SPARSE DIRECT METHODS

➤ See recommended reading list. See also the various sparse matrix sites

- Introduction. Goals of sparse techniques.
- Building blocks for sparse direct solvers
- SPD case. Sparse Column Cholesky/
- Elimination Trees - Symbolic factorization
- Supernodes
- Multifrontal Approach

Direct Sparse Matrix Methods

Problem addressed: Linear systems

$$Ax = b$$

➤ We will consider mostly Cholesky –

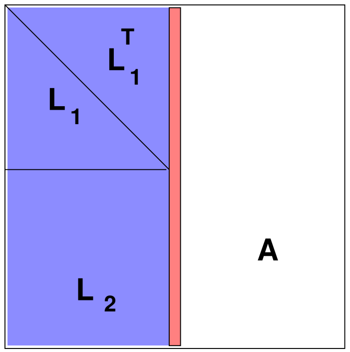
➤ We will consider some implementation details and tricks used to develop efficient solvers

Basic principles:

- Separate computation of structure from rest [symbolic factorization]
- Do as much work as possible statically
- Take advantage of clique formation (supernodes, mass-elimination).

SPARSE COLUMN CHOLESKY

For $j = 1, \ldots, n$ Do:
  $l(j:n, j) = a(j:n, j)$
  For $k = 1, \ldots, j-1$ Do:
    // cmod(k,j):
    $l_{j:n,j} := l_{j:n,j} - l_{j,k} * l_{j:n,k}$
  EndDo
  // cdiv (j) [Scale]
  $l_{j,j} = \sqrt{l_{j,j}}$
  $l_{j+1:n,j} := l_{j+1:n,j} / l_{jj}$
EndDo



Complexity measures

Space:

➤ Determined by $|E^F|$:

$$\boxed{\sum_v \#\text{neighbors of } v \text{ in } G^F}$$

Time:

➤ Number of operations +/- :

$$\boxed{\sum_v \#(\text{neighbors of } v \text{ in } G^F)^2}$$

The four essential stages of a solve

1. Reordering: $A \longrightarrow A := PAP^T$

➤ Preprocessing: uses graph [Min. deg, AMD, Nested Dissection]

2. Symbolic Factorization: Build static data structure.

➤ Exploits 'elimination tree', uses graph only.

➤ Also: 'supernodes'

3. Numerical Factorization: Actual factorization $A = LL^T$

➤ Pattern of L is known. Uses static data structure. Exploits supernodes (blas3)

4. Triangular solves: Solve $Ly = b$ then $L^Tx = y$

Computational Kernels in Sparse Column Cholesky

Two types of computational tasks:

(1) Do column modifications: $comd(k, j), k = 1, .., j - 1$

(2) Scale column j, called $cdiv(j)$.

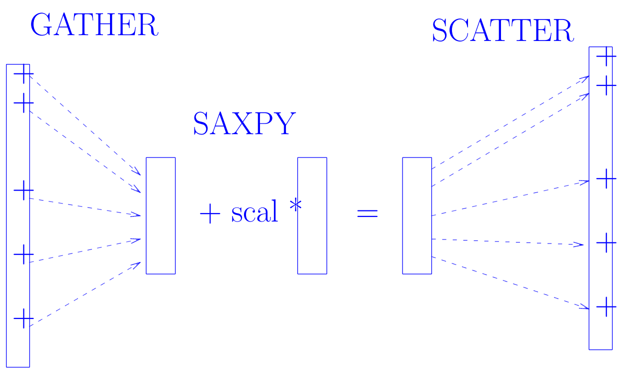
To perform (1):

- first expand column j into a full vector y.
- Then do a succession of sparse SAXPY's.

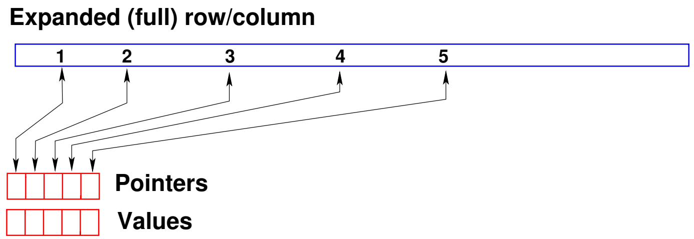
➤ Operation referred to as a 'sparse accumulator'


```
      Do 100 i=1, nz
         y(index(i)) = y(index(i)) + scal * colk(i)
 100  continue
```

Requires (1) gather (2) a saxpy and (3) a scatter



➤ Implementation of a sparse accumulator



➤ To add a nonzero entry - (j, val_j):

- If full_row(j) != zero: modify vals in location full_row(j).
- If full_row(j) == zero. Then create new entry: expand pointer and vals arrays.

ELIMINATION TREES

The notion of elimination tree

➤ Elimination trees are useful in many different ways [theory, symbolic factorization, etc..]

➤ For a matrix whose graph is a tree, parent of column $j < n$ is defined by

$$\boxed{Parent(j) = i, \text{ where } a_{ij} \neq 0 \text{ and } i > j}$$

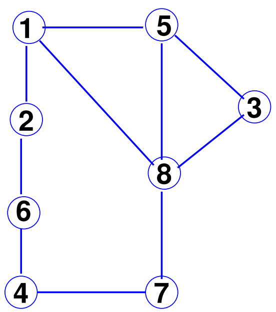
➤ For a general matrix matrix, consider $A = LL^T$, and G^F = 'filled' graph = graph of $L + L^T$. Then

$$\boxed{Parent(j) = \min(i) \; s.t. \; a_{ij} \neq 0 \text{ and } i > j}$$

➤ Defines a tree rooted at column n (Elimintion tree).

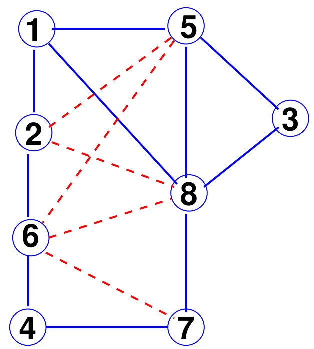
Example: Original matrix and Graph

$$\begin{bmatrix} 1 & \star & & & \star & & & \star \\ \star & 2 & & & & \star & & \\ & & 3 & & \star & & & \star \\ & & & 4 & & \star & \star & \\ \star & & \star & & 5 & & & \star \\ & \star & & \star & & 6 & & \\ & & & \star & & & 7 & \star \\ \star & & \star & & \star & & \star & 8 \end{bmatrix}$$

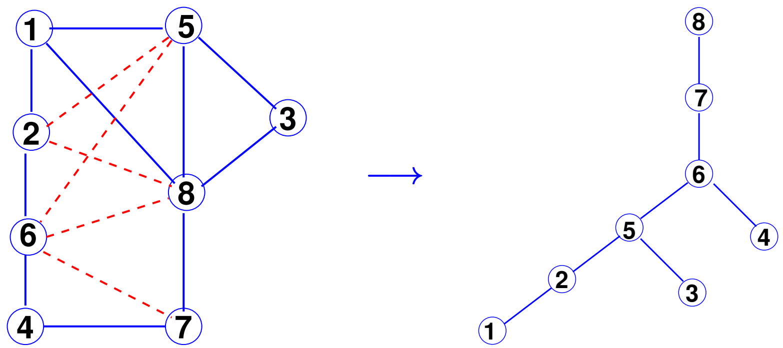


Filled matrix+graph

$$\begin{bmatrix} 1 & \star & & & \star & & & \star \\ \star & 2 & & & \blacksquare & \star & & \blacksquare \\ & & 3 & & \star & & & \star \\ & & & 4 & & \star & \star & \\ \star & \blacksquare & \star & & 5 & \blacksquare & & \star \\ & \star & & \star & \blacksquare & 6 & \blacksquare & \blacksquare \\ & & & \star & & \blacksquare & 7 & \star \\ \star & \blacksquare & \star & & \star & \blacksquare & \star & 8 \end{bmatrix}$$



Corresponding Elimination Tree

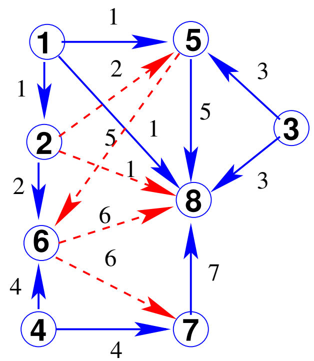


- Parent(i) = 'first nonzero entry in L(i+1:n,i)'
- Parent(i) = min $\{j > i \mid j \in Adj_{G^F}(i)\}$

Where does the elimination tree come from?

- Answer in the form of an excercise.

Consider the elimination steps for the previous example. A directed edge means a row (column) modification. It shows the task dependencies. There are unnecessary dependencies. For example: $\mathbf{1} \to \mathbf{5}$ can be removed because it is subsumed by the path $\mathbf{1} \to \mathbf{2} \to \mathbf{5}$.



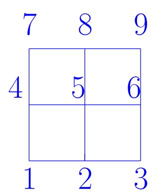
To do: Remove all the redundant dependencies.. What is the result?

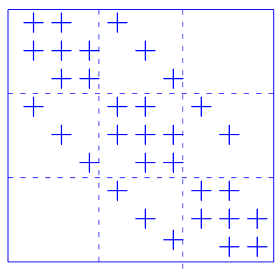
Facts about elimination trees

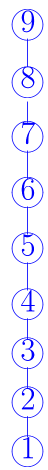
- Elimination Tree defines dependencies between columns.
- The root of a subtree cannot be used as pivot before any of its descendents is processed.
- Elimination tree depends on ordering;
- Can be used to define 'parallel' tasks.
- For parallelism: flat and wide trees → good; thin and tall (e.g. of tridiagonal systems) → Bad.
- For parallel executions, Nested Dissection gives better trees than Minimun Degree ordering.

Elim. tree depends on ordering (Not just the graph)

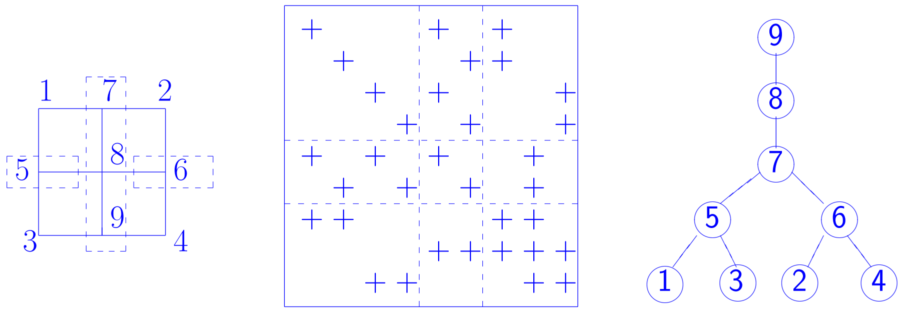
Example: 3×3 grid for 5-point stencil [natural ordering]







➤ Same example with nested dissection ordering

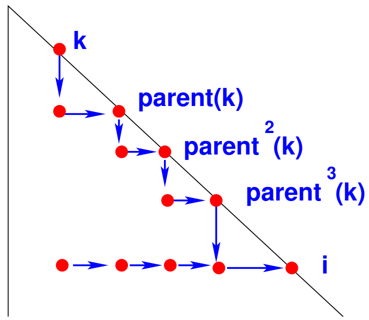


Properties

➤ The elimination tree is a spanning tree of the filled graph [a tree containing all vertices] - obtained by removing edges.

➤ If $l_{ik} \neq 0$ then i is an ancestor of k in the tree

In the previous example: follow the creation of the fill-in (6,8).

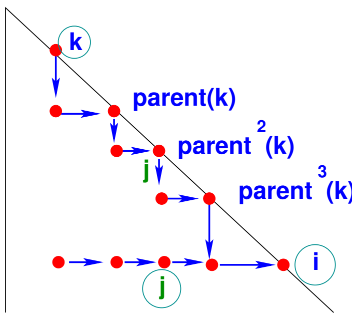


➤ Consequence: no fill-in between branches of the same subtree

Elimination trees and the pattern of L

➤ It is easy to determine the sparsity pattern of L because the pattern of a given column is "inherited" by the ancestors in the tree.

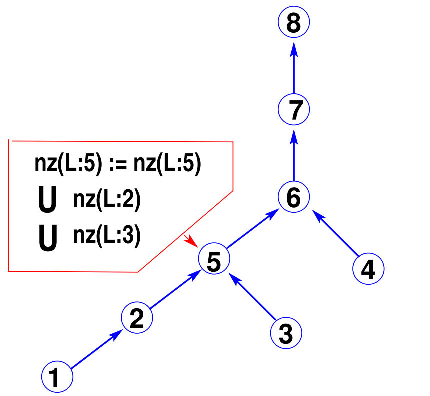
Theorem: For $i > j$, $l_{ij} \neq 0$ iff j is an ancestor of some $k \in Adj_A(i)$ in the elimination tree.



In other words:

$$l_{ij} \neq 0, i > j \ \text{ iff } \ \Big| \ \begin{array}{l} \exists k \in Adj_A(i) s.t. \\ j \rightsquigarrow k \end{array}$$

In theory: To construct the pattern of L, go up the tree and accumulate the patterns of the columns. Initially L has the same pattern as $TRIL(A)$.



➤ However: tree is not available ahead of time

➤ Solution: Parents can be obtained dynamically as the pattern is being built.

➤ This is the basis of symbolic factorization.

Notation :

➤ $nz(X)$ is the pattern of X (matrix or column, or row). A set of pairs (i, j)

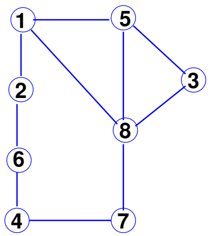
➤ $tril(X)$ = Lower triangular part of pattern [matlab notation] $\{(i, j) \in X \mid i > j\}$

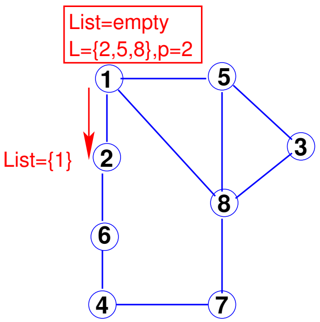
➤ Idea: dynamically create the list of nodes needed to update $L_{:,j}$.

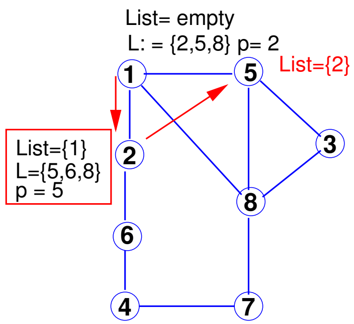
ALGORITHM : 1. *Symbolic factorization*

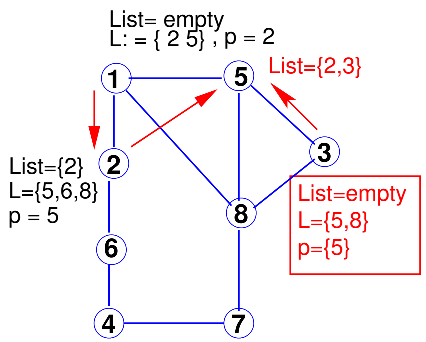
1. *Set:* $nz(L) = tril(nz(A))$,
2. *Set:* $list(j) = \emptyset, j = 1, \cdots, n$
3. *For* $j = 1 : n$
4. *for* $k \in list(j)$ *do*
5. $nz(L_{:,j}) := nz(L_{:,j}) \cup nz(L_{:,k})$
6. *end*
7. $p = \min\{i > j \mid L_{i,j} \neq 0\}$
8. $list(p) := list(p) \cup \{j\}$
9. *End*

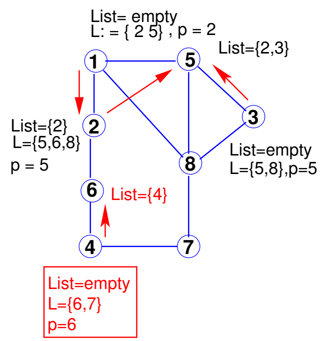
Example: Consider the earlier example:

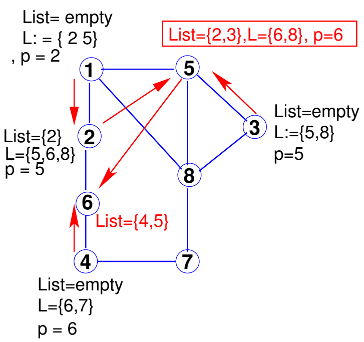


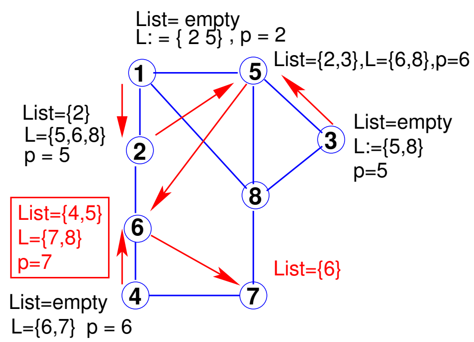









List= empty
L: = { 2 5}
, p = 2
List={2,3},L={6,8}, p=6
List=empty
L:={5,8}
p=5
List={2}
L={5,6,8}
p = 5
List={4,5}
List=empty
L={6,7}
p = 6
1
2
3
4
5
6
7
8


List= empty
L: = { 2 5} , p = 2
List={2,3},L={6,8},p=6
List={2}
L={5,6,8}
p = 5
List=empty
L:={5,8}
p=5
List={4,5}
L={7,8}
p=7
List={6}
List=empty
L={6,7} p = 6
1
2
3
4
5
6
7
8