III Solution of linear systems of equations

1 - Introduction

1.1 - Preliminaries

Given the linear system of equations

$$x + y = 5 \quad (E_1)$$
$$x - y = 1 \quad (E_2)$$

find x and y.

(*) Please note :

— The systems

$$xy + 2 = 3 \qquad \text{(involves products between the variables)}$$
$$x - y = 0$$

or

$$x^2 + y^2 = 25 \qquad \text{(has powers involving the variables)}$$
$$x - y = 1$$

or

$$\sin x + y = 0 \qquad \text{(has transcendental functions)}$$
$$e^y - x = 1$$

are nonlinear

* In order to solve a linear system of equations, we perform a series of operations so that it is transformed in a more easily solvable system

1. Equation E_i can be multiplied by a const. λ, $\lambda \neq 0$, with the resulting equation used in place of E_i : $E_i \to \lambda E_i$ (in some books you find the opposite: $(\lambda E_i) \to (E_i)$)
2. Equation E_j can be multiplied by any constant λ and added to the equation E_i, with the resulting equation used in place of E_i
   $E_i \to E_i + \lambda E_j$
3. Equations E_i and E_j can be transposed in order:
   $E_i \leftrightarrow E_j$

<u>Example</u>:

$$x + y = 5 \quad (E_1)$$
$$x - y = 1 \quad (E_2)$$

$E_1 \to E_1 + E_2 \Rightarrow 2x = 6 \quad (\breve{E}_1)$

$\breve{E}_1 \to \breve{E}_1/2 \Rightarrow x = 3$

$E_2 \to E_1 - E_2 \Rightarrow 2y = 4 \quad (\breve{E}_2)$

$\breve{E}_2 \to \breve{E}_2/2 \Rightarrow y = 2$

(*) <u>Our goal</u>: Make the computer solve linear systems of equations using such operations

<u>General form of a linear system</u>:

$$\begin{aligned}
E_1 &: a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\
E_2 &: a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\
&\vdots \\
E_n &: a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n
\end{aligned}$$

Variables: $x_1, \cdots x_n$ (x_i)

Constants: $a_{11} \cdots a_{nn}$ (a_{ij})

$b_1 \cdots b_n$ (b_j)

(*) Such a system can be written in a more compact way, using a matrix.

1.2- Matrices and vectors

(*) Definition: An $n \times m$ matrix is a rectangular array of elements with n rows and m columns in which not only is the value of an element important, but also its position in the array.

(*) Notation: Matrix: Capital letter: A, M

Matrix element: a_{ij}

$$\begin{array}{cc} a_{i j} & \\ \text{rows} & \text{column} \end{array}$$

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & \vdots & & a_{2m} \\ & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

Example: $A = \begin{bmatrix} 2 & -1 & 7 \\ 3 & 1 & 0 \end{bmatrix}$

is a 2×3 matrix with $a_{11} = 1$, $a_{12} = -1$, $a_{13} = 7$, $a_{21} = 3$, $a_{22} = 1$, and $a_{23} = 0$.

(*) Vectors: The $1 \times n$ matrix

$$A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}$$

is called an n-dimensional column vector and the $1 \times n$ matrix

$$A = [a_{11} \; a_{12} \cdots a_{1n}]$$

is a row vector.

(*) Normally the unnecessary subscripts are omitted and a boldface lower case letter is used for notation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \mathbf{y} = [y_1 \; y_2 \cdots y_n]$$

column vector

row vector

The linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n, \end{aligned}$$

can be represented by a matrix

Procedure :

we construct $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

and combine these matrices to form the augmented matrix

$$\tilde{A} = [A, \mathbf{b}] = \left[\begin{array}{cccc:c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ \vdots & \vdots & & \vdots & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right]$$

↳ separates the coefficients of the unknowns from the values at the RHS of the eqs.

Example: $x+y=5$, $x-y=1$ can be written as $\tilde{A} = \begin{bmatrix} 1 & 1 & \vdots & 5 \\ 1 & -1 & \vdots & 1 \end{bmatrix}$

2. Gaussian elimination with backward substitution

2.1- Problem

We wish to write the matrix of the system as

$$\tilde{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & \vdots & b_1 \\ 0 & a_{22} & \cdots & \vdots & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{nn} \vdots & b_n \end{pmatrix}$$ (upper triangular matrix)

Then the system can be solved iteratively starting from the last equation:

System : $a_{11}x_1 + a_{12}x_2 + \cdots = b_1$

$a_{22}x_2 + \cdots = b_2$

$a_{nn}x_n = b_n$

Example: If $\tilde{A} = \begin{pmatrix} 1 & 1 & \vdots & 5 \\ 1 & -1 & \vdots & 1 \end{pmatrix}$ is written as

$$\tilde{A} = \begin{pmatrix} 1 & 1 & \vdots & 5 \\ 0 & 1 & \vdots & 2 \end{pmatrix}$$

then $y=2$ and by substitution $x=3$

2.2. Procedure

How to do it?

(a) How to obtain an upper-triangular matrix:

- Step 1: Substitute E_j by $E_j - \frac{a_{j1}}{a_{11}} E_1$

  for $j = 2, 3, \cdots n$

  to eliminate the coefficient of x_1

- Step 2: Substitute E_j by $E_j - \frac{a_{j2}}{a_{22}} E_2$

  for $j = 3, \cdots, n$

  to eliminate the coefficient of x_2

  ⋮

- Step $n-1$:
  Substitute E_j by $E_j - \frac{a_{j\,n-1}}{a_{n-1,n-1}} E_n$ for $j = n$

  to eliminate the coefficient of x_{n-1}

(2 nested do loops)

(*) Please note: this procedure is called "Gaussian elimination"

(This is exactly what we have done for our 2×3 matrix when we subtracted ~~rows~~ 1 and 2)

(b) How to find the x_i's $(i = 1, \dots, n)$?

Solving the n-th equation for x_n gives

$$x_n = \frac{b_n}{a_{nn}} \qquad (*)$$

Inserting this in the $(n-1)$st equation:

$$x_{n-1} = \frac{b_{n-1} - a_{n-1,n}\, x_n}{a_{n-1,n-1}}$$

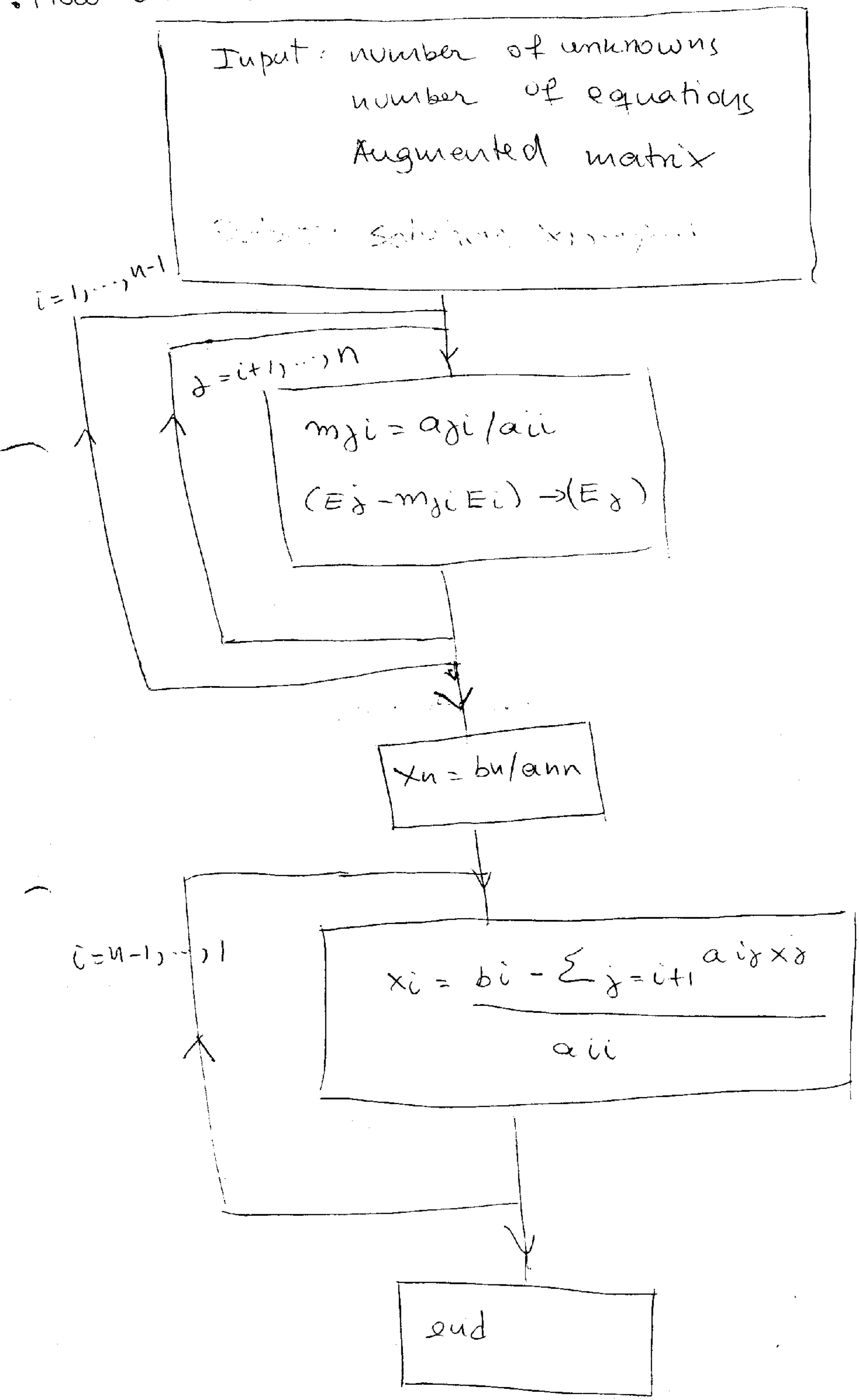
, where x_n is given by (*)

Continuing this process,

$$x_i = \frac{b_i - \sum_{j=i+1}^{n} a_{ij} x_j}{a_{ii}} \qquad \text{(1 do loop)}$$

(*) This procedure is called "backward substitution"

• Flow chart:



* **Problem**: this procedure will break down if any of the elements $a_{11}^{(1)}, a_{22}^{(2)}, a_{33}^{(3)} \cdots, a_{nn}^{(n)}$ is zero

(notation : $a_{ii}^{(i)} \to$ ii^{th} element at the i^{th} iteration)

Reason : either the step

$$\left(E_i - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} E_k\right) \to (E_i)$$

or the backward substitution can not be accomplished.

$\Rightarrow$ The technique for finding a solution must be altered.

Example :

Let us suppose we have a linear system with the augmented matrix

$$\tilde{A} = \tilde{A}^{(1)} = \left[\begin{array}{cccc:c} 1 & -1 & 2 & -1 & -8 \\ 2 & -2 & 3 & -3 & -20 \\ 1 & 1 & 1 & 0 & -2 \\ 1 & -1 & 4 & 3 & 4 \end{array}\right]$$

We start to perform the Gaussian elimination:

$$(E_2 - \underset{\frac{a_{21}}{a_{11}}}{\underset{\shortparallel}{2}} E_1) \to (E_2)$$

$$(E_3 - \underset{\frac{a_{31}}{a_{11}}}{\underset{\shortparallel}{}} E_1) \to (E_3)$$

$(E_4 - E_1) \to (E_4)$

$\quad '' \frac{a_{41}}{a_{11}}$

This yields :

$$\ddot{A}^{(2)} = \left[\begin{array}{cccc|c} 1 & -1 & 2 & -1 & -8 \\ 0 & \boxed{0} & -1 & -1 & -4 \\ 0 & 2 & -1 & 1 & 6 \\ 0 & 0 & 2 & 4 & 12 \end{array}\right]$$

Problem: $a_{22} = 0 \Rightarrow$ if one continues the Gaussian elimination one will end up dividing by zero!

Solution: Swap E_2 with a row for which $a_{22} \neq 0 \Rightarrow (E_2) \leftrightarrow (E_3)$

This allows one to continue with the procedure (one needs an "if statement" in the flow chart)

(*) Please note :

- a_{22} is in this case called "the pivot"
- Swapping the rows and/or columns of a matrix is known as "pivoting".
  - Partial pivoting : swapping only the rows
  - " " the columns
  - Full pivoting : swapping the rows and columns

2.3- Pivoting strategies

2.3.1 - To avoid $a_{kk}^{(k)} = 0$:

Step 1: Check if $a_{kk}^{(k)} = 0$

Step 2: If $a_{kk}^{(k)} = 0$ then:

- find a row p, $p > k$ with $a_{pk}^{(k)} \neq 0$
- switch rows k and p

(need a do loop and an "if" statement)

2.3.2 - To reduce error:

If $a_{kk}^{(k)} \ll a_{jk}^{(k)}$ then $m_{jk} = \dfrac{a_{jk}^{(k)}}{a_{kk}^{(k)}} \gg 1$

Hence, when performing the backward substitut

for $x_n = \left(b_n^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)}\right) / a_{kk}^{(k)}$ the

error will be increased (one is dividing by a very small number)

(*) Procedure to avoid this problem: Move the entry of largest magnitude to the main diagonal and use it to eliminate the remaining entries.

Step 1: Locate the row p in which the element with the largest absolute value lies

Step 2: Switch row k with row p, if $p > k$.

2.4- Problems

- Row degeneracy: one or more rows is a linear combination of the others
- Column degeneracy: all equations contain certain variables only in a certain combination

In this case, the matrix is singular

Examples:

System.

$$x + y = 5$$
$$2x + 2y = 10$$

Matrix

$$A = \begin{pmatrix} 1 & 1 & \vdots & 5 \\ 2 & 2 & \vdots & 10 \end{pmatrix}$$

System

$$x + y + z = 5$$
$$x - y + 2z = 1$$
$$2x + 3z = 6$$

$(E_3 = E_2 + E_1)$

Matrix

$$B = \begin{pmatrix} 1 & 1 & 1 & \vdots & 5 \\ 1 & -1 & 2 & \vdots & 1 \\ 2 & 0 & 3 & \vdots & 6 \end{pmatrix}$$

(*) If the matrices are nearly singular the procedure may fail.

(Nearly singular matrices $\Rightarrow$ $\det(M) \simeq 0$)

Example: Consider the system

$$x + 2y - 2.00 = 0 \quad (*)$$

$$2x + 3y - 3.40 = 0 \quad (**)$$

(solution: $x = 0.8$, $y = 0.6$)

Inserting $x_0 = 1.00$ and $y_0 = 0.48$ gives

$(*) \Rightarrow -0.04$
$(**) \Rightarrow +0.04$ > almost zero

(*) Examples of matrices/vectors (MatLab)

- Matrix $A = \begin{pmatrix} 1 & 2 & 1 & 4 \\ 2 & 0 & 4 & 3 \\ 4 & 2 & 2 & 1 \\ -3 & 1 & 3 & 2 \end{pmatrix}$

$$A = [\underbrace{1\ 2\ 1\ 4}_{\text{Row}};\ \underbrace{2\ 0\ 4\ 3}_{\text{Row}};\ \underbrace{4\ 2\ 2\ 1}_{\text{Row}};\ \underbrace{-3\ 1\ 3\ 2}_{\text{Row}}]$$

- Vector:

$B = [13\ \ 28\ \ 20\ \ 6]$

- Augmented matrix:

$Aug = [A\ \ B]$

- Matrix element $a_{ij} = A(i,j)$

Example: $A(1,2) = 2$

- Vector element $b_i = B(i)$

Example: $B(3) = 20$

In particular, $X = A\backslash B$ is the solution of $AX = B$ by Gaussian elimination with partial pivoting.

3. LU Decomposition (or triangular factorization)

(*) One uses the fact that ~~most~~ triangular matrices are easy to solve, ~~and~~ decompose an arbitrary matrix in such a form.

Definition : A nonsingular matrix A has a triangular factorization if it can be expressed as the product of a lower triangular matrix L and an upper triangular matrix U

$$A = LU \qquad (*)$$

Example : 4×4 matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & m_{32} & 1 & 0 \\ m_{41} & m_{42} & m_{43} & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix}}_{U}$$

(*) **Assumption** :

The coefficient matrix A of

$Ax = B$ can be written as (*)

After finding L and U, the solution X is computed by defining $Y = UX$ and solving 2 systems:

- Step 1 : Solve $LY = B$ for Y using forward substitut
- Step 2 : Solve $UX = Y$ for X using backward substitution

<u>Theorem</u>: If Gaussian elimination can be performed on the linear system $A\vec{x} = \vec{b}$ without row interchanges, then the matrix A can be factored into

$$A = LU$$

where $m_{ji} = \dfrac{a_{ji}^{(i)}}{a_{ii}^{(i)}}$

$$U = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn}^{(n)} \end{bmatrix} \text{ and } L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ m_{21} & 1 & \cdots & \vdots \\ \vdots & & \ddots & 0 \\ m_{n1} & \cdots & m_{n,n-1} & 1 \end{bmatrix}$$

(*) Why does it work?

- First step in the Gaussian elimination procedure:

$$(E_j - m_{j1} E_1) \to (E_j), \quad (j = 2, n) \quad \text{where } m_{j,1} = \frac{a_{j1}^{(1)}}{a_{11}^{(1)}}$$

This is equivalent to multiplying

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \text{ by } M^{(1)} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -m_{21} & 1 & \ddots & \vdots \\ \vdots & 0 & \ddots & 0 \\ -m_{n1} & 0 & \cdots & 1 \end{bmatrix}}_{\text{First gaussian transformation matrix}}$$

Check :

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -m_{21} & 1 & & 0 \\ \vdots & 0 & \ddots & 1 \\ -m_{n1} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \overbrace{-m_{21}a_{11}+a_{21}}^{0} & a_{22}-m_{21}a_{12} & \cdots & a_{2n}-m_{21}a_{1n} \\ -m_{31}a_{11}+a_{31} & a_{32}-m_{31}a_{12} & & a_{3n}-m_{31}a_{3n} \\ \vdots & & & \vdots \end{bmatrix} \quad A^{(2)}\vec{x} = M^{(1)}A^{(1)}\vec{x}$$

. To remove the second column:

Multiply **A** by $M^{(2)} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & & 0 \\ \vdots & 1 & & & 0 \\ \vdots & m_{32} & & & \\ \vdots & \vdots & 0 & & \\ 0 & m_{n2} & 0 & 0 & 1 \end{bmatrix}}_{\text{2nd gaussian transformation matrix}}$

$\vdots$

$$A^{(3)}\vec{x} = M^{(2)}A^{(2)}\vec{x} = M^{(2)}M^{(1)}A^{(1)}\vec{x}$$

$\vdots$

$$A^{(n)}\vec{x} = \underbrace{M^{(n-1)}}_{\substack{(n-1)\text{st} \\ \text{Gaussian} \\ \text{transformation} \\ \text{matrix}}} A^{(n-1)}\vec{x} = \underbrace{M^{\ )}\cdots M^{(1)}}_{\hat{M}} A^{(1)}\vec{x}$$

$$\underbrace{A^{(n)}}_{\substack{\text{upper} \\ \text{triangular} \\ \text{matrix}}} \vec{x} = \underbrace{\hat{M}}_{\substack{\text{product} \\ \text{of} \\ \text{gaussian transf.} \\ \text{matrices}}} \underbrace{A^{(1)}}_{\text{initial matrix}} \vec{x} \qquad (*)$$

with $$\tilde{M} = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ \vdots & 1 & & & \vdots \\ \vdots & -m_{32} & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -m_{n2} & -m_{n2} & \cdots & & \end{pmatrix}$$

- To reverse the effect of the transformation

One sets $(E_j + m_{j,k} E_k) \to (E_j)$, $j = k+1, \dots, n$ in $A^{(n)}$

(In other words, one multiplies (*) with M^{-1} from the left)

$$\tilde{M}^{-1} A^{(n)} \vec{x} = \underbrace{\tilde{M}^{-1} \tilde{M}}_{1\!\!1} A^{(1)} \vec{x}$$

$$\underbrace{A^{(1)}}_{\text{initial matrix}} \vec{x} = \underbrace{\tilde{M}^{-1} A^{(n)}}_{\text{LU matrices}} \vec{x}$$

with $$\tilde{M}^{-1} = \begin{pmatrix} 1 & 0 & \cdots & & 0 \\ & 1 & & & \vdots \\ & -m_{32} & \ddots & & 0 \\ & \vdots & & \ddots & \\ -m_{n2} & -m_{n2} & \cdots & & 1 \end{pmatrix}$$

(*) One can use the Gaussian elimination also to factorize matrices (that is what the LU decomposition is!)

Example : Let us consider the linear system

$$
\begin{array}{rrrrcrl}
x_1 & + x_2 & & + 3x_4 & = & 4 & E_1 \\
2x_1 & + x_2 & - x_3 & + x_4 & = & 1 & E_2 \quad (*) \\
3x_1 & - x_2 & - x_3 & + 2x_4 & = & -3 & E_3 \\
-x_1 & + 2x_2 & + 3x_3 & - x_4 & = & 4 & E_4
\end{array}
$$

Matrix of the system: $A = \begin{pmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{pmatrix}$

Let us convert the system $(*)$ to a triangular system:

$$
\left.\begin{array}{ll}
(E_2 - 2E_1) \to (E_2) & (m_{21} = 2) \\
(E_3 - 3E_1) \to (E_3) & (m_{31} = 3) \\
(E_4 - (-1)E_1) \to (E_4) & (m_{41} = -1)
\end{array}\right\} \text{got rid of the entries in the } 1^{st} \text{ column}
$$

$$
\left.\begin{array}{ll}
(E_3 - 4E_2) \to (E_3) & (m_{32} = 4) \\
(E_4 - (-3)E_2) \to (E_4) & (m_{42} = -3)
\end{array}\right\} \text{got rid of the entries in the 2nd column}
$$

This gives

$$
\begin{array}{rrrrcr}
x_1 & + x_2 & & + 3x_4 & = & 4 \\
& - x_2 & - x_3 & - 5x_4 & = & -7 \\
& & 3x_3 & + 13x_4 & = & 13 \\
& & & -13x_4 & = & -13
\end{array}
$$

Hence $U = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix}$

$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix}$ (from what we saw before)

Hence

$$A = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix}$$

(*) One can now solve any linear system involving A

Example

$$A\vec{x} = \underbrace{\begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix}}_{\vec{b}} = L\underbrace{U\vec{x}}_{\vec{y}}$$

- 1st Step : Solve $L\vec{y} = \vec{b}$ (forward substitution)

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix} \to y_1 = 8$$

$16 + y_2 = 7 \to y_2 = -9$

$\vdots$

$y_3 = 26$

$y_4 = -26$

- 2nd step: solve $U\vec{x} = \vec{y}$ (backward substitution)

$$\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\text{known}} \to x_4 = \frac{-26}{-13} = 2$$

$3x_3 + 13x_4 = 26$

$x_3 = 0$

$-\underbrace{x_2 - x_3}_{0} + \underbrace{5x_4}_{10} = -9$

$x_2 = 1$

Algorithm

Input: Dimension n; the entries a_{ij}, $1 \le i, j \le n$ of A; the diagonal $l_{11} = \cdots = l_{nn} = 1$ of L

Output: • The entries l_{ij}, $1 \le j \le i$, $1 \le i \le n$ of L

• The entries u_{ij}, $i \le j \le n$, $1 \le i \le n$ of U

Step 1: Sets value for u_{11}:

If $a_{11} = 0$ then

OUTPUT
('Factorization impossible')
STOP
else $u_{11} = \dfrac{a_{11}}{l_{11}}$

Step 2 For $j = 2, \cdots, n$ set $u_{1j} = a_{1j}/l_{11}$ (first row of U)

$l_{j1} = a_{j1}/u_{11}$ (first column of L

Step 3 For $i = 2, \cdots, n-1$ do steps 4 and 5

Step 4 $u_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik} u_{ki}$

If $u_{ii} = 0$ then
OUTPUT('Factorization impossible')
STOP
(computes entries in the diagonal.)

Step 5: For $j = i+1, \cdots, N$

Set $u_{ij} = \dfrac{1}{l_{ii}} \left[a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \right]$;

(i-th row of U

$$l_{ji} = \frac{1}{u_{ii}}\left[a_{ji} - \sum_{k=1}^{i-1} l_{jk} u_{ki}\right] \quad \text{(ith column of L)}$$

Step 6 $$u_{nn} = \frac{1}{l_{nn}}\left[a_{nn} - \sum_{k=1}^{n-1} l_{nk} u_{kn}\right]$$

(Note: If $l_{nn} u_{nn} = 0$ then $A = LU$ but A is singular)

Step 7 : Output (l_{ij} for $j = 1, \dots, i$ and $i = 1, \dots, n$)
Output (u_{ij} for $j = i, \dots, n$ and $i = 1, \dots, n$)

Subsequently:

- Solution to a linear system of the form :

$$A\vec{x} = L \underbrace{U\vec{x}}_{\vec{y}} = \vec{b}$$

Step 1 : Set $y_1 = \frac{b_1}{l_{11}}$

Step 2 for $i = 2, 3, \dots, n$
$$y_i = \frac{1}{l_{ii}}\left[b_i - \sum_{j=1}^{i-1} l_{ij} y_j\right]$$

(forward substitution)

(*) <u>Please note:</u> the above-stated procedure does not allow ROW interchanges

⇒ This problem can be solved by using a Permutation matrix P and performing the decomposition PA = LU

∴ • P is obtained by rearranging the rows of the identity matrix

Example : $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \left(I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$

$$\therefore PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Similarly

$$AP = \begin{bmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{bmatrix}$$

So $A = \underbrace{P^{-1}}_{\text{inverse permut. matrix}} L\,U$

<u>Remark</u> : The matlab command [L, U, P] = lu(A) creates :

- The lower triangular matrix L
- The " " triangular matrix U
- The permutation matrix P

4- Special types of matrices

- Allow more effective factorization techniques

or

- Have some properties which guarantee that row interchanges are not necessary

4.1 - Strictly diagonally dominant matrices

- **Definition**: The $n \times n$ matrix A is said to be strictly diagonally dominant if

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ji}|$$

holds for each $i = 1, 2, \dots, n$.

Example:

$$A = \begin{bmatrix} 7 & 2 & 0 \\ 3 & 5 & -1 \\ 0 & 5 & -6 \end{bmatrix}$$

$|a_{11}| = 7 \Rightarrow |a_{12}| = 2$, $|a_{13}| = 0$ $\quad |a_{11}| > |a_{12}| + |a_{13}|$

$|a_{22}| = 5 \Rightarrow |a_{21}| = 3$, $|a_{23}| = 1$ $\quad |a_{22}| > |a_{21}| + |a_{23}|$

(*) Please note:

- A strictly diagonally dominant matrix A is nonsingular
- One can perform Gaussian elimination on any linear system of the form $A\vec{x} = \vec{b}$ without row interchanges
- The computations will be stable with respect to the growth of round-off errors

4.2 - Band matrices

Definition: An $n \times n$ matrix is called a band matrix if there exist integers p and q; $p > 1$, $q < n$ such that $a_{ij} = 0$ whenever $p \leq j - i$ or $q \leq i - j$.
(bandwidth: $w = p + q - 1$)

Example: Tridiagonal matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n-1} & a_{nn} \end{bmatrix}$$

(bandwidth 3; $p = q = 2$)

- Factorization procedures can be considerably simplified in this case.

Suppose $A = \mathbf{LU}$, with

$$A \equiv \begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & & \vdots \\ & & \ddots & & 0 \\ 0 & & \ddots & & a_{n+1,n} \\ \vdots & & & & \\ 0 & \cdots & 0 & a_{n-1,n} & a_{nn} \end{bmatrix}$$

$$L = \begin{bmatrix} l_{11} & 0 & \cdots & & 0 \\ l_{21} & l_{22} & \ddots & & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & & & & \\ 0 & \cdots & 0 & l_{n,n-1} & l_{nn} \end{bmatrix}, \quad U = \begin{bmatrix} 1 & u_{12} & 0 & \cdots & 0 \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & & u_{n-1,n} \\ 0 & \cdots & & 0 & 1 \end{bmatrix}$$

multiplying LU we have

$$LU = \begin{bmatrix} l_{11} & l_{11}u_{12} & 0 & \cdots & 0 \\ l_{21} & l_{21}u_{12}+l_{22} & \ddots & & 0 \\ & \ddots & \ddots & & l_{n-1,n-1}u_{n-1,n} \\ 0 & & & & \\ \vdots & & & & \\ 0 & \cdots & 0 & l_{n,n-1} & l_{n,n-1}u_{n-1,n}+l_{nn} \end{bmatrix}$$

- In general:

$$a_{11} = l_{11}$$

$$a_{i,i-1} = l_{i,i-1}, \quad i = 2, \ldots, n$$

$$a_{ii} = l_{i,i-1}u_{i-1,i} + l_{ii}, \quad i = 2, 3, \ldots, n$$

$$a_{i,i+1} = l_{ii}u_{i,i+1}, \quad i = 1, 2, \ldots, n-1$$

We will employ such equations to factorize A

(*) Crout factorization for tridiagonal linear systems

Algorithm

Input: the dimension n, the non-vanishing entries of A

Output: nonvanishing entries of L, U

- Step 1 : Set $l_{11} = a_{11}$;

  $u_{12} = a_{12} / l_{11}$

- Step 2 : For $i = 2, \dots, n-1$ set $l_{i,i-1} = a_{i,i-1}$

  $l_{ii} = a_{ii} - l_{i,i-1} u_{i-1,i}$

  $u_{i,i+1} = a_{i,i+1} / l_{ii}$

- Step 3 : Set $l_{n,n-1} = a_{n,n-1}$

  $l_{nn} = a_{nn} - l_{n,n-1} u_{n-1,n}$

- Step 4 Output

  Stop

(*) Please note: This also considerably simplifies the solutions of linear systems.

$$A\vec{x} = \vec{b}$$

$$L\underbrace{U\vec{x}}_{\vec{y}} = b \quad \text{solve for } \vec{y}$$

Solve $U\vec{x} = \vec{y}$ for $\vec{x}$

4.3 - Symmetric and positive definite matrices

- Symmetric : $a_{ij} = a_{ji}$
- Positive definite: $\vec{v} \cdot \vec{A} \cdot \vec{v} > 0$ for all vectors $\vec{v}$

$$[v_1, \dots, v_n] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_i v_j$$

Other criteria :

- A has an inverse ($\det A \neq 0$)
- $a_{ii} > 0$, $i = 1, 2, \dots, n$
- $\max_{1 \leq k, j \leq n} |a_{kj}| \leq \max_{1 \leq i \leq n} |a_{ii}|$
- $(a_{ij})^2 < a_{ii} a_{jj}$ for each $i \neq j$

Example :

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$ is symmetric and positive definite

- Symmetric: $a_{12} = 1$

$a_{21} = 1 = a_{12}$ ✓

- Positive definite :

$$[v_1 \quad v_2] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = [v_1 \quad v_2] \begin{bmatrix} 2v_1 + v_2 \\ v_1 + 2v_2 \end{bmatrix}$$

$$= 2v_1^2 + 2v_2^2 + 2v_1v_2 =$$

$$= v_1(2v_1 + v_2) + v_2(v_1 + 2v_2) = 2v_1^2 + 2v_2^2 + 2v_1v_2 =$$

$$= (v_1 + v_2)^2 + v_1^2 + v_2^2 > 0 \quad \text{unless } v_1 = v_2 = 0 .$$

Note that: (a) $a_{11} = 2 > 0$ ✓

$a_{22} = 2 > 0$

(b) $\det A = 4 - 1 = 3 \neq 0$ ✓

(c) $\underbrace{\max_{1 \le k, j \le n} |a_{kj}| = 1}_{(*)}$, $\underbrace{\max |a_{ii}| = 2}_{(**)}$

$(*) < (**)$ ✓

(d) $(a_{ij})^2 = 1$

$a_{ii}a_{jj} = 4$ $\longrightarrow$ $1 < 4$ ✓

(*) Choleski decomposition

$$A = \underbrace{L}_{\substack{\text{lower}\\ \text{diagonal}\\ \text{matrix}}} \cdot \underbrace{L^T}_{\text{transpose of } L}$$

Remember:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & & a_{nn} \end{bmatrix}$$

$$A^T = \begin{bmatrix} a_{11} & \cdots & a_{n1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nn} \end{bmatrix}$$

$\Rightarrow a^T_{ij} = a_{ji}$

Symmetric matrices: $A^T = A$

Example : 3 x 3 matrix

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{(*)} = \underbrace{\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}}_{L^T}$$

$$L L^T = \underbrace{\begin{bmatrix} l_{11}^2 & l_{11} l_{21} & l_{11} l_{31} \\ l_{21} l_{11} & l_{21}^2 + l_{22}^2 & l_{21} l_{31} + l_{22} l_{32} \\ l_{31} l_{11} & l_{31} l_{21} + l_{32} l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}}_{(**)}$$

$(*) = (**)$

$$l_{11}^2 = a_{11}$$
$$l_{22}^2 = a_{22} - l_{21}^2 \qquad \Rightarrow \quad l_{ii}^2 = a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2$$
$$l_{33}^2 = a_{33} - l_{31}^2 - l_{32}^2 \qquad l_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)^{1/2}$$

$$\overbrace{a_{12} = a_{21}}^{\text{symmetric}} = l_{21} l_{11} \rightarrow l_{21} = a_{12} / l_{11}$$

$$a_{31} = a_{13} = l_{11} l_{31} \rightarrow l_{31} = a_{13} / l_{11}$$

$$a_{32} = a_{23} = l_{31} l_{21} + l_{32} l_{22} \rightarrow l_{32} = \frac{1}{l_{22}} \left[\overset{a_{32}}{\overset{\shortparallel}{a_{23}}} - l_{31} l_{21} \right]$$

$\vdots$

$$l_{ji} = \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik} \right) \Big/ l_{ii} \quad , j \neq i$$

Algorithm:

Input: dimension n, entries a_{ij}, for $1 \le i, j \le n$ of A

Output: the entries l_{ij}, for $1 \le j \le i$, $1 \le i \le n$ of L

(the entries of $U = L^T$ are $u_{ij} = l_{ji}$, for $i \le j \le n$ and $1 \le i \le n$)

Step 1: Set $l_{11} = \sqrt{a_{11}}$

Step 2: For $j = 2, \ldots, n$ set $l_{j1} = a_{j1} / l_{11}$ (first column of L)

Step 3: For $i = 2, \ldots, n-1$ do steps 4 and 5

Step 4 Set $l_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2\right)^{1/2}$

Step 5 For $j = i+1, \ldots, n$

Set $l_{ji} = \left(a_{ji} - \sum_{k=1}^{i-1} l_{jk} l_{ik}\right) / l_{ii}$

Step 6: Set $l_{nn} = \left(a_{nn} - \sum_{k=1}^{n-1} l_{nk}^2\right)^{1/2}$

Step 7: Output (l_{ij}, for $j = 1, \ldots, i$ and $i = 1, \ldots, n$)