

Exercise: PageRank by the power method

Let $H \geq 0$ be an adjacency matrix: $H_{ij} > 0$ represents a link from page i to page j . Set $d = H\mathbf{1}$, $z_i = 1$ if $d_i = 0$ and $z_i = 0$ otherwise. Let $v = \mathbf{1}/n$ and $0 < \gamma < 1$.

$$S_{ij} = \begin{cases} H_{ij}/d_i, & d_i > 0, \\ 0, & d_i = 0, \end{cases} \quad A = \gamma(S + zv^T) + (1 - \gamma)\mathbf{1}v^T.$$

Implement `w = pagerank(H,gamma,maxit,tol)` to approximate the normalized left eigenvector $w^T A = w^T$, $\mathbf{1}^T w = 1$.

- Start from $w = v$ and compute each product without forming A :

$$w_{\text{new}} = \gamma(S^T w + v(z^T w)) + (1 - \gamma)v.$$

- Store H, S as sparse matrices. Handle dangling pages ($d_i = 0$) explicitly. Stop on $\|w_{\text{new}} - w\|_1 \leq \text{tol}$ or the iteration limit; check $\|A^T w - w\|_1$ using the same product formula.

PageRank: tests and interpretation

Use this small graph first (page 4 has no outgoing links):

$$H = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

- Convert H with `sparse(H)`. Check $w \geq 0$, $\sum_i w_i = 1$ and the eigenvector residual.
- Compare rankings and iteration counts for $\gamma = 0.5, 0.85, 0.95, 0.99$. Use `[importance,page] = sort(w,'descend')`.
- Repeat on a larger sparse graph of your choice, recording its source and the convention for link direction.

For larger graphs, compare runtime and memory usage as the number of nodes and links grows.