# Numerical Methods for Astrophysics: SOLUTION OF LINEAR EQUATIONS

Michela Mapelli

Systems of linear equations are common in many fields of physics and engineering (e.g. electrical circuits)

How to solve a system of n linear algebraic equations in n unknowns

$$A_{11} x_1 + A_{12} x_2 + \dots + A_{1n} x_n = b_1$$

$$A_{21} x_1 + A_{22} x_2 + \dots + A_{2n} x_n = b_2$$

$$\vdots$$

$$A_{n1} x_1 + A_{n2} x_2 + \dots + A_{nn} x_n = b_n$$

where  $A_{ij}$  and  $b_{ij}$  are known, while  $x_i$  are unknown.

In matrix form:



or in compact form **A x** = **b** 

A system like A x = b admits a unique solution (the vector x) provided that its determinant is non zero (if det = 0 we have a singular matrix) If the determinant is zero, the system has no or infinite solutions

Two different approaches to solve linear equations:

#### **1. DIRECT METHODS:**

transform the original equations into equivalent equations (i.e. equations that have the same solution) that can be solved more easily.

The transformation can apply the 3 elementary operations (do not affect the solution but change the determinant):

#### i) exchanging two equations

ii) multiplying an equation by a non-zero constant

iii) multiplying an equation by a non-zero constant and then subtracting it from another equation

Examples of direct methods: Gauss elimination, LU decomposition

Two different approaches to solve linear equations:

#### 2. INDIRECT METHODS:

Start with a guess of the solution (Ansatz) and then repeatedly refine it until a given convergence criterion is satisfied

Indirect methods are less efficient than direct ones but more used because

- simpler to implement,
- more efficient if matrix is large and sparse (lots of zeros)

**DRAWBACK: do not always converge** 

**Examples of indirect methods: Gauss - Seidel method** 

Suppose we must solve

$$2w + x + 4y + z = -4$$
$$3w + 4x - y - z = 3$$
$$w - 4x + y + 5z = 9$$

$$2w - 2x + y + 3z = 7$$

Or in matrix form

$$\begin{bmatrix} 2 & 1 & 4 & 1 \\ 3 & 4 & -1 & -1 \\ 1 & -4 & 1 & 5 \\ 2 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ 9 \\ 7 \end{bmatrix}$$

To solve analytically I should find the INVERSE of matrix A and then calculate x = A<sup>-1</sup> b This is very inefficient

FINAL GOAL: produce a upper triangular matrix with diagonal members = 1



Two basic rules adopted in Gauss elimination:

- We can multiply any of our equations by a constant and it is still the same equation (= the solution does not change) if we multiply any row of A and the corresponding row of b by any constant, then the solution does not change.
- We can take any linear combination of two equations to get another correct equation.

If we add to or subtract from any row of **A** a multiple of any other row, and we do the same for the vector **b**, the solution does not change.

#### Name convention

$$\begin{bmatrix} A_{00} & A_{01} & A_{02} & A_{03} \\ A_{10} & A_{11} & A_{12} & A_{13} \\ A_{20} & A_{21} & A_{22} & A_{23} \\ A_{30} & A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

STEP 1: divide first row of A by  $A_{00}$  to change  $A_{00}$  from 2 to 1



STEP 1: divide first row of A by  $A_{00}$  to change  $A_{00}$  from 2 to 1



**STEP 2: zero A\_{10}** by subtracting 3 times the first row from the second row



**STEP 2: zero A\_{10}** by subtracting 3 times the first row from the second row



STEPs 3 and 4: zero A<sub>20</sub> and A<sub>30</sub> by doing a similar trick with the first row (subtract 1 time the first row from the third and 2 times the first row from the fourth)

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 2.5 & -7 & -2.5 \\ 1 & -4 & 1 & 5 \\ 2 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \\ 9 \\ 7 \end{bmatrix}$$

STEPs 3 and 4: zero A<sub>20</sub> and A<sub>30</sub> by doing a similar trick with the first row (subtract 1 time the first row from the third and 2 times the first row from the fourth)

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 2.5 & -7 & -2.5 \\ 0 & -4.5 & -1 & 4.5 \\ 2 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \\ 11 \\ 7 \end{bmatrix}$$

STEPs 3 and 4: zero A<sub>20</sub> and A<sub>30</sub> by doing a similar trick with the first row (subtract 1 time the first row from the third and 2 times the first row from the fourth)

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 2.5 & -7 & -2.5 \\ 0 & -4.5 & -1 & 4.5 \\ 0 & -3 & -3 & 2.0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \\ 11 \\ 11 \end{bmatrix}$$

**STEP 5: change A**<sub>11</sub> from 2.5 to 1 by dividing the row by A<sub>11</sub>

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 2.5 & -7 & -2.5 \\ 0 & -4.5 & -1 & 4.5 \\ 0 & -3 & -3 & 2.0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \\ 11 \\ 11 \end{bmatrix}$$

**STEP 5: change A**<sub>11</sub> from 2.5 to 1 by dividing the row by A<sub>11</sub>

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & -4.5 & -1 & 4.5 \\ 0 & -3 & -3 & 2.0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ 11 \\ 11 \end{bmatrix}$$

STEPs 6 and 7: zero  $A_{21}$  and  $A_{31}$  by subtracting

- 4.5 times the second row from the third, and
- 3.0 times the second row from the fourth

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & -4.5 & -1 & 4.5 \\ 0 & -3 & -3 & 2.0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ 11 \\ 11 \end{bmatrix}$$

STEPs 6 and 7: zero  $A_{21}$  and  $A_{31}$  by subtracting

- 4.5 times the second row from the third, and
- 3.0 times the second row from the fourth

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & 0 & -13.6 & 0 \\ 0 & -3 & -3 & 2.0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ 27.2 \\ 11 \end{bmatrix}$$

STEPs 6 and 7: zero  $A_{21}$  and  $A_{31}$  by subtracting

- 4.5 times the second row from the third, and
- 3.0 times the second row from the fourth

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & 0 & -13.6 & 0 \\ 0 & 0 & -11.4 & -1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ 27.2 \\ 21.8 \end{bmatrix}$$

STEP 8: change A<sub>22</sub> from –13.6 to 1 by dividing the entire row by –13.6

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & 0 & -13.6 & 0 \\ 0 & 0 & -11.4 & -1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ 27.2 \\ 21.8 \end{bmatrix}$$

STEP 8: change A<sub>22</sub> from –13.6 to 1 by dividing the entire row by –13.6

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -11.4 & -1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ -2 \\ 21.8 \end{bmatrix}$$

STEP 9: zero A<sub>23</sub> by subtracting –11.4 times the third row from the fourth one

$$\begin{bmatrix} 1 & 0.5 & 2 & 0.5 \\ 0 & 1 & -2.8 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -11.4 & -1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 3.6 \\ -2 \\ 21.8 \end{bmatrix}$$

STEP 9: zero A<sub>23</sub> by subtracting –11.4 times the third row from the fourth one



# STEP 10: change $A_{_{33}}$ to 1 by dividing the fourth row by $A_{_{33}}$



# STEP 10: change $A_{_{33}}$ to 1 by dividing the fourth row by $A_{_{33}}$



#### THE MATRIX IS NOW UPPER TRIANGULAR

SUMMARY:

- first we divide a row by its diagonal element to obtain
  A<sub>ij</sub> = 1 if i == j
- then we subtract N times an upper row i from a lower row k to zero the element of column j below the diagonal

where  $N = A_{kj}$ 

When the matrix is upper triangular

The system of equations looks like

$$\begin{bmatrix} 1 & a_{01} & a_{02} & a_{03} \\ 0 & 1 & a_{12} & a_{13} \\ 0 & 0 & 1 & a_{23} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Which is the same as

$$w + a_{01}x + a_{02}y + a_{03}z = b_0$$
$$x + a_{12}y + a_{13}z = b_1$$
$$y + a_{23}z = b_2$$
$$z = b_3$$

The system 
$$w + a_{01}x + a_{02}y + a_{03}z = b_0$$
  
 $x + a_{12}y + a_{13}z = b_1$   
 $y + a_{23}z = b_2$   
 $z = b_3$ 

can be solved by backsubstitution, i.e. by starting from the last equation and using it to solve the previous one and so on:

$$z = b3$$
  
 $y = -a23 z + b2$   
 $x = -a12 y - a13 z + b1$   
 $w = -a01 x - a02 y - a03z + b0$ 

#### **EXERCISE:**

Produce a python script to implement the Gauss elimination method with backsubstitution and solve the system 14. Solution of the exercise: [2. -1. -2.1.]

System 14 of the notes is our example:



Since it is your first difficult script let's do it together

**Advice 1: decompose your problem in smaller steps:** 

**STEP 0: upload the matrix and vector** 

**STEP 1: divide the lines by their diagonal elements** 

STEP 2: zero the elements below the diagonal starting from 2<sup>nd</sup> row

**STEP 3: generalize the zeroing to all the k rows below the 1<sup>st</sup> one** 

STEP 4: generalize the zeroing to all the rows (not only the 1<sup>st</sup> one)

**STEP 5: back substitution** 

STEP 6: program debugged and running. How can I improve it? Make it faster, more elegant, add comments

#### **Linear equations.** Pivoting

# Suppose we want to calculate a matrix which has some ZEROS on the diagonal

$$\begin{bmatrix} 0 & 1 & 4 & 1 \\ 3 & 4 & -1 & -1 \\ 1 & -4 & 1 & 5 \\ 2 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ 9 \\ 7 \end{bmatrix}$$

The Gauss elimination does not work: divide by zero I can try to change the order of the rows (PIVOTING)

$$\begin{bmatrix} 3 & 4 & -1 & -1 \\ 0 & 1 & 4 & 1 \\ 1 & -4 & 1 & 5 \\ 2 & -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ -4 \\ 9 \\ 7 \end{bmatrix}$$

Method does not work if matrix is too sparse

A modification of Gaussian elimination + pivoting

Suppose we want to solve many sets of equations in the form Ax = b, in which A is the same but b changes

→ We want to keep track of the transformations we make to A, to avoid re-calculating them for each b

The LU decomposition is a modification of Gaussian elimination + pivoting that allows remembering the transformations

Consider a generic 4x4 matrix A	<i>a</i> <sub>00</sub>	$a_{01}$	$a_{02}$	$a_{03}$
	<i>a</i> <sub>10</sub>	$a_{11}$	<i>a</i> <sub>12</sub>	<i>a</i> <sub>13</sub>
	<i>a</i> <sub>20</sub>	<i>a</i> <sub>21</sub>	a <sub>22</sub>	a <sub>23</sub>
	a <sub>30</sub>	<i>a</i> <sub>31</sub>	a <sub>32</sub>	a <sub>33</sub>

STEPs 1, 2 and 3 of Gauss elimination can be written as (check yourself)

$$\frac{1}{a_{00}}\begin{bmatrix}1&0&0&0\\-a_{10}&a_{00}&0&0\\-a_{20}&0&a_{00}&0\\-a_{30}&0&0&0&a_{00}\end{bmatrix}\begin{bmatrix}a_{00}&a_{01}&a_{02}&a_{03}\\a_{10}&a_{11}&a_{12}&a_{13}\\a_{20}&a_{21}&a_{22}&a_{23}\\a_{30}&a_{31}&a_{32}&a_{33}\end{bmatrix}=\begin{bmatrix}1&b_{01}&b_{02}&b_{03}\\0&b_{11}&b_{12}&b_{13}\\0&b_{21}&b_{22}&b_{23}\\0&b_{31}&b_{32}&b_{33}\end{bmatrix}$$

- divide first row by a<sub>00</sub>,
- then subtract the first row  $a_{10}$  times from the second
- then subtract the first row a<sub>20</sub> times from the third
- then subtract the first row  $a_{30}$  times from the fourth

#### Let's name $L_0$ the first part of this



$$\mathbf{L}_{\mathbf{0}} = \frac{1}{a_{00}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -a_{10} & a_{00} & 0 & 0 \\ -a_{20} & 0 & a_{00} & 0 \\ -a_{30} & 0 & 0 & a_{00} \end{bmatrix}$$

#### It can be shown that the next step of Gaussian elim. is

$$\frac{1}{b_{11}} \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -b_{21} & b_{11} & 0 \\ 0 & -b_{31} & 0 & b_{11} \end{bmatrix} \begin{bmatrix} 1 & b_{01} & b_{02} & b_{03} \\ 0 & b_{11} & b_{12} & b_{13} \\ 0 & b_{21} & b_{22} & b_{23} \\ 0 & b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} 1 & c_{01} & c_{02} & c_{03} \\ 0 & 1 & c_{12} & c_{13} \\ 0 & 0 & c_{22} & c_{23} \\ 0 & 0 & c_{32} & c_{33} \end{bmatrix}$$

Let's define L<sub>1</sub>

$$\mathbf{L}_{1} = \frac{1}{b_{11}} \begin{bmatrix} b_{11} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -b_{21} & b_{11} & 0 \\ 0 & -b_{31} & 0 & b_{11} \end{bmatrix}$$

# It can be shown that the last steps of Gauss elimination correspond to multiply the results by these matrices

$$\mathbf{L}_{2} = \frac{1}{c_{22}} \begin{bmatrix} c_{22} & 0 & 0 & 0 \\ 0 & c_{22} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -c_{32} & c_{22} \end{bmatrix}, \qquad \mathbf{L}_{3} = \frac{1}{d_{33}} \begin{bmatrix} d_{33} & 0 & 0 & 0 \\ 0 & d_{33} & 0 & 0 \\ 0 & 0 & d_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In compact form A x = b becomes

$$L_{3} L_{2} L_{1} L_{0} A x = L_{3} L_{2} L_{1} L_{0} b$$

where

 $L_3 L_2 L_1 L_0 A$  is a UPPER TRIANGULAR matrix, because it is the result of Gauss elimination

 $L_3 L_2 L_1 L_0 b$  is a known vector

 $\rightarrow$  I can calculate vector x by backsubstitution

In practice, the LU decomposition is implemented in a slightly modified way

**1.** The INVERSE of matrices  $L_0$ ,  $L_1$ ,  $L_2$  and  $L_3$  are:

$$\mathbf{L}_{0}^{-1} = \begin{bmatrix} a_{00} & 0 & 0 & 0 \\ a_{10} & 1 & 0 & 0 \\ a_{20} & 0 & 1 & 0 \\ a_{30} & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{L}_{1}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & b_{11} & 0 & 0 \\ 0 & b_{21} & 1 & 0 \\ 0 & 0 & b_{31} & 1 \end{bmatrix}$$
$$\mathbf{L}_{2}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_{22} & 0 \\ 0 & 0 & c_{32} & 1 \end{bmatrix} \quad \mathbf{L}_{3}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & d_{33} \end{bmatrix}$$

2. Define 
$$\mathbf{L} = \mathbf{L}_{0}^{-1} \mathbf{L}_{1}^{-1} \mathbf{L}_{2}^{-1} \mathbf{L}_{3}^{-1}$$
  
$$\mathbf{L} = \begin{bmatrix} a_{00} & 0 & 0 & 0 \\ a_{10} & b_{11} & 0 & 0 \\ a_{20} & b_{21} & c_{22} & 0 \\ a_{30} & b_{31} & c_{32} & d_{33} \end{bmatrix}$$

**3. Define**  $U = L_3 L_2 L_1 L_0 A$ 

4. By definition of inverse matrix we have that

L U = A

**5.** Hence, equation A x = b can be written as

LUx = b

# Writing a script for LU decomposition can be complex BUT USE NUMPY

```
from numpy.linalg import solve
x=solve(A,b)
```

In general, mathematical functions in python are much better optimized than you can ever achieve by writing your own function. So, if there is a pre-defined function doing exactly what you need, use it without re-inventing the wheel. On the other hand, please avoid using python functions without knowing what they do exactly, i.e. which algorithms they implement and what are the validity limitations of these algorithms.

#### **Linear equations.** Gauss – Seidel method

**Indirect method. Faster for sparse matrices** 

1. Write 
$$A x = b$$
 in scalar notation  

$$\sum_{j=1}^{n} A_{ij} x_j = b_i \quad i = 1, 2, ..., n$$

2. Extract the diagonal terms from the sum

$$A_{ii} x_i + \sum_{j=1, j \neq i}^n A_{ij} x_j = b_i \quad i = 1, 2, ..., n$$

#### **Linear equations.** Gauss – Seidel method

**3.** Reshuffle so that we have only x<sub>i</sub> on the left side

$$x_{i} = \frac{1}{A_{ii}} \left( b_{i} - \sum_{j=1, j \neq i}^{n} A_{ij} x_{j} \right) \quad i = 1, 2, ..., n$$

4. The above equation suggests a solution BY ITERATION

$$x_i \leftarrow \frac{1}{A_{ii}} \left( b_i - \sum_{j=1, j \neq i}^n A_{ij} x_j \right) \quad i = 1, 2, ..., n$$

- We start with a GUESS on x that we plug into the right-hand terms
- We re-calculate the left-hand terms
- We iterate this till the difference between the previous and the next version of x<sub>i</sub> becomes "small enough"

Linear equations. Pros and cons of the algorithms

#### LU decomposition and Gaussian elimination:

**PROS:** 

- easy to implement
- fast
- exact solution

CONS:

- not always possible to use

#### **Gauss-Seidel method:**

**PROS**:

- very easy to implement
- fast (especially if sparse matrices)

CONS:

- approximate solution
- might fail to converge
- not always possible to use

# **Linear equations.** Exercise

#### **EXERCISE:**

Write a script to implement the Gauss-Seidel method and use it to solve the following equations.

$$\begin{array}{ccc} 4 & -1 & 1 \\ -1 & 4 & -2 \\ 1 & -2 & 4 \end{array} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ -1 \\ 5 \end{bmatrix}$$
(45)

Solution of the exercise: [3.0, 1.0, 1.0].

### **Linear equations.** Exercise

#### **EXERCISE:**

Use the Gauss-Seidel method and the Gauss elimination method with backsubstitution (or the LU method with backsubstitution) to solve the circuit of resistors shown in Figure 23. All the resistors have the same resistance R. The power rail at the top is at voltage  $V_+ = 5$  V. What are the other voltages  $V_1$  and  $V_4$ ?

Suggestion to solve the problem:

1. Write down the system of linear equations using Ohm's law and Kirchhoff current law.

For example, for the junction at Voltage V<sub>1</sub> we have

$$\frac{V_1 - V_2}{R} + \frac{V_1 - V_3}{R} + \frac{V_1 - V_4}{R} + \frac{V_1 - V_+}{R} = 0$$
(46)

2. Adapt your Gauss-Seidel and Gauss elimination scripts to solve the equations.

Solution of the exercise: [3., 1.66666667, 3.3333333, 2].

#### **Linear equations.** Exercise

