

Solving Systems of Linear Equations

Y

$$\text{Consider } a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ \vdots \quad \vdots \quad \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m$$

These set of equations can be written as, $Ax = b$ for $x \in \mathbb{R}^n$

where $A = (a_{ik}) = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$

Definitions

① Leading principal submatrix

For $k \leq n$, the (k, k) matrix A_k , formed by deleting rows and columns $k+1, \dots, n$ from A_{nn} is called the leading principal submatrix.

② Nonsingular

A_{nn} is nonsingular if $\det A \neq 0$.

Strongly nonsingular, if all the determinants of the leading principal submatrix are non zero.

③ Triangular matrix, either L (lower triangular), or R (upper t.)

④ Permutation matrix.

Only has one nonzero entry of 1 and $n-1$ zero entries in all rows and columns.

⑤ LR factorization

Every strongly nonsingular A_{nn} matrix can be uniquely

%

factored into a product $A = LR$ where

$$L = \begin{bmatrix} 1 & & & \\ \cdot & 1 & & \\ \cdot & \cdot & 1 & \\ \cdot & \cdot & \cdot & 1 \end{bmatrix} \quad R = \begin{bmatrix} & & & \\ & & & \\ & & & \\ 0 & & & \end{bmatrix}$$

- ⑥ For every nonsingular matrix A there exists an (n,n) permutation matrix P , so that $PA = LR$
- ⑦ Orthogonal matrix.

If $A^T A = I$, then A is called orthogonal.

- ⑧ Householder matrix

If $\vec{v} \neq 0 \in \mathbb{R}^n$ and I_n is the identity matrix,
then the Householder matrix

$$H = I_n - \frac{2}{\|v\|^2} vv^T$$

Where H is a symmetric and orthogonal matrix.

- ⑨ Band matrix

An (n,n) matrix A is called a band matrix, if its entries vanish outside of a band parallel to the main diagonal.

- ⑩ Diagonally dominant

A is diagonally dominant if $|a_{ii}| \geq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}|$ for all $i = 1, \dots, n$.

- ⑪ Positive definite

A symmetric matrix $A = A^T$ is called positive definite if $x^T A x > 0$ for all $x \neq 0, x \in \mathbb{R}^n$

Norms and Condition Numbers

Norms of vectors and matrices provides us with some notion of the 'size' of vectors and matrices.

p-norm is defined as

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

special cases :

1) 1-norm : $\|x\|_1 = \sum_{i=1}^n |x_i|$

also called the Manhattan-norm.

2) 2-norm : $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}$

which corresponds to the usual notion of distance in Euclidean space.

3) ∞ -norm : $\|x\|_\infty = \max_i |x_i| \quad \text{as } p \rightarrow \infty.$

In general, for any vector x in \mathbb{R}^n , we have

1) $\|x\|_1 \geq \|x\|_2 \geq \|x\|_\infty$

2) $\|x\|_1 \leq \sqrt{n} \|x\|_2, \|x\|_2 \leq \sqrt{n} \|x\|_\infty, \text{ and } \|x\|_1 \leq n \|x\|_\infty$

Thus for a given n , any two norms differ by at most a constant, and therefore if one is small, then the others will also be proportionally small.

7

vector norm axioms :

- 1) $\|x\| > 0 \quad \forall x \in \mathbb{R}^n$ with $x \neq 0$
- 2) $\|x\| = 0$ if and only if $x = 0$
- 3) $\|\gamma x\| = |\gamma| \cdot \|x\|$ for any scalar γ .
- 4) $\|x+y\| \leq \|x\| + \|y\| \quad \rightarrow$ triangle inequality. $\forall x, y \in \mathbb{R}^n$

Matrix Norms

- A matrix norm is compatible with a vector norm if for every matrix A and every vector x the inequality $\|Ax\| \leq \|A\| \|x\|$ holds

$$\|A\|_1 = \max_{1 \leq k \leq n} \sum_{i=1}^n |a_{ik}| \quad \rightarrow \text{column sum norm.}$$

$$\|A\|_\infty = \max_{1 \leq k \leq n} \sum_{i=1}^n |a_{ik}| \quad \rightarrow \text{row sum norm.}$$

$$\|A\|_2 = \sqrt{\sum_{i,k=1}^n |a_{ik}|^2} \quad \rightarrow \text{Frobenius norm.}$$

matrix norm axioms :

- 1) $\|A\| > 0$ if $A \neq 0$
- 2) $\|\gamma A\| = |\gamma| \cdot \|A\|$ for any scalar γ
- 3) $\|A+B\| \leq \|A\| + \|B\|$
- 4) $\|AB\| \leq \|A\| \cdot \|B\|$
- 5) $\|Ax\| \leq \|A\| \cdot \|x\|$ for any vector x .

Errors and the condition number

To approximate the error of an approximate solution, $x^{(0)}$ of the true solution of $Ax = b$, we may use

$$r^{(0)} = b - Ax^{(0)}, \text{ where } r^{(0)} \text{ is the residual vector}$$

- If $x^{(0)}$ is close to the solution, then $r^{(0)}$ must be a very small value.

- However if $|r^{(0)}| < \epsilon$, where ϵ is a small number need not imply that $x^{(0)}$ is a good approximation for x because A may not be well-conditioned.

- Therefore the residual vector only is not a good measure of the quality of $x^{(0)}$.

The following measures of the linear system can also be used to determine how well-conditioned the system is.

i) Hadamard Condition Number

$$K_H(A) = \frac{|\det A|}{\alpha_1 \alpha_2 \dots \alpha_n} \quad \text{with } \alpha_i = \sqrt{\alpha_{11}^2 + \dots + \alpha_{ii}^2} \quad \text{for } i=1, \dots, n.$$

A matrix A is considered ill-conditioned if $K_H(A) \ll 1$.

Current experience has shown that:

$$K_H(A) < 0.01 \Rightarrow \text{ill-conditioned}$$

$$0.01 \leq K_H(A) \leq 0.1 \Rightarrow \text{no assertion can be made}$$

$$K_H(A) > 0.1 \Rightarrow \text{well-conditioned}$$

2) Condition number

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

If A is singular, then $\text{cond}(A) = \infty$.

$$\|A\| \cdot \|A^{-1}\| = \left(\max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \right) \cdot \left(\min_{x \neq 0} \frac{\|Ax\|}{\|x\|} \right)^{-1}$$

- The condition number of a matrix bounds the ratio of the relative change in the solution of a linear system to a given relative change in the input data.
- $\text{cond}(A)$ measures how close a matrix is to being singular, and it is a better measure than $\det(A)$, since the value of $\det(A)$ gives no measure of how close A is to being singular.
- properties of the condition number are:
 - 1) For any A , $\text{cond}(A) \geq 1$.
 - 2) For any A and scalar γ , $\text{cond}(\gamma A) = \text{cond}(A)$
 - 3) For any diagonal matrix $D = \text{diag}(d_i)$,
$$\text{cond}(D) = \frac{\max |d_i|}{\min |d_i|}.$$
- computing the condition number is an expensive process, since A^{-1} is needed; therefore we need methods to estimate the condition number.

3) Spectral condition Number

$$\mu(A) = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|} \quad \text{where } \lambda_i \text{ are the eigenvalues of the matrix}$$

Estimating condition numbers

1) Hadamard condition number.

If LR factorization is known, then the $\det(A)$ can be computed and $K_H(A)$ only requires $O(n^2)$ operations where n is the size of the matrix.

2) $\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$ requires $O\left(\frac{4}{3}n^3\right)$ operations and hence its an order of magnitude more expensive than the Hadamard condition number. Therefore we must be able to estimate the value.

Ⓐ Forsythe and Moler method.

Step 1 : Calculate $r^{(0)} = b - Ax^{(0)}$ in double prec. (DP)

Step 2 : Solve $Az^{(1)} = r^{(0)}$ via factorization method of A.

Step 3 : Calculate $\text{cond}(A) \approx \frac{\|z^{(1)}\|}{\|x^{(0)}\|} \cdot \frac{1}{\text{EPS}}$

where EPS is the machine constant which is defined as $5 \cdot 10^{-t}$, where t is the number of mantissas.

$$\text{cond} \approx \frac{\|A^{-1}b - x^{(0)}\|}{\|x^{(0)}\|} \cdot \frac{1}{\text{EPS}}$$

Improving the Condition Number

8/

1) Row scaling

We multiply each row of A by a constant such that all the row vectors of the matrix have the same norm. The system can be written as,

$$D_1 A x = D_1 b \quad \text{where } D_1 \text{ is a non-singular and diagonal matrix.}$$

2) column scaling

similar to row scaling, here we multiply each column by a constant.

$$D_2 y = x$$

$$A D_2 y = b$$

3) Joint row and column scaling

$$D_1 A D_2 y = D_1 b$$

4) Row replacement.

Replacing certain rows of A with linear combinations of the given rows.

Iterative Refinement

7

For a linear system $Ax = b$, where $x^{(0)}$ is an approximate solution and it contains rounding errors, then the following iterative procedure can improve the solution to the system.

Suppose $r^{(0)} = b - Ax^{(0)}$, then find the solution to

$$Az^{(1)} = b - Ax^{(0)} = r^{(0)} \quad \text{and set } x^{(1)} = x^{(0)} + z^{(1)}$$

Since A , b , and $x^{(0)}$ are known, then $r^{(0)}$ can be easily computed.

Then compute $z^{(1)}$ by solving $Az^{(1)} = r^{(0)}$.

We can then find an approximate solution to $x^{(0)}$.

$$x^{(1)} = x^{(0)} + z^{(1)}$$

We then repeat the process using the following algorithm.

$$Az^{(v+1)} = b - Ax^{(v)} = r^{(v)} \quad \text{for } v = 0, 1, 2, \dots$$

This process continues until

$$\frac{\|z^{(v+1)}\|_\infty}{\|x^{(v)}\|_\infty} < \epsilon \text{ for } \epsilon >$$

Accuracy of Solutions

10/

The most obvious way to check the validity of a solution is to substitute it into the equation to see how closely the two sides match.

Let \hat{x} be an approximate solution, then for a linear system $Ax=b$, we can define the residual vector as

$$r = b - A\hat{x}$$

If A is nonsingular, then the error $\|\hat{x} - x\| = 0$ iff $\|r\| = \epsilon$

\uparrow
if and only if

If \hat{x} exactly satisfies

$$(A+E)\hat{x} = b$$

$$E\hat{x} = b - A\hat{x}$$

$$\text{then } \|r\| = \|b - A\hat{x}\| = \|E\hat{x}\| \leq \|E\| \cdot \|\hat{x}\|$$

thus we have

$$\frac{\|r\|}{\|A\|} \leq \frac{\|E\| \cdot \|\hat{x}\|}{\|A\|}$$

$$\frac{\|r\|}{\|A\| \cdot \|\hat{x}\|} \leq \frac{\|E\|}{\|A\|}$$

- The inequality relates the 'relative residual' to the relative change in the matrix.

- If the relative residual is large, then this implies that the backward error in the matrix is large.

To bound $\|E\|$, Wilkinson showed that for LU factorization by Gaussian Elimination, a bound of the form .

$$\frac{\|E\|}{\|A\|} \leq p n \cdot \text{EPS} , \text{ where } p \text{ is the growth factor.}$$

In practice, with pivoting there is little or no growth, so that

$$\frac{\|E\|}{\|A\|} \propto n \cdot \text{EPS}.$$

Example: using 3-digit decimal arithmetic to solve

$$\begin{bmatrix} 0.641 & 0.242 \\ 0.321 & 0.121 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.883 \\ 0.442 \end{bmatrix}$$

Gaussian Elimination with partial pivoting yields,

$$\begin{bmatrix} 0.641 & 0.242 \\ 0 & 0.000242 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.883 \\ -0.000383 \end{bmatrix}$$

and back-substitution gives $x = \begin{bmatrix} 0.782 \\ 1.58 \end{bmatrix}$

$$\text{The residual is } r = b - Ax = \begin{bmatrix} -0.000622 \\ -0.000202 \end{bmatrix}$$

but the exact solution is $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

The cause is due to the fact that the matrix is nearly singular and due to rounding error since only 3-digits are used.

Estimating Accuracy

12/

Let x be the solution to the nonsingular linear system $Ax = b$, and \hat{x} be the solution to the system $A\hat{x} = b + \Delta b$, where Δb is a perturbation of the right-hand-side.

If we define $\Delta x = \hat{x} - x$, then

$$b + \Delta b = A\hat{x} = A(x + \Delta x) = Ax + A\Delta x$$

$$\text{since } Ax = b, \text{ then } A\Delta x = \Delta b \text{ or } \Delta x = A^{-1}\Delta b$$

$$\text{so, } \|\Delta x\| \leq \|A\| \cdot \|x\| \quad \text{and} \quad \|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\|$$

$$\text{Then } \frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A\| \cdot \|A^{-1}\| \cdot \|\Delta b\|}{\|b\|}$$

From the definition of the condition number, $\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$

$$\text{so, } \frac{\|\Delta x\|}{\|x\|} \leq \text{cond}(A) \cdot \frac{\|\Delta b\|}{\|b\|}$$

Therefore for a given relative change in the right-hand-side, the condition number of A determines the relative change to the solution.

Similarly, if we change matrix A by E , then

$$(A+E)\hat{x} = b$$

$$\text{Then } x = A^{-1}b$$

$$x - \hat{x} = A^{-1}b - \hat{x} = A^{-1}b - A^{-1}A\hat{x}$$

$$\Delta x = A^{-1}(b - A\hat{x})$$

From $(A+E)\hat{x} = b$

$$A\hat{x} + E\hat{x} = b$$

$$E\hat{x} = b - A\hat{x}$$

Then $\Delta x = A^{-1}E\hat{x}$

$$\|\Delta x\| \leq \|A^{-1}\| \cdot \|E\| \cdot \|\hat{x}\|$$

$$\frac{\|\Delta x\|}{\|\hat{x}\|} \leq \|A^{-1}\| \cdot \|E\| \cdot \frac{\|A\|}{\|A\|} = \text{cond}(A) \cdot \frac{\|E\|}{\|A\|}$$

The same can be said.

If both the left-and-right-hand-sides are changed then we can use the following approach.

We define $A(t) = A + tE$, where t is a real-valued parameter and $b(t) = b + t\Delta b$.

For the linear system, $A(t)x(t) = b(t)$

$$A'(t)x(t) + A(t)x'(t) = b'(t)$$

$$A(t)x'(t) = b'(t) - A'(t)x(t)$$

$$x'(t) = A^{-1}(t)b'(t) - A^{-1}(t)A'(t)x(t)$$

$$\|x'(t)\| \leq \|A^{-1}\| \|b'\| + \|A^{-1}\| \|A'\| \|x\|$$

$$\frac{\|x'\|}{\|x\|} \leq \|A^{-1}\| \cdot \frac{\|b'\|}{\|x\|} \cdot \frac{\|A\|}{\|A\|} + \|A^{-1}\| \cdot \frac{\|A'\|}{\|A\|} \cdot \|A'\| \cdot \frac{\|x\|}{\|x\|}$$

$$\frac{\|x'\|}{\|x\|} \leq \text{cond}(A) \cdot \frac{\|b'\|}{\|b\|} + \text{cond}(A) \cdot \frac{\|A'\|}{\|A\|}$$

$$\frac{\|x'\|}{\|x\|} \leq \text{cond}(A) \left(\frac{\|b'\|}{\|b\|} + \frac{\|A'\|}{\|A\|} \right)$$

Again we see that the relative change in the solution is bounded by the condition number times the sum of the relative changes to b and A .

If A and b are accurate to machine precision,

then

$$\frac{\|x'\|}{\|x\|} \leq \text{cond}(A) \cdot \text{EPS}$$

$$\text{or } \frac{\|x'\|}{\|x\|} \approx \text{cond}(A) \cdot \text{EPS}$$

The computed solution losses about $\log_{10}(\text{cond}(A))$ decimal digits of accuracy relative to the accuracy of the input.

Factorization Principle

- Direct methods primarily rely on factoring the matrix A into a product of LR, where L is a lower triangular matrix and R and upper triangular matrix.

- The steps are as presented below,

1) Find a lower triangular matrix L so that $L^{-1}A = R$ is upper triangular

2) Solve $C = L^{-1}b$

3) Solve $Rx = C$ using back substitution.

- If L^{-1} cannot be found, due to instability of the method, then the standard procedure is to exchange rows in A and replace it with PA, where P is a permutation matrix.

Gauß Algorithm

The idea behind the algorithm with column pivot search is to transform a linear system of the form presented on page 1 of these notes, where $m=n$, into a triangular system

$$r_{11}x_1 + r_{12}x_2 + \dots + r_{1n}x_n = c_1$$

$$r_{22}x_2 + \dots + r_{2n}x_n = c_2$$

$$\vdots \quad \vdots \quad \vdots$$

$$r_{nn}x_n = c_n$$

by eliminating the lower triangular coef. of A.

The algorithm achieves this by performing row exchanges to bring the largest element of a column into its diagonal position.

The following is a simplified procedure for the Gaussian algorithm.

- 1) Perform row exchanges to ensure that the element of the largest size appears in the (1,1) position

$$A^{(0)} = PA \quad \text{and} \quad b^{(0)} = Pb \\ = (a_{ik}^{(0)}) \quad \quad \quad = (b_i^{(0)})$$

Thus the equation can be written as

$$\sum_{k=1}^n a_{ik}^{(0)} x_k = b_i^{(0)} \quad \text{for } i=1, \dots, n.$$

- 2) Eliminate x_1 from equations $i=2, \dots, n$ by multiplying the first equation by $\frac{-a_{ii}^{(0)}}{a_{11}^{(0)}}$ and add it to the i^{th} equation for all $i=2, \dots, n$.

$$a_{11}^{(0)} x_1 + a_{12}^{(0)} x_2 + \dots + a_{1n}^{(0)} x_n = b_1^{(0)} \\ \tilde{a}_{22}^{(1)} x_2 + \dots + \tilde{a}_{2n}^{(1)} x_n = \tilde{b}_2^{(1)} \\ \vdots \\ \tilde{a}_{nn}^{(1)} x_n = \tilde{b}_n^{(1)}$$

where

$$\tilde{a}_{ik}^{(1)} = \begin{cases} 0 & \text{for } k=1 \text{ and } i=2, \dots, n \\ a_{ik}^{(0)} - a_{1k}^{(0)} \frac{a_{11}^{(0)}}{a_{11}^{(0)}} & \text{otherwise} \end{cases}$$

$$\tilde{b}_i^{(1)} = b_i^{(0)} - b_1^{(0)} \frac{a_{11}^{(0)}}{a_{11}^{(0)}} \quad \text{for } i=2, \dots, n.$$

- Steps 1 and 2 are repeated until the triangular form on page 15 is realized. Once the triangular form or matrix R is achieved, then a backsubstitution is performed to compute x_i recursively.
- The operation count for this algorithm is approximately $\frac{n^3}{3}$.
- If during the procedure, a nonzero pivot element does not exist, then $\det(A) = 0$ and A is numerically a singular matrix.

Matrix Inversion with Gauß Algorithm

If $Ax_i = e_i$ where e_i denotes the i th unit vector, then we can form $AX = I_n$ where I_n is the identity matrix, and X contains the vectors x_i in each column. From the definition,

$$X = A^{-1}.$$

Other Pertaining Information on the Gauß Algorithm

- 1) It also goes by Gaussian Elimination / Reduction or LU factorization decomposition. (see below).
- 2) The pivot steps and row multiplications will eventually form the L matrix.

Gauß in LU factorization form

15/

If M represents the elimination matrix or Gauss transformation, then the linear system can be written as,

$$MA = M_{n-1} \dots M_1 A x = M_{n-1} \dots M_1 b = Mb$$

where $U = MA$ where U is the upper triangular matrix.

The lower triangular matrix L , can then be computed by

$$L = M^{-1} = (M_{n-1} \dots M_1)^{-1} = M_1^{-1} \dots M_{n-1}^{-1} = L_1 \dots L_{n-1}$$

Example

Given $\begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix}$

$$M_2 M_1 A = M_2 M_1 b$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} x = M_2 M_1 b$$

$$U = M_2 M_1 A = \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix}$$

$$L = L_1 L_2 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{bmatrix} = LU$$

Pivoting

To ensure that the error is minimized and to reduce the amplification of rounding errors, we want the pivot to be the largest value, and thus the multipliers would be closer to one.

The equations are transformed to

$$MA = U \quad \text{where} \quad M = M_{n-1} P_{n-1} \dots M_1 P_1, \quad \text{where } P \text{ is the permutation matrix.}$$

Implementing the Gaussian Elimination Algorithm.

The algorithm is generally represented as a triple-nested loop.

```

for —
  for —
    for —
       $a_{ij} = a_{ij} - (a_{ik}/a_{kk})a_{kj}$ 
      end
    end
  end

```

- To improve efficiency some of the operations are moved outside of the loop.
- The basic algorithm has to be modified to take advantage of the different memory access patterns (row-wise) or (column-wise). Algorithm must also take advantage of the architectural features of the computer such as (cache, vectorization, multiprocessors)
- To reduce storage L occupies the lower triangular portion of A and R or U occupies the upper triangular.
- To minimize data exchange, row exchanges and pivoting are

accomplished through an additional integer vector to keep track of the new row order instead of physically moving the original locations in memory as well as in the matrix. 20/

Gauß-Jordan Method

In this method, the matrix is reduced to a diagonal form rather than just an upper triangular.

Special types of Linear Systems

1) Symmetric Positive Definite Systems

Cholesky Decomposition

- If for a linear system $Ax = b$, the matrix A is symmetric and positive definite ($x^T Ax > 0$) for all $x \neq 0$, then we can use the Cholesky decomposition.

- Algorithm:

step 1) Factor $A = LL^T$, where $L = L^T$

step 2) Update $c = L^{-1}b$

step 3) Backsubstitute $L^T x = c$

- For large n it uses only half as many operations and storage as the Gauß algorithm.
- No pivoting is required for numerical stability.

2) Linear systems with Tridiagonal Matrices (or Five-diagonal)

- simple algorithms such as the Thomas Algo. to factorize A into an LU form for such matrices.

3) Band matrices

- Are matrices whose elements are zero outside a band parallel to the diagonal.
- The tridiagonal and five-diagonal are special cases.

- simple algorithms exists that transform these matrices into an LU form.

21

4) Sparse matrices.

- most entries of A are zero.

- The algorithms are generally more ~~expensive~~ complicated than for dense algorithms due to the need to efficiently handle 'fill-ins' in the L and U factors. 'Fill-in's are those entries that change from an initial zero to a non-zero entry during the execution of the algorithm.

- A typical sparse matrix solver follows these steps.

1) An ordering step that reorders the rows and columns such that the matrix has a special structure.

2) An analysis step to create a suitable data structure for the factors.

3) Compute L and U

4) Perform forward and backward substitution to compute \mathbf{x} .

- Cuthill - McKee Algorithm. for Sparse symmetric matrices.

Step 1) use cuthill-mcKee algorithm to transform the sparse symmetric matrix to a band matrix with the lowest band possible.

Step 2) Use Rosen algorithm to further reduce the band width.

Step 3) use specially adapted Gauss or Cholesky algorithms to factor the matrix into L, U form.

Operations Count

22

Procedure

Gauß (Gauss)

Cholesky
Decomposition

Gauß-Jordan

Algorithms for
tridiagonal matrices

Algorithms for
five-diagonal
matrices

Iterative
methods

Operations count

$$\frac{n^3}{3} + n^2 - \frac{n}{3}$$

$$\frac{n^3}{6} + O(n^2)$$

$$\frac{n^3}{2} + n^2 + \frac{n}{2}$$

$$5n - 4$$

$$11n - 16$$

$$2n^2 - 2n$$