Chapter 2
Linear Equations

Gaussian elimination

2.1 Introduction to Systems of Equations

Linear Equations

- Any straight line in xy-plane can be represented algebraically by an equation of the form: $$a_1 x + a_2 y = b$$
- General form: define a linear equation in the *n* variables $x_1, x_2, \ldots, x_n$: $$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b$$
  - Where $a_1, a_2, \ldots, a_n,$ and b are real constants.
  - The variables in a linear equation are sometimes called unknowns.

Linear Systems (1/2)

- A finite **set** of linear equations in the variables $x_1, x_2, \ldots, x_n$ is called a system of linear equations or a linear system .

- A sequence of numbers $s_1, s_2, \ldots, s_n$ is called a solution of the system.

- A system has **no** solution is said to be inconsistent ; if there is **at least** one solution of the system, it is called consistent.

$$
\begin{array}{l}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\vdots \qquad \vdots \qquad \qquad \vdots \qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m
\end{array}
$$

↑ An arbitrary system of *m* linear equations in *n* unknowns

Elementary Row Operations

- The basic method for solving a system of linear equations is to replace the given system by **a new system that has the same solution set** but which is **easier** to solve.

- Since the **rows** of an augmented matrix correspond to the **equations** in the associated system. new systems is generally obtained in a series of steps by applying the following three types of operations to eliminate unknowns systematically. These are called elementary row operations.
  1. Multiply an equation through by an nonzero constant.
  2. Interchange two equation.
  3. Add a multiple of one equation to another.

2.2 Gaussian Elimination

Echelon Forms

- This matrix which have following properties is in reduced row-echelon form (Example 1, 2).

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a leader 1.
2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leader 1 in the lower row occurs farther to the right than the leader 1 in the higher row.
4. Each column that contains a leader 1 has zeros everywhere else.

- A matrix that has the first three properties is said to be in row-echelon form (Example 1, 2).
- A matrix in reduced row-echelon form is of necessity in row-echelon form, but not conversely.

Example 1
Row-Echelon & Reduced Row-Echelon form

- reduced row-echelon form:

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

- row-echelon form:

$$\begin{bmatrix} 1 & 4 & -3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Example 2
More on Row-Echelon and Reduced Row-Echelon form

- All matrices of the following types are in **row-echelon form** (any real numbers substituted for the *'s.) :

$$\begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

- All matrices of the following types are in **reduced row-echelon form** (any real numbers substituted for the *'s.) :

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & * \\ 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

Example 3
Solutions of Four Linear Systems (a)

Suppose that the augmented matrix for a system of linear equations have been reduced by row operations to the given reduced row-echelon form. Solve the system.

$$\text{(a)}\ \begin{bmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

Solution (a)

the corresponding system of equations is : $\longrightarrow$

$$\begin{aligned} x \qquad\qquad &= 5 \\ y \qquad &= -2 \\ z &= 4 \end{aligned}$$

Example 3
Solutions of Four Linear Systems (b1)

$$\text{(b)} \begin{bmatrix} 1 & 0 & 0 & 4 & -1 \\ 0 & 1 & 0 & 2 & 6 \\ 0 & 0 & 1 & 3 & 2 \end{bmatrix}$$

Solution (b)

1. The corresponding system of equations is :

free variables

$$\begin{aligned} x_1 \qquad\qquad & + 4x_4 = -1 \\ x_2 \qquad & + 2x_4 = 6 \\ x_3 & + 3x_4 = 2 \end{aligned}$$

leading variables

Example 3
Solutions of Four Linear Systems (b2)

$$x_1 = -1 - 4x_4$$
$$x_2 = 6 - 2x_4$$
$$x_3 = 2 - 3x_4$$

2. We see that the free variable can be assigned an arbitrary value, say t, which then determines values of the leading variables.

3. There are infinitely many solutions, and the general solution is given by the formulas

$$x_1 = -1 - 4t,$$
$$x_2 = 6 - 2t,$$
$$x_3 = 2 - 3t,$$
$$x_4 = t$$

Example 3
Solutions of Four Linear Systems (c1)

$$(c)\ \begin{bmatrix} 1 & 6 & 0 & 0 & 4 & -2 \\ 0 & 0 & 1 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solution (c)

1. The 4th row of zeros leads to the equation places no restrictions on the solutions (why?). Thus, we can omit this equation. $\longrightarrow$

$$\begin{aligned} x_1 + 6x_2 \qquad\qquad + 4x_5 &= -2 \\ x_3 \quad + 3x_5 &= 1 \\ x_4 + 5x_5 &= 2 \end{aligned}$$

Example 3
Solutions of Four Linear Systems (c2)

Solution (c)

2. Solving for the leading variables in terms of the free variables: $\rightarrow$

$$
\begin{aligned}
x_1 &= -2 - 6x_2 - 4x_5 \\
x_3 &= 1 - 3x_5 \\
x_4 &= 2 - 5x_5
\end{aligned}
$$

3. The free variable can be assigned an arbitrary value,there are infinitely many solutions, and the general solution is given by the formulas. $\longrightarrow$

$$
\begin{aligned}
x_1 &= -2 - 6s - 4t\ , \\
x_2 &= s \\
x_3 &= 1 - 3t \\
x_4 &= 2 - 5t, \\
x_4 &= t
\end{aligned}
$$

Example 3
Solutions of Four Linear Systems (d)

$$\text{(d)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution (d):

the last equation in the corresponding system of equation is

$$0x_1 + 0x_2 + 0x_3 = 1$$

Since this equation cannot be satisfied, there is **no solution** to the system.

Elimination Methods (1/7)

- We shall give a step-by-step elimination procedure that can be used to reduce any matrix to reduced row-echelon form.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Elimination Methods (2/7)

- Step1. Locate the leftmost column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Leftmost nonzero column

- Step2. Interchange the top row with another row, to bring a nonzero entry to top of the column found in Step1.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The 1th and 2th rows in the preceding matrix were interchanged.

Elimination Methods (3/7)

- Step3. If the entry that is now at the top of the column found in Step1 is a, multiply the first row by 1/a in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The 1st row of the preceding matrix was multiplied by 1/2.

- Step4. Add suitable multiples of the top row to the rows below so that all entires below the leading 1 become zeros.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

-2 times the 1st row of the preceding matrix was added to the 3rd row.

Elimination Methods (4/7)

- Step5. Now cover the top row in the matrix and begin again with Step1 applied to the submatrix that remains. Continue in this way until the entire matrix is in row-echelon form.

$$\begin{bmatrix} \boxed{1 \quad 2 \quad -5 \quad 3 \quad 6 \quad 14} \\ 0 \quad 0 \quad -2 \quad 0 \quad 7 \quad 12 \\ 0 \quad 0 \quad -5 \quad 0 \quad -17 \quad -29 \end{bmatrix}$$

Leftmost nonzero column in the submatrix

$$\begin{bmatrix} \boxed{1 \quad 2 \quad -5 \quad 3 \quad 6 \quad 14} \\ 0 \quad 0 \quad 1 \quad 0 \quad -\frac{7}{2} \quad -6 \\ 0 \quad 0 \quad 5 \quad 0 \quad -17 \quad -29 \end{bmatrix}$$

The 1st row in the submatrix was multiplied by -1/2 to introduce a leading 1.

Elimination Methods (5/7)

- Step5 (cont.)

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

-5 times the 1st row of the submatrix was added to the 2nd row of the submatrix to introduce a zero below the leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

The top row in the submatrix was covered, and we returned again Step1.

Leftmost nonzero column in the new submatrix

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

The first (and only) row in the new submetrix was multiplied by 2 to introduce a leading 1.

- The **entire** matrix is now in **row-echelon form**.

Elimination Methods (6/7)

- Step6. Beginning with las nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

7/2 times the 3rd row of the preceding matrix was added to the 2nd row.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

-6 times the 3rd row was added to the 1st row.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

5 times the 2nd row was added to the 1st row.

- The **last** matrix is in **reduced row-echelon form**.

Elimination Methods (7/7)

- Step1~Step5: the above procedure produces a row-echelon form and is called Gaussian elimination.
- Step1~Step6: the above procedure produces a reduced row-echelon form and is called Gaussian-Jordan elimination.
- Every matrix has **a unique reduced row-echelon** form but a row-echelon form of a given matrix is not unique.

Example 4
Gauss-Jordan Elimination(1/4)

- Solve by Gauss-Jordan Elimination

$$
\begin{aligned}
x_1 + 3x_2 - 2x_3 \qquad\quad + 2x_5 \qquad\quad &= 0 \\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\
5x_3 + 10x_4 \qquad\quad + 15x_6 &= 5 \\
2x_1 + 6x_2 \qquad\quad + 8x_4 + 4x_5 - 18x_6 &= 6
\end{aligned}
$$

- Solution:

The augmented matrix for the system is

$$
\begin{bmatrix}
1 & 3 & -2 & 0 & 2 & 0 & 0 \\
2 & 6 & -5 & -2 & 4 & -3 & -1 \\
0 & 0 & 5 & 10 & 0 & 15 & 5 \\
2 & 6 & 0 & 8 & 4 & 18 & 6
\end{bmatrix}
$$

Example 4 Gauss-Jordan Elimination(2/4)

- Adding -2 times the 1st row to the 2nd and 4th rows gives

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix}$$

- Multiplying the 2nd row by -1 and then adding -5 times the new 2nd row to the 3rd row and -4 times the new 2nd row to the 4th row gives

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 2 \end{bmatrix}$$

Example 4 Gauss-Jordan Elimination(3/4)

- Interchanging the 3rd and 4th rows and then multiplying the 3rd row of the resulting matrix by 1/6 gives the row-echelon form.

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Adding -3 times the 3rd row to the 2nd row and then adding 2 times the 2nd row of the resulting matrix to the 1st row yields the reduced row-echelon form.

$$\begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 4
Gauss-Jordan Elimination(4/4)

- The corresponding system of equations is

$$
\begin{aligned}
x_1 + 3x_2 \qquad + 4x_4 + 2x_5 \qquad &= 0 \\
x_3 + 2x_4 \qquad &= 0 \\
x_6 &= \tfrac{1}{3}
\end{aligned}
$$

- Solution

The augmented matrix for the system is

$$
\begin{aligned}
x_1 &= -3x_2 - 4x_4 - 2x_5 \\
x_3 &= -2x_4 \\
x_6 &= \tfrac{1}{3}
\end{aligned}
$$

- We assign the free variables, and the general solution is given by the formulas:

$$x_1 = -3r - 4s - 2t,\ x_2 = r,\ x_3 = -2s,\ x_4 = s,\ x_5 = t,\ \ x_6 = \tfrac{1}{3}$$

Back-Substitution

- It is sometimes preferable to solve a system of linear equations by using Gaussian elimination to bring the augmented matrix into row-echelon form **without continuing all the way to the reduced row-echelon form**.
- When this is done, the corresponding system of equations can be solved by solved by a technique called back-substitution.
- Example 5

Example 5
ex4 solved by Back-substitution(1/2)

- From the computations in Example 4, a row-echelon form from the augmented matrix is

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- To solve the corresponding system of equations

$$\begin{aligned} x_1 + 3x_2 \qquad + 4x_4 + 2x_5 \qquad &= 0 \\ x_3 + 2x_4 \qquad &= 0 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

- Step1. Solve the equations for the leading variables.

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 \\ x_3 &= 1 - 2x_4 - 3x_6 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Example5
ex4 solved by Back-substitution(2/2)

- Step2. Beginning with the bottom equation and working upward, successively substitute each equation into all the equations above it.
  - Substituting x6=1/3 into the 2nd equation

$$x_1 = -3x_2 + 2x_3 - 2x_5$$
$$x_3 = -2x_4$$
$$x_6 = \tfrac{1}{3}$$

  - Substituting x3=-2 x4 into the 1st equation

$$x_1 = -3x_2 + 2x_3 - 2x_5$$
$$x_3 = -2x_4$$
$$x_6 = \tfrac{1}{3}$$

- Step3. Assign free variables, the general solution is given by the formulas.

$$x_1 = -3r - 4s - 2t,\ x_2 = r,\ x_3 = -2s,\ x_4 = s,\ x_5 = t,\ \ x_6 = \tfrac{1}{3}$$

Example 6
Gaussian elimination(1/2)

- Solve by Gaussian elimination and back-substitution. (ex3 of Section1.1)

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + 6y - 5z &= 0 \end{aligned}$$

- Solution
  - We convert the augmented matrix $\begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$
  - to the ow-echelon form $\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$
  - The system corresponding to this matrix is

$$x + y + 2z = 9,\ \ y - \tfrac{7}{2}z = -\tfrac{17}{2},\ \ z = 3$$

Example 6 Gaussian elimination(2/2)

- Solution
  - Solving for the leading variables

$$x = 9 - y - 2z,$$
$$y = -\tfrac{17}{2} + \tfrac{7}{2}z,$$
$$z = 3$$

  - Substituting the bottom equation into those above

$$x = 3 - y,$$
$$y = 2,$$
$$z = 3$$

  - Substituting the 2nd equation into the top

$$x = 1,\ \ y = 2,\ \ z = 3$$

Homogeneous Linear Systems(1/2)

- A system of linear equations is said to be homogeneous if the constant terms are all zero; that is , the system has the form : $\longrightarrow$

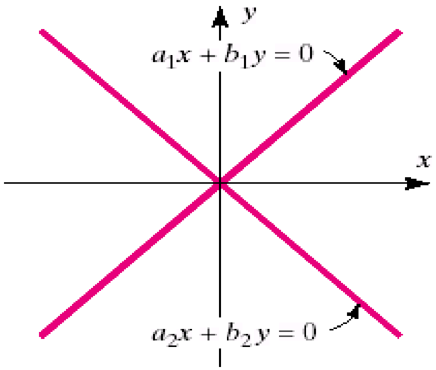
$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}$$

- Every homogeneous system of linear equation is **consistent**, since all such system have $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ as a solution. This solution is called the trivial solution; if there are another solutions, they are called nontrivial solutions.
- There are only two possibilities for its solutions:
  - The system has **only** the trivial solution.
  - The system has **infinitely** many solutions in addition to the trivial solution.

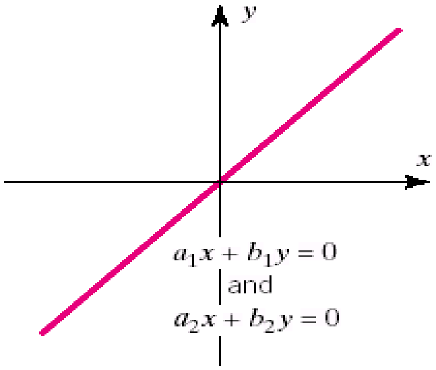
Homogeneous Linear Systems(2/2)

- In a special case of a homogeneous linear system of two linear equations in two unknowns: (fig1.2.1)

$$a_1x + b_1y = 0 \; (a_1, b_1 \text{ not both zero})$$
$$a_2x + b_2y = 0 \; (a_2, b_2 \text{ not both zero})$$



(*a*) Only the trivial solution



(*b*) Infinitely many solutions

Figure 1.2.1

Example 7
Gauss-Jordan Elimination(1/3)

- Solve the following homogeneous system of linear equations by using Gauss-Jordan elimination.

$$\begin{aligned} 2x_1 + 2x_2 - x_3 \phantom{{}-3x_4} + x_5 &= 0 \\ -x_1 - x_2 + 2x_3 - 3x_4 + x_5 &= 0 \\ x_1 + x_2 - 2x_3 \phantom{{}-3x_4} - x_5 &= 0 \\ x_3 + x_4 + x_5 &= 0 \end{aligned}$$

- Solution
  - The augmented matrix

$$\begin{bmatrix} 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & -1 & 2 & -3 & 1 & 0 \\ 1 & 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

  - Reducing this matrix to reduced row-echelon form

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Example 7
Gauss-Jordan Elimination(2/3)

Solution (cont)

- The corresponding system of equation

$$\begin{aligned} x_1 + x_2 \qquad\qquad\quad + x_5 &= 0 \\ x_3 \qquad + x_5 &= 0 \\ x_4 \qquad\quad &= 0 \end{aligned}$$

- Solving for the leading variables is

$$\begin{aligned} x_1 &= -x_2 - x_5 \\ x_3 &= -x_5 \\ x_4 &= 0 \end{aligned}$$

- Thus the general solution is

$$x_1 = -s - t\ ,\ \ x_2 = s,\ \ x_3 = -t,\ \ x_4 = 0,\ \ x_5 = t$$

- Note: the trivial solution is obtained when s=t=0.

Example7 Gauss-Jordan Elimination(3/3)

- Two important points:
  - Non of the three row operations alters the final column of zeros, so the system of equations corresponding to the reduced row-echelon form of the augmented matrix must also be a homogeneous system.
  - If the given homogeneous system has m equations in n unknowns with m<n, and there are r nonzero rows in reduced row-echelon form of the augmented matrix, we will have r<n. It will have the form:

$$
\begin{array}{llll}
\cdots x_{k1} & & + \sum () = 0 & \\
& \cdots x_{k2} & + \sum () = 0 & \\
& \ddots & \vdots & \\
& & x_r + \sum () = 0 & \quad (1)
\end{array}
\qquad
\begin{array}{ll}
x_{k1} = -\sum () & \\
x_{k2} = -\sum () & \\
\vdots & \\
x_r = -\sum () & \quad (2)
\end{array}
$$

Computer Solution of Linear System

- Most computer algorithms for solving large linear systems are based on Gaussian elimination or Gauss-Jordan elimination.
- Issues
  - Reducing roundoff errors
  - Minimizing the use of computer memory space
  - Solving the system with maximum speed

Chapter 2
Linear Equations

Gaussian elimination

A given linear system of equations has the form

$$Ax = b \, ,$$

Where

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

Easy to consider that the solution to $Ax = b$ can be written $x = A^{-1}b$

.

- The problem is how to obtain the inverse,

In fact, the inverse requires more arithmetic.

$$Ax=b, \quad \text{where} \quad A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

$$A \to A^{(1)} \to A^{(2)} \to \cdots \to A^{(n)}, \quad \text{and}$$

$$b \to b^{(1)} \to b^{(2)} \to \cdots \to b^{(n)}, \quad \text{where}$$

$$A^{(1)}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & a_{33}^{(1)} & \cdots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & a_{n3}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix},$$

$$
A^{(2)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix},
$$

$$
A^{(3)} = \begin{bmatrix} a_{11}^{(3)} & a_{12}^{(3)} & a_{13}^{(3)} & \cdots & a_{1n}^{(3)} \\ 0 & a_{22}^{(3)} & a_{23}^{(3)} & \cdots & a_{2n}^{(3)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(3)} \end{bmatrix}
$$

。。。 。。。 。。。

$A \to A^{(1)}$ means: $L_0 A = A^{(1)}$, where

$$L_0 = \begin{bmatrix} 1 & & & & & \\ -m_{21} & 1 & & & & \\ -m_{31} & & 1 & & & \\ -m_{41} & & & 1 & & \\ \vdots & & & & \ddots & \\ -m_{n1} & & & & & 1 \end{bmatrix}$$

For general we have

$$L_{j-1} = \begin{bmatrix} 1 & & & & & \\ -m_{2j} & 1 & & & & \\ -m_{3j} & & 1 & & & \\ -m_{4j} & & & 1 & & \\ \vdots & & & & \ddots & \\ -m_{nj} & & & & & 1 \end{bmatrix}, \qquad j = 1, 2, 3, \cdots, n$$

What is the m_{ij} ?

We will show you next!

The recursive relation

$$L_k A^{(k)} = A^{(k+1)} \qquad L_k b^{(k)} = b^{(k)}$$

Here $A^{(n)}$ is upper triangular, denotes U

$$L_{n-1} \cdots L_1 L_0 A = A^{(n)}$$

$$L_{n-1} \cdots L_1 L_0 b = b^{(n)}$$

Thus $$A = L_0^{-1} L_1^{-1} \cdots L_{n-1}^{-1} U = LU,$$

$$L = L_0^{-1} L_1^{-1} \cdots L_{n-1}^{-1}.$$

L is unit lower triangular.

Theorem on LU-Decomposition for matrix

If all n leading principal minors of the nxn matrix A are nonsingular, then A has an LU-decomposition.

- Suppose that A can be factored into the product of a lower triangular matrix L and an upper triangular matrix U: $A=LU$.

Then, to solve the system of equations $Ax=b$,

It is enough to solve this problem in two stages:

$Lz=b$ solve for z

$Ux=z$ solve for x

Here, solving these two triangular system is simple!

$$A = LU$$

An example for LU-decomposition

$$\begin{bmatrix} 6 & -2 & 2 & 4 \\ 12 & -8 & 6 & 10 \\ 3 & -13 & 9 & 3 \\ -6 & 4 & 1 & -18 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ 34 \\ 27 \\ -38 \end{bmatrix}$$

First step:

- Subtract 2 times the first equation from the second.
- Subtract ½ times the first equation from the third.
- Subtract -1 times the first equation from the fouth.
- recode

$$L = \begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ \frac{1}{2} & & 1 & \\ -1 & & & 1 \end{bmatrix}$$

Second step:

VECTOR SPACES AND SUBSPACES

- **Definition:** A **vector space** is a nonempty set *V* of objects, called *vectors*, on which are defined two operations, called *addition and multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors **u**, **v**, and **w** in *V* and for all scalars *c* and *d*.
  1. The sum of **u** and **v**, denoted by $\mathbf{u}+\mathbf{v}$, is in *V*.
  2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.
  3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.
  4. There is a zero vector **0** in *V* such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.

5. For each **u** in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
6. The scalar multiple of **u** by c, denoted by $c\mathbf{u}$, is in V.
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$.
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$.
9. $c(d\mathbf{u})=(cd)\mathbf{u}$.
10. $1\mathbf{u}=\mathbf{u}$.

Matrix and Vector Norms

How can we compare the size of vectors, matrices (and functions!)? For scalars it is easy (absolute value). The generalization of this concept to vectors, matrices and functions is called a **norm**. Formally the norm is a function from the space of vectors into the space of scalars denoted by

$$\boxed{\|(.)\|}$$

with the following properties:

Definition: Norms.

1. $||v|| > 0$ *for any* $v \in 0$ *and* $||v|| = 0$ *implies* $v=0$
2. $||av||=|a|\ ||v||$
3. $||u+v|| \leq ||v||+||u||$ *(Triangle inequality)*

We will only deal with the so-called l_p Norm.

Vector Norms

On a vector space V, a norm is a function $\|\bullet\|$ from the space V to the set of nonnegative reals that obeys these three postulates:

$$\|x\| > 0 \qquad if \qquad x \neq 0, \qquad x \in V$$

$$\|\lambda x\| = |\lambda| \|x\| \qquad if \qquad \lambda \in R, \qquad x \in V$$

$$\|x + y\| \leq \|x\| + \|y\| \qquad if \qquad x, y \in V \, (triangle - inequality)$$

We can think of $\|x\|$ as the length of the vector x

- The most familiar norm on R^n is the Euclidean $l_2 - norm$ defined by

$$\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{1/2}, where \quad x = (x_1, x_2, \ldots, x_n)^T$$

In numerical analysis, other norms also used.

$$\|x\|_\infty = \max_{1\le i\le n} |x_i|$$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

For example,

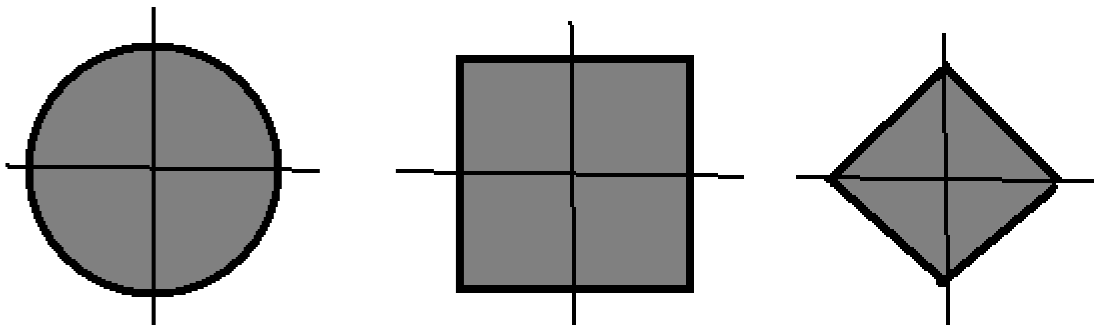
$$x = (0.2 \quad 0.4 \quad 0.6 \quad 0.8)$$

$$\|x\|_1 = 2$$

$$\|x\|_2 = 1.0954$$

$$\|x\|_\infty = 0.8$$

Unit cells in R^2 for three norms



$\|x\|_2 \leq 1$ $\qquad$ $\|x\|_\infty \leq 1$ $\qquad$ $\|x\|_1 \leq 1$

The l_p-Norm

The l_p- Norm for a vector x is defined as ($p \geq 1$):

$$\boxed{\|x\|_{l_p} = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}}$$

Examples:

- for p=2 we have the ordinary euclidian norm: $$\|x\|_{l_2} = \sqrt{x^T x}$$

- for p= ∞ the definition is $$\|x\|_{l_\infty} = \max_{1 \leq i \leq n} |x_i|$$

-

Condition number

- The condition number:

$$M = \max \frac{\|Ax\|}{\|x\|}, \qquad m = \min \frac{\|Ax\|}{\|x\|},$$

$$\kappa(A) = \frac{M}{n}$$

Matrix norm

$$x \in R^n, A \in R^{n\times n}$$

$$\|A\|_v = \max_{x\neq 0} \frac{\|Ax\|_v}{\|x\|_v}$$

Here , $\|x\|_v$ **is the norm of the vector**

$$v = 1, 2, \infty.$$

Easy to compute the matrix norms corresponding to the $l_\infty, l_1, l_2,$ vector norms

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

$\|A\|_\infty$ 、$\|A\|_1$ **----easy to compute,**

$\|A\|_2$ **----not easy to compute.**

However, It is helpful for theoretical analysis.

- 1 -2
- -3 4

$$\|Ax\| \leq \|A\|\|x\|,$$

$$\|AB\| \leq \|A\|\|B\|$$

$$\|x\| = 0 \Leftrightarrow x = o,$$

$$\|A\| = 0 \Leftrightarrow A = O$$

Sparse matrices

- **Sparsity:**

  How big the per-cent for the number of the element being zero in a matrix .

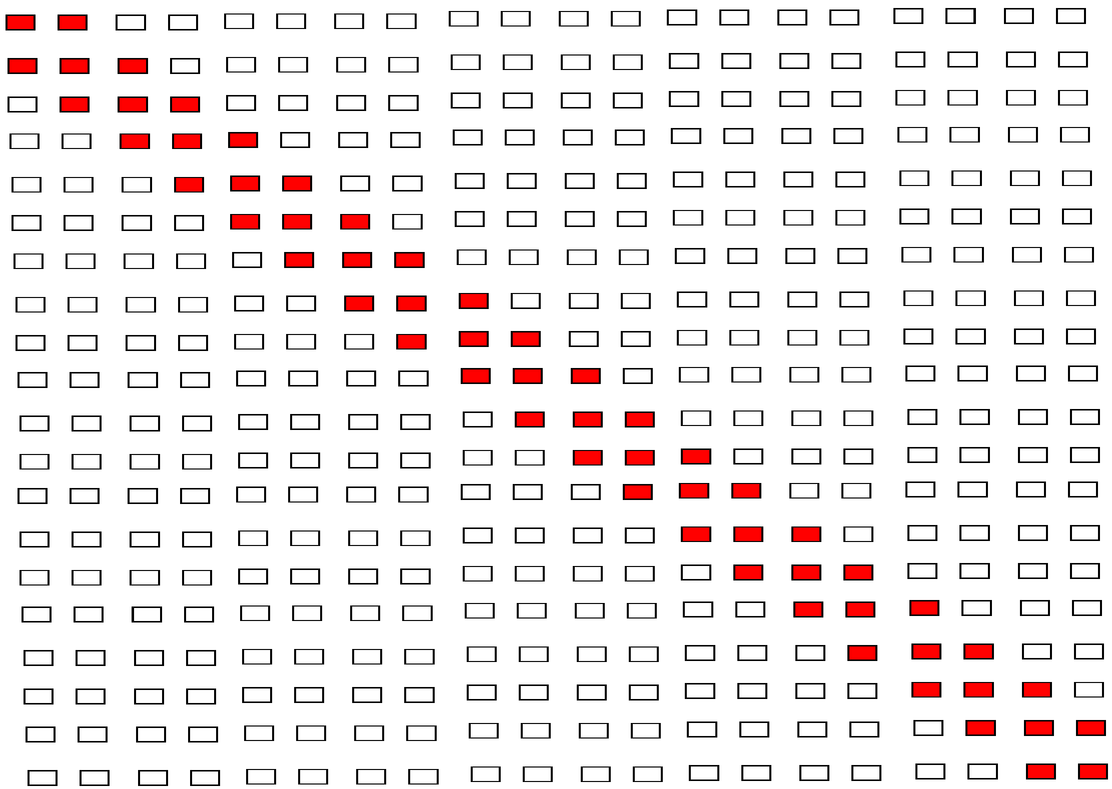
A **sparse matrix** is a matrix whose sparsity is nearly equal to 1.

A **band matrix** is a matrix whose bandwidth is small.

- ***bandwidth:***

  The bandwidth of a matrix is the maximum distance of the nonzero elements from the main diagonal

A matrix with bandwidth equal to 1 is a tridiagonal matrix, it is very important case!



By using Gaussian elimination to solve the special tridiaganal system, the algorithm is called (追赶法)。

Diagonally Dominant Matrix

- The matrix has the property that

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \qquad (1 \le i \le n)$$

- Example:

$$\begin{bmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 4 & 0 & -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 4 & 0 & 0 & 0 & -1 & 0 \\ 0 & -1 & -1 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 4 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 4 \end{bmatrix}$$

Iterative Refinement

- The notion of a convergent sequence of vectors

$$v^{(1)}, v^{(2)}, v^{(3)}, \cdots$$

means

$$\lim_{k \to \infty} \left\| v^{(k)} - v \right\| = 0$$

Example---------

Theorem:

- If A is an nxn matrix such that $\|A\| < 1$, then $I - A$ is invertible, and

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

Proof: If it is not invertible, then it is singular, and there exists a vector x satisfying $\|x\| = 1$,

and $(I - A)x = 0$ from this we have

$$1 = \|x\| = \|Ax\| \leq \|A\|\|x\| = \|A\|$$

(contradicts). Next, we will show that

$$\sum_{k=0}^{m} A^k \to (I-A)^{-1} \quad \Leftrightarrow \quad (I-A)\sum_{k=0}^{m} A^k \to I,$$

$$as \quad m \to \infty$$

$$(I-A)\sum_{k=0}^{m} A^k = \sum_{k=0}^{m} (A^k - A^{k+1})$$

$$= A^0 - A^{m+1} = I - A^{m+1},$$

$$\left\| A^{m+1} \right\| \le \left\| A \right\|^{m+1} \to 0, \quad as \quad m \to \infty$$

…………

From the theorem we have

$$\left\|(I-A)^{-1}\right\| \le \sum_{k=0}^{\infty}\left\|A^{k}\right\| \le \sum_{k=0}^{\infty}\left\|A\right\|^{k} = \frac{1}{1-\|A\|}$$

Solution of Equations
by Iterative Methods

Solution of Equations by Iterative Methods

- The Gaussian Algorithm is termed **Direct** method for solving the system $Ax = b$

It works through a finite number of steps and produce a solution x that would be completely accurate ,if it has not any roundoff errors.

An **indirect** method produces a sequence of vectors that converges to the solution.

Iterative---A simple process is applied repeatedly to generate the sequence.

The advantage of iterative methods is that they are usually stable.

- To convey the general idea, we describe two fundamental iterative methods.

$$\begin{cases} 7x_1 - 6x_2 = 3 \\ -8x_1 + 9x_2 = -4 \end{cases}$$

1.Jacobi method:

$$\begin{cases} 7x_1 - 6x_2 = 3 \\ -8x_1 + 9x_2 = -4 \end{cases} \Rightarrow \begin{cases} x_1 = \dfrac{6}{7}x_2 + \dfrac{3}{7} \\ x_2 = \dfrac{8}{9}x_1 - \dfrac{4}{9} \end{cases}$$

$$\begin{cases} x_1 = \dfrac{6}{7}x_2 + \dfrac{3}{7} \\ x_2 = \dfrac{8}{9}x_1 - \dfrac{4}{9} \end{cases} \Rightarrow \begin{cases} x_1^{(k)} = \dfrac{6}{7}x_2^{(k-1)} + \dfrac{3}{7} \\ x_2^{(k)} = \dfrac{8}{9}x_1^{(k-1)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots$$

- Initially when k=1, we select $x_1^{(0)}, x_2^{(0)},$

  Simply set them to 0.

then we can obtain $x_1^{(1)}, x_2^{(1)},$

Next, $x_1^{(2)}, x_2^{(2)};$ $x_1^{(3)}, x_2^{(3)};$ $\cdots\cdots$

2.Gauss-Seidel method

$$X=\begin{bmatrix}x_1\\x_2\end{bmatrix},\qquad X^{(k)}=\begin{bmatrix}x_1^{(k)}\\x_2^{(k)}\end{bmatrix}$$

$$\begin{cases}x_1=\dfrac{6}{7}x_2+\dfrac{3}{7}\\[2ex]x_2=\dfrac{8}{9}x_1-\dfrac{4}{9}\end{cases}\Rightarrow\begin{cases}x_1^{(k)}=\dfrac{6}{7}x_2^{(k-1)}+\dfrac{3}{7}\\[2ex]x_2^{(k)}=\dfrac{8}{9}x_1^{(k)}-\dfrac{4}{9}\end{cases}\qquad k=1,2,3,4,\cdots$$

- Jacobi:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + \dfrac{3}{7} \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k-1)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots \qquad \begin{matrix} x_1^{(1)} = \dfrac{3}{7}, \\ x_2^{(1)} = -\dfrac{4}{9}, \end{matrix}$$

- Seidel:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + \dfrac{3}{7} \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots \qquad \begin{matrix} x_1^{(1)} = \dfrac{3}{7}, \\ x_2^{(1)} = \dfrac{8}{9}(\dfrac{3}{7}) - \dfrac{4}{9} = -\dfrac{4}{9} + \dfrac{8}{21}, \end{matrix}$$

- Jacobi:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + 1 \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k-1)} - 1 \end{cases}$$

$$Let \qquad x_1^{(0)} = 0, \quad x_2^{(0)} = 0.$$

$$find \qquad x_1^{(1)} = ?, \quad x_2^{(1)} = ?$$

$$x_1^{(2)} = ?, \quad x_2^{(2)} = ?$$

- Seidel:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + 1 \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k)} - 1 \end{cases}$$

$$Let \qquad x_1^{(0)} = 0, \quad x_2^{(0)} = 0.$$

$$find \qquad x_1^{(1)} = ?, \quad x_2^{(1)} = ?$$

$$x_1^{(2)} = ?, \quad x_2^{(2)} = ?$$

The form as $x = Gx + c$

- The iteration formula:

$$x^{(k)} = Gx^{(k-1)} + c, \qquad k = 1, 2, 3, \cdots$$

to produce a sequence converging to $(I - G)^{-1}c$, for any starting vector $x^{(0)}$, it is **necessary and sufficient that the spectral radius of G be less than 1.**

- **Proof** Suppose that $\rho(G)<1.$

We have the following fact: $\rho(G)=\inf\|G\|,\ for\|\bullet\|.$

Means $\|G\|<1$.

Thus, $x^{(1)}=Gx^{(0)}+c$

$$x^{(2)}=G^2x^{(0)}+Gc+c$$

$$x^{(3)}=G^3x^{(0)}+G^2c+Gc+c$$

$$\cdots\cdots$$

$$x^{(k)}=G^kx^{(0)}+\sum_{j=0}^{k-1}G^jc \qquad *$$

$$\left\|G^kx^{(0)}\right\|\le\left\|G^k\right\|\left\|x^{(0)}\right\|\le\left\|G\right\|^k\left\|x^{(0)}\right\|\to 0,\quad as \quad k\to\infty$$

By the Theorem:

- If A is an nxn matrix such that $\|A\| < 1$, then

$I - A$ is invertible, and

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k$$

Here we have

$$\sum_{j=0}^{\infty} G^j c = (I - G)^{-1} c$$

By letting $k \to \infty$, and from *, we obtain

$$\lim_{k \to \infty} x^{(k)} = (I - G)^{-1} c$$

Suppose the converse that $\rho(G) \geq 1$,

select u and λ so that

$$Gu = \lambda u, \qquad \lambda \geq 1 \qquad u \neq 0$$

Let c=u and $x^{(0)} = 0,$

from $x^{(k)} = G^k x^{(0)} + \sum_{j=0}^{k-1} G^j c$ *

$$x^{(k)} = \sum_{j=0}^{k-1} \lambda^j c$$

$$diverges \qquad as \qquad \lambda = 1 or \lambda \neq 1$$

The form as $x = Gx + c$

- The iteration formula:

$$x^{(k)} = Gx^{(k-1)} + c, \qquad k = 1, 2, 3, \cdots$$

to produce a sequence converging to $(I - G)^{-1}c$, for any starting vector $x^{(0)}$, it is **necessary and sufficient that the spectral radius of G be less than 1.**

This is the principle to construct a iterative scheme for solving linear system!

- Jacobi:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + \dfrac{3}{7} \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k-1)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots \qquad \begin{aligned} x_1^{(1)} &= \frac{3}{7}, \\ x_2^{(1)} &= -\frac{4}{9}, \end{aligned}$$

- Seidel:

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + \dfrac{3}{7} \\ x_2^{(k)} = \dfrac{8}{9} x_1^{(k)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots \qquad \begin{aligned} x_1^{(1)} &= \frac{3}{7}, \\ x_2^{(1)} &= \frac{8}{9}(\frac{3}{7}) - \frac{4}{9} = -\frac{4}{9} + \frac{8}{21}, \end{aligned}$$

- Jacobi:

$$\begin{cases} x_1^{(k)} = \frac{6}{7} x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9} x_1^{(k-1)} - \frac{4}{9} \end{cases}$$

$$x^{(k)} = Gx^{(k-1)} + c$$

$$G = \begin{bmatrix} 0 & \frac{6}{7} \\ \frac{8}{9} & 0 \end{bmatrix}, \qquad c = \begin{bmatrix} \frac{3}{7} \\ -\frac{4}{9} \end{bmatrix}$$

- Seidel:

$$\begin{cases} x_1^{(k)} = \frac{6}{7} x_2^{(k-1)} + \frac{3}{7} \\ x_2^{(k)} = \frac{8}{9} x_1^{(k)} - \frac{4}{9} \end{cases}$$

$$k = 1, 2, 3, 4, \cdots$$

$$x_1^{(1)} = \frac{3}{7},$$

$$x_2^{(1)} = \frac{8}{9}\left(\frac{3}{7}\right) - \frac{4}{9} = -\frac{4}{9} + \frac{8}{21},$$

$$\begin{cases} x_1^{(k)} = \dfrac{6}{7} x_2^{(k-1)} + \dfrac{3}{7} \\[2ex] x_2^{(k)} = \dfrac{8}{9} x_1^{(k)} - \dfrac{4}{9} \end{cases} \qquad k = 1, 2, 3, 4, \cdots$$

$$x_1{}^{(k)} = \frac{6}{7} x_2{}^{(k-1)} + \frac{3}{7}$$

$$x_2{}^{(k)} = \frac{8}{9} x_1{}^{(k)} - \frac{4}{9} = \frac{8}{9}[\frac{6}{7} x_2{}^{(k-1)} + \frac{3}{7}]$$

$$= \frac{16}{21} x_2{}^{(k-1)} + (\frac{8}{21} - \frac{4}{9})$$

$$= \frac{16}{21}(\frac{8}{9} x_1{}^{(k-1)} - \frac{4}{9}) + (\frac{8}{21} - \frac{4}{9})$$

$$x^{(k)} = G x^{(k-1)} + c :$$

$$G = \begin{bmatrix} 0 & \dfrac{6}{7} \\[2ex] \dfrac{16}{21} \times \dfrac{8}{9} & 0 \end{bmatrix}$$

Iterative Methods for Solving Linear Systems of Equations

Iterative Methods

An iterative technique to solve **Ax=b** starts with an initial approximation $\mathbf{x}^{(0)}$ and generates a sequence $\left\{\mathbf{x}^{(k)}\right\}_{k=0}^{\infty}$

First we convert the system **Ax=b** into an equivalent form $\mathbf{x} = \mathbf{T}\mathbf{x} + \mathbf{c}$

And generate the sequence of approximation by

$$\boxed{\mathbf{x}^{(k)} = \mathbf{T}\mathbf{x}^{(k-1)} + \mathbf{c},\ k = 1,2,3\ldots}$$

This procedure is similar to the fixed point method.

The stopping criterion: $$\frac{\left\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\right\|}{\left\|\mathbf{x}^{(k)}\right\|} < \varepsilon$$

Iterative Methods (Example)

$$
\begin{aligned}
E_1 &: \quad 10x_1 - x_2 + 2x_3 = 6 \\
E_2 &: \quad -x_1 + 11x_2 - x_3 + 3x_4 = 25 \\
E_3 &: \quad 2x_1 - x_2 + 10x_3 - x_4 = -11 \\
E_4 &: \quad 3x_2 - x_3 + 8x_4 = 15
\end{aligned}
$$

We rewrite the system in the **x=Tx+c** form

$$
\begin{aligned}
x_1 &= \frac{1}{10}x_2 - \frac{1}{5}x_3 + \frac{3}{5} \\
x_2 &= \frac{1}{11}x_1 + \frac{1}{11}x_3 - \frac{3}{11}x_4 + \frac{25}{11} \\
x_3 &= -\frac{1}{5}x_1 + \frac{1}{10}x_2 + \frac{1}{10}x_4 - \frac{11}{10} \\
x_4 &= -\frac{3}{8}x_2 + \frac{1}{8}x_3 + \frac{15}{8}
\end{aligned}
$$

Iterative Methods (Example) – cont.

and start iterations with $\mathbf{x}^{(0)} = (0, 0, 0, 0)$

$$x_1^{(1)} = \frac{1}{10}x_2^{(0)} - \frac{1}{5}x_3^{(0)} + \frac{3}{5} = 0.6000$$

$$x_2^{(1)} = \frac{1}{11}x_1^{(0)} + \frac{1}{11}x_3^{(0)} - \frac{3}{11}x_4^{(0)} + \frac{25}{11} = 2.2727$$

$$x_3^{(1)} = -\frac{1}{5}x_1^{(0)} + \frac{1}{10}x_2^{(0)} + \frac{1}{10}x_4^{(0)} - \frac{11}{10} = -1.1000$$

$$x_4^{(1)} = -\frac{3}{8}x_2^{(0)} + \frac{1}{8}x_3^{(0)} + \frac{15}{8} = 1.8750$$

Continuing the iterations, the results are in the Table:

k	0	1	2	3	4	5	6	7	8	9	10
$x_1^{(k)}$	0.0000	0.6000	1.0473	0.9326	1.0152	0.9890	1.0032	0.9981	1.0006	0.9997	1.0001
$x_2^{(k)}$	0.0000	2.2727	1.7159	2.0533	1.9537	2.0114	1.9922	2.0023	1.9987	2.0004	1.9998
$x_3^{(k)}$	0.0000	−1.1000	−0.8052	−1.0493	−0.9681	−1.0103	−0.9945	−1.0020	−0.9990	−1.0004	−0.9998
$x_4^{(k)}$	0.0000	1.8750	0.8852	1.1309	0.9739	1.0214	0.9944	1.0036	0.9989	1.0006	0.9998

The Jacobi Iterative Method

The method of the Example is called the Jacobi iterative method

$$x_i^{(k)} = \frac{\sum\limits_{\substack{j=1 \\ j \neq i}} \left(-a_{ij} x_j^{(k-1)} \right) + b_i}{a_{ii}}, \qquad i = 1, 2, \ldots., n$$

Algorithm: Jacobi Iterative Method

Jacobi Iterative

To solve $A\mathbf{x} = \mathbf{b}$ given an initial approximation $\mathbf{x}^{(0)}$:

INPUT the number of equations and unknowns n; the entries a_{ij}, $1 \le i, j \le n$ of the matrix A; the entries b_i, $1 \le i \le n$ of $\mathbf{b}$; the entries XO_i, $1 \le i \le n$ of $\mathbf{XO} = \mathbf{x}^{(0)}$; tolerance TOL; maximum number of iterations N.

OUTPUT the approximate solution $x_1, \ldots, x_n$ or a message that the number of iterations was exceeded.

Step 1 Set $k = 1$.

Step 2 While ($k \le N$) do Steps 3–6.

Step 3 For $i = 1, \ldots, n$

$$\text{set } x_i = \frac{-\sum_{\substack{j=1 \\ j \ne i}}^{n} (a_{ij} XO_j) + b_i}{a_{ii}}.$$

Step 4 If $||\mathbf{x} - \mathbf{XO}|| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$;
(*Procedure completed successfully.*)
STOP.

Step 5 Set $k = k + 1$.

Step 6 For $i = 1, \ldots, n$ set $XO_i = x_i$.

Step 7 OUTPUT ('Maximum number of iterations exceeded');
(*Procedure completed unsuccessfully.*)
STOP.

The Jacobi Method: x=Tx+c Form

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & & . \\ . & . & & . \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} =$$

$$= \underbrace{\begin{bmatrix} a_{11} & 0 \ldots\ldots\ldots\ldots 0 \\ 0 \quad a_{22} & \ldots\ldots\ldots\ldots 0 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots 0 \\ 0 \ldots\ldots\ldots\ldots 0 & a_{nn} \end{bmatrix}}_{\mathbf{D}} - \underbrace{\begin{bmatrix} 0 \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots 0 \\ -a_{21} & \ldots\ldots\ldots\ldots 0 \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ -a_{n1} \ldots\ldots -a_{n,n-1} & 0 \end{bmatrix}}_{-\mathbf{L}} - \underbrace{\begin{bmatrix} 0 & -a_{12} \ldots\ldots -a_{1n} \\ 0 & \ldots\ldots\ldots\ldots -a_{2n} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots & \ldots\ldots -a_{n-1,n} \\ 0 \ldots\ldots\ldots\ldots & \ldots\ldots 0 \end{bmatrix}}_{-\mathbf{U}}$$

$$\boxed{\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}}$$

The Jacobi Method: x=Tx+c Form (cont)

$$\mathbf{A} = \mathbf{D} - \mathbf{L} - \mathbf{U}$$

and the equation **Ax=b** can be transformed into

$$(\mathbf{D} - \mathbf{L} - \mathbf{U})\mathbf{x} = \mathbf{b}$$

$$\mathbf{Dx} = (\mathbf{L} + \mathbf{U})\mathbf{x} + \mathbf{b}$$

$$\mathbf{x} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U})\mathbf{x} + \mathbf{D}^{-1}\mathbf{b}$$

Finally

$$\mathbf{T} = \mathbf{D}^{-1}(\mathbf{L} + \mathbf{U}) \qquad \mathbf{c} = \mathbf{D}^{-1}\mathbf{b}$$

The Gauss-Seidel Iterative Method

The idea of GS is to compute $\mathbf{x}^{(k)}$ using most recently calculated values. In our example:

$$
\begin{aligned}
x_1^{(k)} &= \frac{1}{10}x_2^{(k-1)} - \frac{1}{5}x_3^{(k-1)} + \frac{3}{5} \\
x_2^{(k)} &= \frac{1}{11}x_1^{(k)} + \frac{1}{11}x_3^{(k-1)} - \frac{3}{11}x_4^{(k-1)} + \frac{25}{11} \\
x_3^{(k)} &= -\frac{1}{5}x_1^{(k)} + \frac{1}{10}x_2^{(k)} + \frac{1}{10}x_4^{(k-1)} - \frac{11}{10} \\
x_4^{(k)} &= -\frac{3}{8}x_2^{(k)} + \frac{1}{8}x_3^{(k)} + \frac{15}{8}
\end{aligned}
$$

Starting iterations with $\mathbf{x}^{(0)} = (0, 0, 0, 0)$, we obtain

k	0	1	2	3	4	5
$x_1^{(k)}$	0.0000	0.6000	1.030	1.0065	1.0009	1.0001
$x_2^{(k)}$	0.0000	2.3272	2.037	2.0036	2.0003	2.0000
$x_3^{(k)}$	0.0000	−0.9873	−1.014	−1.0025	−1.0003	−1.0000
$x_4^{(k)}$	0.0000	0.8789	0.9844	0.9983	0.9999	1.0000

The Gauss-Seidel Iterative Method

$$x_i^{(k)} = \frac{-\sum_{j=1}^{i-1}\left(a_{ij}x_j^{(k)}\right) - \sum_{j=i+1}^{n}\left(a_{ij}x_j^{(k-1)}\right) + b_i}{a_{ii}}, \qquad i = 1, 2, \ldots, n$$

Gauss-Seidel in $\mathbf{x}^{(k)} = \mathbf{T}\mathbf{x}^{(k-1)} + \mathbf{c}$ form (the Fixed Point)

$$\mathbf{Ax} = (\mathbf{D} - \mathbf{L} - \mathbf{U})\mathbf{x} = \mathbf{b}$$

$$(\mathbf{D} - \mathbf{L})\mathbf{x} = \mathbf{Ux} + \mathbf{b}$$

$$(\mathbf{D} - \mathbf{L})\mathbf{x}^{(k)} = \mathbf{Ux}^{(k-1)} + \mathbf{b}$$

Finally

$$\boxed{\mathbf{x}^{(k)} = \underbrace{(\mathbf{D} - \mathbf{L})^{-1}\mathbf{U}}_{\mathbf{T}}\mathbf{x}^{(k-1)} + \underbrace{(\mathbf{D} - \mathbf{L})^{-1}\mathbf{b}}_{\mathbf{c}}}$$

Algorithm: Gauss-Seidel Iterative Method

Gauss-Seidel Iterative

To solve $A\mathbf{x} = \mathbf{b}$ given an initial approximation $\mathbf{x}^{(0)}$:

INPUT the number of equations and unknowns n; the entries a_{ij}, $1 \le i, j \le n$ of the matrix A; the entries b_i, $1 \le i \le n$ of $\mathbf{b}$; the entries XO_i, $1 \le i \le n$ of $\mathbf{XO} = \mathbf{x}^{(0)}$; tolerance TOL; maximum number of iterations N.

OUTPUT the approximate solution $x_1, \ldots, x_n$ or a message that the number of iterations was exceeded.

Step 1 Set $k = 1$.

Step 2 While ($k \le N$) do Steps 3–6.

Step 3 For $i = 1, \ldots, n$

$$\text{set } x_i = \frac{-\sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}XO_j + b_i}{a_{ii}}.$$

Step 4 If $||\mathbf{x} - \mathbf{XO}|| < TOL$ then OUTPUT $(x_1, \ldots, x_n)$;
(*Procedure completed successfully.*)
STOP.

Step 5 Set $k = k + 1$.

Step 6 For $i = 1, \ldots, n$ set $XO_i = x_i$.

Step 7 OUTPUT ('Maximum number of iterations exceeded');
(*Procedure completed unsuccessfully.*)
STOP.

Conjugate Gradient Methods

(special methods for solving Ax=b)

- For the case when A is a nXn, real, symmetric, and positive definite matrix.
- In the case, the problem of solving Ax=b is equivalent to the problem of minimizing the quadratic form

$$q(x) = (x, Ax) - 2(x, b).$$

- **Proof** we consider x+tv, where x, v are vectors, and t is a scalar.

$$q(x+tv) = (x+tv, A(x+tv)) - 2(x+tv, b)$$

$$= ......$$

$$**$$

$$= q(x) + 2t(v, Ax-b) + t^2(v, Av)$$

For any $v \neq 0, (v, Av) > 0$, so the quadratic function has a minimum. Compute the derivative with respect to t as

$$\frac{d}{dt} q(x+tv) = 2(v, Ax-b) + 2t(v, Av) \overset{!}{=} 0$$

$$\Rightarrow \hat{t} = \frac{(v, b-Ax)}{(v, Av)},$$

$$q(x+\hat{t}v) = \ldots\ldots$$

$$= q(x) - \frac{(v, b-Ax)^2}{(v, Av)}$$

Start from $x^{(k)}$, and a suitable search direction $v^{(k)}$ is chosen, the next point in the sequence is

$$x^{(k+1)} = x^{(k)} + t_k v^{(k)}$$

$$where \qquad t_k = \frac{(v^{(k)}, b - ax^{(k)})}{(v^{(k)}, Av^{(k)})},$$

$$k = 0, 1, 2, 3, 4, \cdots$$