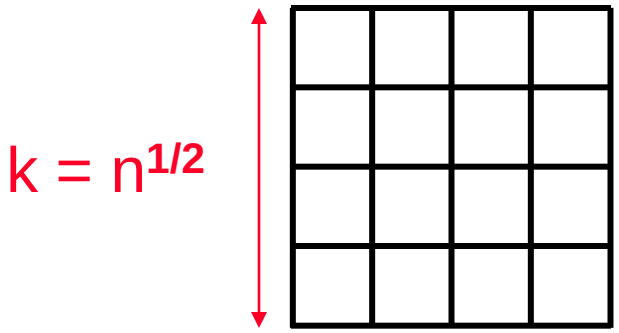
The "model problem"



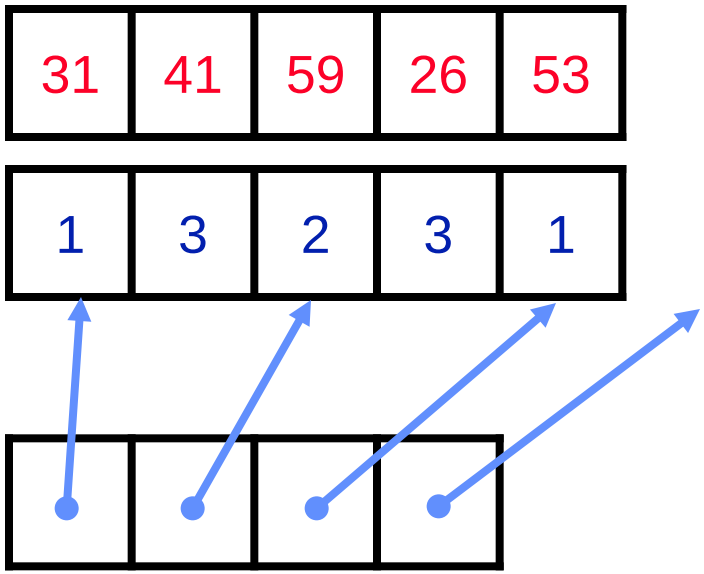
- Graph is a regular square grid with $n = k^2$ vertices.
- For each i from 1 to n, except on the boundaries:

$$-x(i-k) - x(i-1) + 4*x(i) - x(i+1) - x(i+k) = 0$$

- n equations in n unknowns: A*x = b
- Each row of A has at most 5 nonzeros.

Sparse matrix data structure (one example)

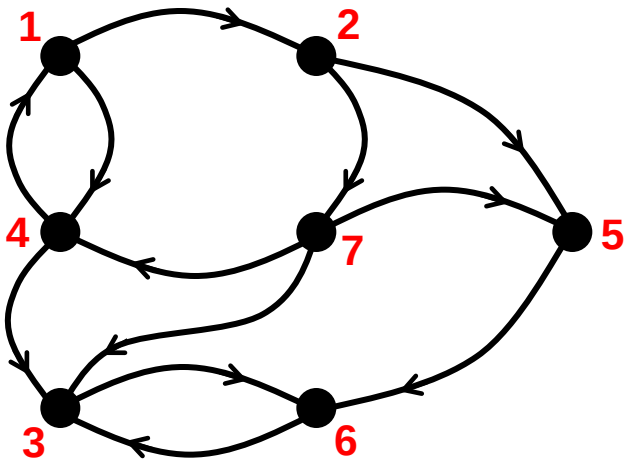
31	0	53
0	59	0
41	26	0

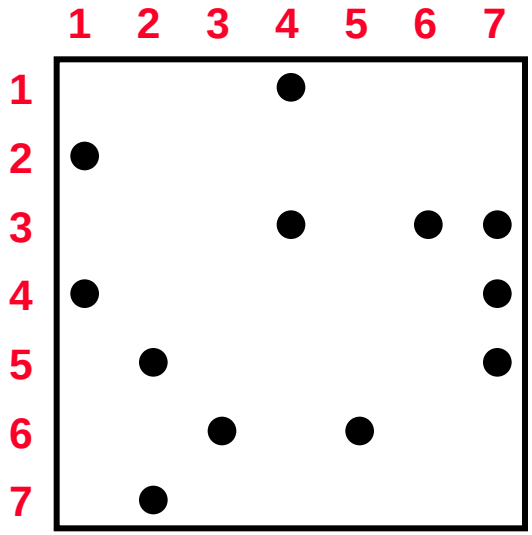


- Full:
  - 2-dimensional array of real or complex numbers
  - (nrows*ncols) memory
- Sparse:
  - compressed column storage (CSC)
  - about (2*nzs + ncols) memory

Graphs and Matrices

Graph

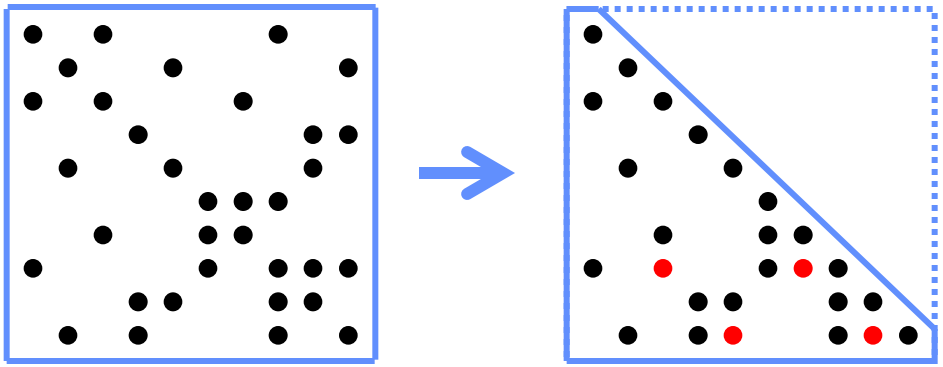




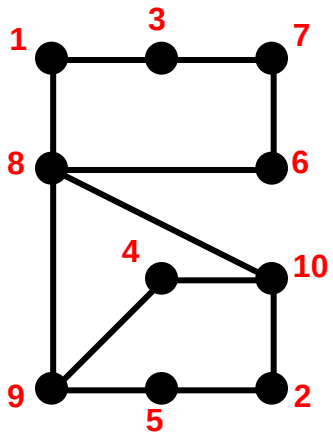
Matrix

- Starting with the matrix:
  - One graph vertex for each row (or column) of the matrix
  - One graph edge (i,j) for each nonzero A(j,i) in the matrix
  - (Some people point the edges the opposite way, from rows to columns; either way is ok as long as it's consistent.)
- Or, starting with the graph:
  - The adjacency matrix has A(j,i)=1 if (i,j) is an edge.

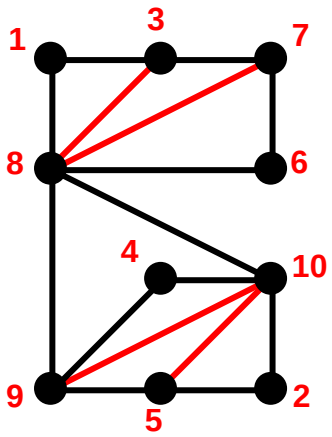
Graphs and Sparse Matrices: Cholesky factorization



Fill: new nonzeros in factor



G(A)



$G^+(A)$
[chordal]

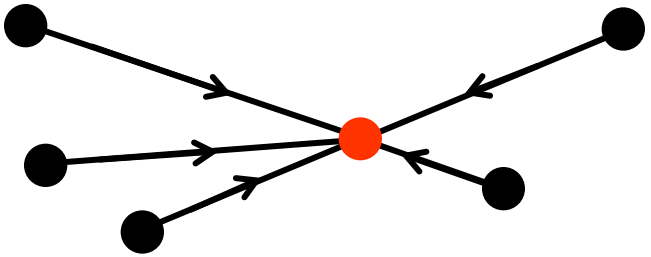
Symmetric Gaussian elimination:

for j = 1 to n

add edges between j's higher-numbered neighbors

Google and the Random Surfer

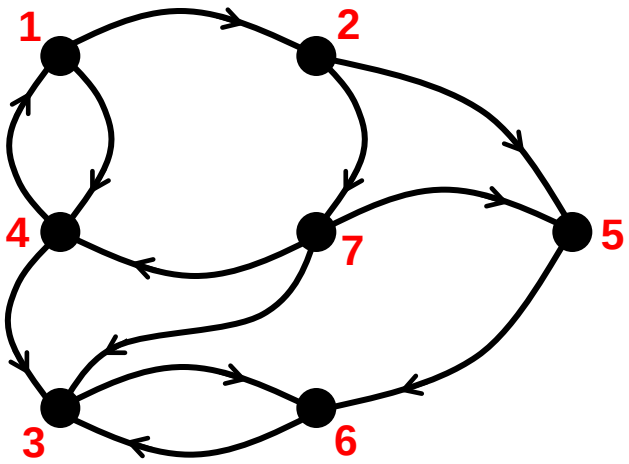
How does Google figure out which web pages are most important?

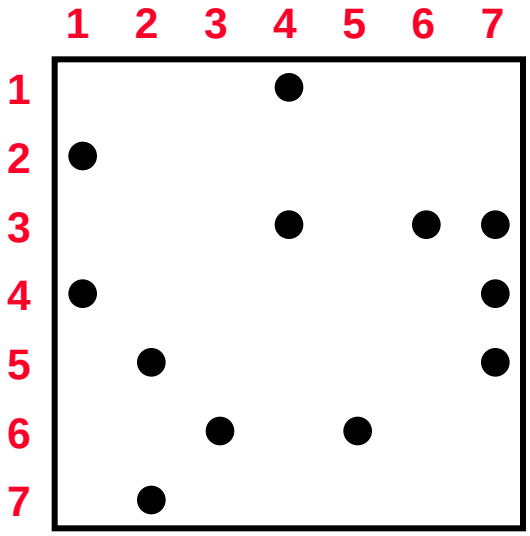


- An important page is one that lots of important pages point to.
- Start at any web page and follow links at random. Forever.
- You'll see "important" pages more often than unimportant ones.

Analyzing the Web with graphs and matrices

Graph



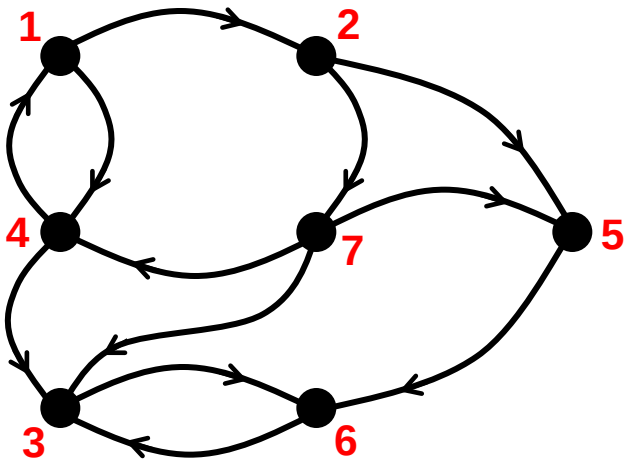


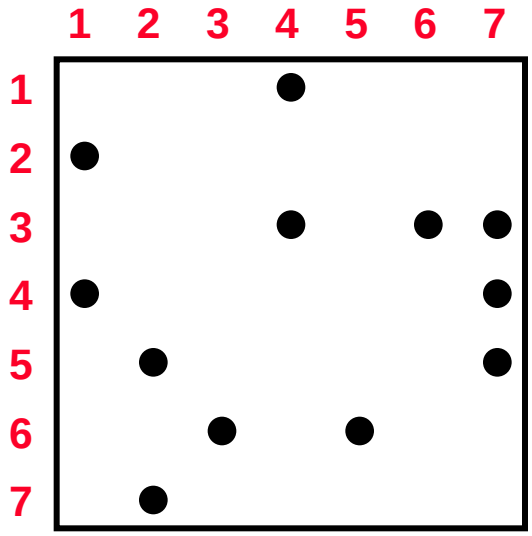
Matrix

- Graph nodes are web pages
- Arrows between nodes are links between web pages
- Matrix entries are links from “column” pages to “row” pages
- The Page Rank comes from algebra on the matrix

Analyzing the Web with graphs and matrices

Graph

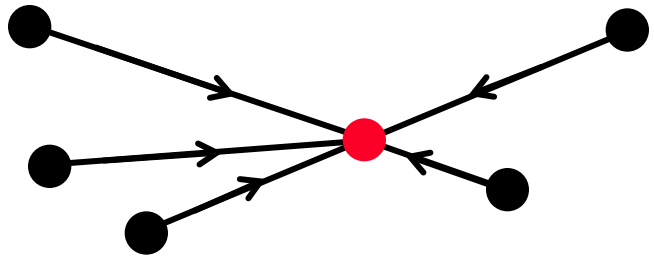




Matrix

- Graph nodes are web pages
- Arrows between nodes are links between web pages
- Matrix entries are links from “column” pages to “row” pages
- The Page Rank comes from algebra on the matrix
- The matrix has about 1,000,000,000,000 rows & columns

Random Surfer Rule



An important page is one that many important pages point to.

- If there are no links out of this page, choose a page at random.
- Otherwise, with probability p (= .85), follow a random link out of this page.
- Or, with probability 1-p (= .15), choose a page at random.

A Page Rank Matrix

- Importance ranking of web pages

- Stationary distribution of a Markov chain

- Power method: matvec and vector arithmetic

- Matlab*P page ranking demo (from SC'03) on a web crawl of mit.edu (170,000 pages)