Pinto

O PageRank do Google: cálculo e métodos de aceleração

O motor de busca Google tem tido um extraordinário sucesso em boa parte devido ao seu algoritmo PageRank. Este algoritmo classifica a importância das páginas existentes na internet de acordo com um vector próprio da matriz de adjacência do grafo associado às ligações entre essas páginas. Normalmente o cálculo do vector próprio é feito pelo método da potência cuja velocidade de convergência depende da separação entre os primeiro e segundo valores próprios, supondo uma numeração por ordem decrescente de valor absoluto. Este método tem uma grande vantagem de só precisar da matriz de iteração para efectuar produtos matriz-vector e, claro, a simplicidade de programação.

Quando os valores próprios são próximos o método da potência é lento, o cálculo do vector PageRank poderá demorar vários dias. O rápido crescimento da Web torna necessária a utilização de métodos que acelerem a sua convergência, tais como métodos de extrapolação, métodos adaptativos, técnicas de agregação/desagregação, entre outros.

Além do método da potência existem outros métodos iterativos mais rápidos em número de iterações, embora computacionalmente mais caros por iteração, como por exemplo toda a classe de métodos de tipo Arnoldi.