Solutions of Linear Equation Systems

Introduction

- Many engineering and scientific problems can be formulated in terms of systems of simultaneous linear equations.
- When these systems consist of only a few equations, a solution can be found analytically using the standard methods from algebra, such as substitution.
- However, complex problems may involve a large number of equations that cannot realistically be solved using analytical methods.
- In these cases, we will need to find the solution numerically using computers.

Example: Material Purchasing for Manufacturing

- Let us assume that a manufacturer is marketing a product made of an alloy material meeting a certain specified composition.
- The three critical ingredients of the alloy are manganese, silicon, and copper. The specifications require 15 pounds of manganese, 22 pounds of silicon, and 39 pounds of copper for each ton of alloy to be produced.
- This mix of ingredients requires the manufacturer to obtain inputs from three different mining suppliers.
- Ore from the different suppliers has different concentrations of the alloy ingredients, as detailed in Table.
- Given this information, the manufacturer must determine the quantity of ore to purchase from each supplier so that the alloy ingredients are not wasted.

Solution

TABLE 5.1 Concentration of Alloy Ingredients for Three Suppliers

	Supplier 1 (lb/ton of Ore)	Supplier 2 (lb/ton of Ore)	Supplier 3 (lb/ton of Ore)
Manganese	1	3	2
Silicon	2	4	3
Copper	3	4	7

X_j = amount of ore purchased from supplier j

C_i = amount of ingredient i required per ton of alloy

a_{ij} = amount of ingredient i contained in each ton of ore shipped from supplier j

Using these notations, we can formulate a general equation that defines

(1) the relationships among the compositions of the ore shipped by the different suppliers,
(2) the amount of ore needed from each supplier, and
(3) the required composition of the final alloy as

$$\sum_{j=1}^{n} a_{ij} X_j = C_i \qquad \text{for } i = 1, 2, \ldots, m$$

in which *m* is the number of ingredients and *n* is the number of suppliers. For the case under consideration, both *m* and *n* equal 3·

$$X_1 + 3X_2 + 2X_3 = 15$$

$$2X_1 + 4X_2 + 3X_3 = 22$$

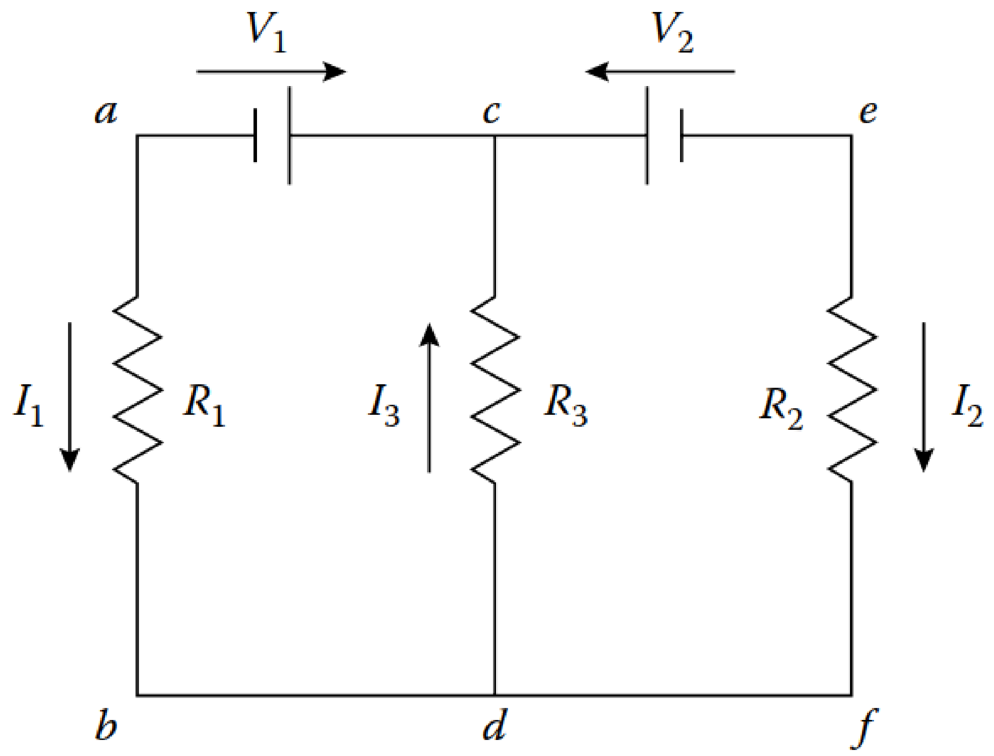
$$3X_1 + 4X_2 + 7X_3 = 39$$

Example 2: Electrical Circuit Analysis

Current flows in circuits are governed by Kirchhoff's laws.

- Kirchhoff's first law states that the algebraic sum of the currents flowing into a junction of a circuit must equal zero.
- Kirchhoff's second law states that the algebraic sum of the electromotive forces around a closed circuit must equal the sum of the voltage drops around the circuit, where a voltage drop equals the product of the current and the resistance·

Example 2: Electrical Circuit Analysis



Applying Kirchhoff's first law at junction *c*

$$I_1 + I_2 - I_3 = 0$$

Applying Kirchhoff's second law to network loop *acdb*

$$V_1 = R_1 I_1 + R_3 I_3$$

Applying Kirchhoff's second law to network loop *aefb*

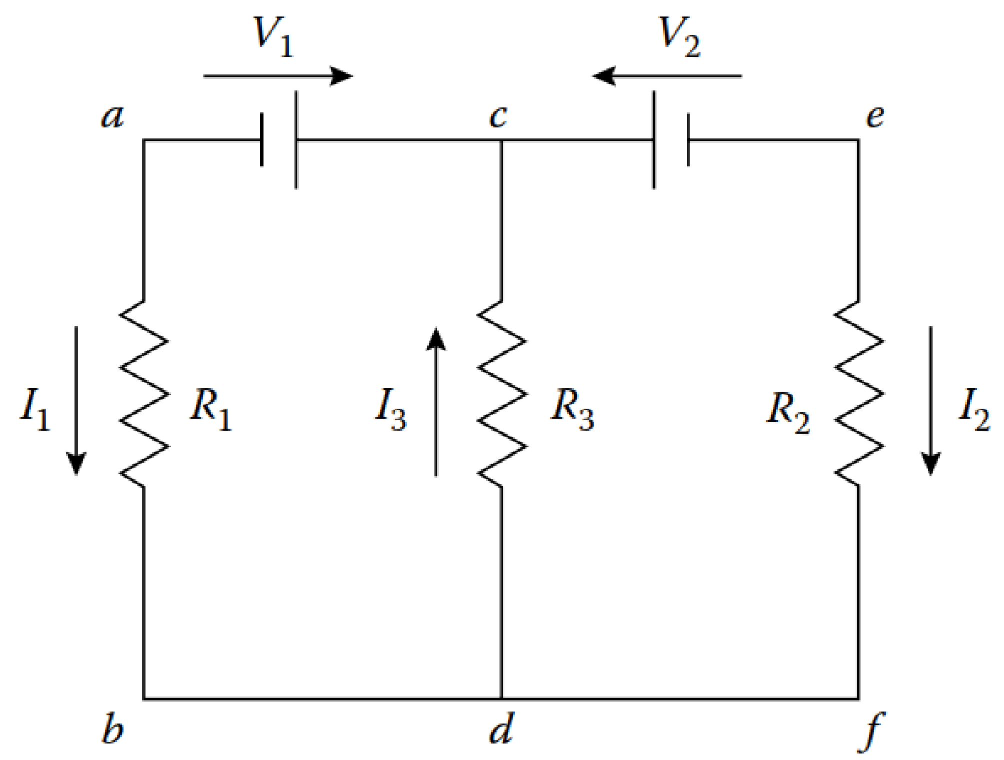
$$V_1 - V_2 = R_1 I_1 - R_2 I_2$$

Assume that $R_1 = 2$, $R_2 = 4$, $R_3 = 5$, $V_1 = 6$, and $V_2 = 2$

$$I_1 + I_2 - I_3 = 0$$

$$2I_1 + 5I_3 = 6$$

$$2I_1 - 4I_2 = 4$$



General Form For A System of Equations

$$a_{11}X_1 + a_{12}X_2 + \cdots + a_{1n}X_n = C_1$$
$$a_{21}X_1 + a_{22}X_2 + \cdots + a_{2n}X_n = C_2$$
$$\vdots$$
$$a_{n1}X_1 + a_{n2}X_2 + \cdots + a_{nn}X_n = C_n$$

in which the a_{ij} terms are the **known coefficients** of the equations, the X_j terms are the **unknown variables**, and the C_i terms are the **known constants**.

Since values for both the a_{ij} and C_i terms will be known for any problem, the system of equations represents *n* linear equations with *n* unknowns.

Iterative Equation-Solving Methods

- Linear equations can be solved by
  - ***Direct* equation-solving methods** like the Gaussian elimination method
    - the solution is found after a fixed, predictable number of operations
  - A trial-and-error procedure or ***iterative* methods**.
    - the number of operations required to obtain a solution is not fixed
    - a major advantage of iterative methods is that they can be used to solve nonlinear simultaneous equations, a task that is not possible using direct elimination methods

Two of the most common methods,

- **The Jacobi and**
- **Gauss– Seidel procedures**

Jacobi Iteration

Idea: $ax = b$ for a single linear equation with a single unknown, it is straightforward to solve for the unknown $x = \frac{b}{a}$

what if: $a_{11}x_1 + a_{12}x_2 = b_1$ **solve eq 1 for x_1** $x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2)$

$a_{21}x_1 + a_{22}x_2 = b_2$ **solve eq 2 for x_2** $x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1)$

But, we need values for x_1 and x_2 on the right-hand-side...

Jacobi Iteration

What if we start with **guesses** for those values, call those guesses: x_1^0 and x_2^0

$$x_1 = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^0\right) \qquad x_2 = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^0\right)$$

The equations above act as an "update" or "correction" to the guesses. To make things clearer, rename the updates to: x_1^1 and x_2^1

$$x_1^1 = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^0\right) \qquad x_2^1 = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^0\right)$$

Now repeat!

Jacobi Iteration

$$x_1^1 = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^0\right) \qquad x_2^1 = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^0\right)$$

Now repeat!

$$x_1^2 = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^1\right) \qquad x_2^2 = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^1\right)$$

$$x_1^3 = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^3\right) \qquad x_2^3 = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^3\right)$$

$$\vdots \qquad \vdots$$

k is an iteration counter

$$x_1^{k+1} = \frac{1}{a_{11}}\left(b_1 - a_{12}x_2^k\right) \qquad x_2^{k+1} = \frac{1}{a_{22}}\left(b_2 - a_{21}x_1^k\right)$$

Jacobi Iteration Generalization

Example: 3 linear equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

$$x_1^{k+1} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^k - a_{13}x_3^k) \longrightarrow x_1^{k+1} = x_1^k + \frac{1}{a_{11}}(b_1 - a_{11}x_1^k - a_{12}x_2^k - a_{13}x_3^k)$$

$$x_2^{k+1} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^k - a_{23}x_3^k) \longrightarrow x_2^{k+1} = x_2^k + \frac{1}{a_{22}}(b_2 - a_{21}x_1^k - a_{22}x_2^k - a_{23}x_3^k)$$

$$x_3^{k+1} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^k - a_{32}x_2^k) \longrightarrow x_3^{k+1} = x_3^k + \frac{1}{a_{33}}(b_2 - a_{31}x_1^k - a_{32}x_2^k - a_{33}x_3^k)$$

Jacobi Iteration Generalization

In general, for an $n \times n$ system

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

Jacobi Iteration Generalization

Jacobi Algorithm:

1. Guess x_i
2. Update guess for each x_i, starting with $i = 0$ according to

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

3. Repeat step 2 until iteration count has been reached

or

The acceptable difference is set by the user and influenced by the need for accuracy·

Example: Jacobi Iteration

Given the following set of equations, solve for values of the unknowns using Jacobi iteration:

$$3X_1 + X_2 - 2X_3 = 9$$

$$-X_1 + 4X_2 - 3X_3 = -8$$

$$X_1 - X_2 + 4X_3 = 1$$

Solution

$$X_1 = \frac{9 - X_2 + 2X_3}{3}$$

$$X_2 = \frac{-8 + X_1 + 3X_3}{4}$$

$$X_3 = \frac{1 - X_1 + X_2}{4}$$

values of X_1 = X_2 = X_3 = 1 for this initial estimate are assumed.

$$X_1 = \frac{9 - 1 + 2(1)}{3} = \frac{10}{3}$$

$$X_2 = \frac{-8 + 1 + 3(1)}{4} = -1$$

$$X_3 = \frac{1 - 1 + 1}{4} = \frac{1}{4}$$

These new values for X_1, X_2, and X_3 are then used as the new solution estimate.

$$X_1 = \frac{9-(-1)+2\left(\frac{1}{4}\right)}{3} = \frac{7}{2}$$

$$X_2 = \frac{-8+\frac{10}{3}+3\left(\frac{1}{4}\right)}{4} = -\frac{47}{48}$$

$$X_3 = \frac{1-\frac{10}{3}+(-1)}{4} = -\frac{5}{6}$$

This process is repeated until the differences between the previous values and the new values are small.

Iteration	X_1	$\|\Delta X_1\|$	X_2	$\|\Delta X_2\|$	X_3	$\|\Delta X_3\|$
0	1	–	1	–	1	–
1	3.333	2.333	–1.000	2.000	0.250	0.750
2	3.500	0.167	–0.979	0.021	–0.833	1.083
3	2.771	0.729	–1.750	0.771	–0.870	0.036
4	3.003	0.233	–1.960	0.210	–0.880	0.010
5	3.066	0.063	–1.909	0.050	–0.991	0.111
6	2.976	0.090	–1.976	0.067	–0.994	0.003
7	2.996	0.020	–2.001	0.025	–0.988	0.006
8	3.008	0.012	–1.992	0.009	–0.999	0.011
9	2.998	0.011	–1.997	0.005	–1.000	0.001
10	2.999	0.001	–2.001	0.003	–0.999	0.001
11	3.001	0.002	–1.999	0.001	–1.000	0.001
12	3.000	0.001	–2.000	0.000	–1.000	0.001

If a maximum absolute change of less than 0.05

Using a fixed number of iterations can be inefficient.

We need a way to tell the solver to stop the iterations.

Convergence: When to Stop Iterating?

Successive calculations (iteration) continue until the tolerance value (TD) is satisfied

$$\left|x_i^{(k+1)} - x_i^{(k)}\right| \leq TD \qquad (i = 1,2,\ldots,n)$$

Gauss–Seidel Iteration (Multi-Step Iteration)

It is quite similar to the Jacobi method.

- The only difference is; Substituting the calculated x_i value into the next equation.

Idea: Always use most recent information.

Gauss-Seidel Algorithm

1. Guess x_i.
2. Update guess for each x_i, starting with i=1. Use most recent information when calculating each x_i.

   Gauss-Seidel Update: $x_i = x_i + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{n} a_{ij}x_j\right)$
3. Repeat step 2 until convergence is obtained.

Example

Solve: the following system of equations by taking TD = 0.05

An initial solution estimate of $X_1 = X_2 = X_3 = 1$

$$-X_1 + 4X_2 - 3X_3 = -8 \longrightarrow X_1 = 8 + 4X_2 - 3X_3$$

$$3X_1 + X_2 - 2X_3 = 9 \longrightarrow X_2 = 9 - 3X_1 + 2X_3$$

$$X_1 - X_2 + 4X_3 = 1 \longrightarrow X_3 = \frac{1 - X_1 + X_2}{4}$$

Divergence of Gauss–Seidel Iteration

Iteration	X_1	$\lvert\Delta X_1\rvert$	X_2	$\lvert\Delta X_2\rvert$	X_3	$\lvert\Delta X_3\rvert$
0	1	–	1	–	1	–
1	9	8	–16	17	–6	7
2	–38	47	111	127	37.5	43.5
3	339.5	377.5	–934.5	1045.5	–318.25	355.75

If we re-arrange equations like

$$-X_1 + 4X_2 - 3X_3 = -8 \quad \rightarrow \quad 3X_1 + X_2 - 2X_3 = 9$$

$$3X_1 + X_2 - 2X_3 = 9 \quad \rightarrow \quad -X_1 + 4X_2 - 3X_3 = -8$$

$$X_1 - X_2 + 4X_3 = 1 \quad \rightarrow \quad X_1 - X_2 + 4X_3 = 1$$

The first iteration cycle

$$X_1 = \frac{9 - X_2 + 2X_3}{3}$$

$$X_2 = \frac{-8 + X_1 + 3X_3}{4}$$

$$X_3 = \frac{1 - X_1 + X_2}{4}$$

$$X_1 = \frac{9 - 1 + 2(1)}{3} = 3.333$$

$$X_2 = \frac{-8 + 3.333 + 3(1)}{4} = -0.417$$

$$X_3 = \frac{1 - 3.333 + (-0.417)}{4} = -0.688$$

The second iteration cycle

$$X_1 = \frac{9-(-0.417)+2(-0.688)}{3} = 2.680$$

$$X_2 = \frac{-8+2.680+3(-0.688)}{4} = -1.845$$

$$X_3 = \frac{1-2.680+(-1.845)}{4} = -0.882$$

Iteration	X_1	$\lvert\Delta X_1\rvert$	X_2	$\lvert\Delta X_2\rvert$	X_3	$\lvert\Delta X_3\rvert$
0	1	–	1	–	1	–
1	3.333	2.333	–0.417	1.417	–0.688	1.688
2	2.680	0.348	–1.845	1.428	–0.882	0.194
3	3.027	0.346	–1.904	0.059	–0.983	0.101
4	2.979	0.048	–1.992	0.088	–0.993	0.010
5	3.002	0.023	–1.994	0.002	–0.999	0.006
6	2.999	0.003	–2.000	0.006	–1.000	0.001
7	3.000	0.001	–2.000	0.000	–1.000	0.000
8	3.000	0.000	–2.000	0.000	–1.000	0.000

PIVOTING

- ***Pivoting*** is the displacement of rows in the coefficient matrix so that the diagonal elements are maximized in absolute value.

Row Pivoting

$$\begin{bmatrix} 0 & 1 & 3 \\ 0 & 2 & 1 \\ 4 & 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$\begin{bmatrix} 4 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 1 & 3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_3 \\ b_2 \\ b_1 \end{pmatrix}$$

Column Pivoting

$$\begin{bmatrix} 0 & 1 & 3 \\ 0 & 2 & 1 \\ 4 & 1 & 2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

$$\begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 2 & 1 & 4 \end{bmatrix} \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

MATRICES

The case of determining the values $x_1, x_2, \ldots, x_n$ that simultaneously satisfy a set of equations

(General set of equations)

$$
\begin{aligned}
f_1(x_1, x_2, \ldots, x_n) &= 0 \\
f_2(x_1, x_2, \ldots, x_n) &= 0 \\
\vdots \qquad & \quad \vdots \\
f_n(x_1, x_2, \ldots, x_n) &= 0
\end{aligned}
$$

- Such systems can be either *linear* or *nonlinear*. *Linear algebraic equations* that are of the general form

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

where

the a's are constant coefficients,

the b's are constants, and

n is the number of equations.

- The system of linear equations given can be represented in matrix form:

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

$$[A]\{x\} = \{b\}$$

where $[A]$ is *n x n* Coefficient matrix

$\{x\}$ is *n x 1* Unknown vector

$\{b\}$ is *n x 1* Right-hand side (RHS) vector

4.2 Matrix

- A *matrix* consists of a rectangular array of elements represented by a single symbol.
- $[A]$ is the shorthand notation for the matrix and a_{ij} designates an individual *element* of the matrix.
- A horizontal set of elements is called a *row* and a vertical set is called a *column*. The first subscript i always designates the number of the row in which the element lies. The second subscript j designates the column.

Column 3

$n \times m$ Matrix

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{bmatrix} \leftarrow \text{Row 2}$$

Row Vector: Matrices with row dimension n = 1, such as

$$[B] = [b_1 \quad b_2 \quad \cdots \quad b_m]$$

Column Vector: Matrices with column dimension n = 1, such as

$$[C] = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}$$

Square Matrix: Matrices where n = m are called *square matrices*. For example, a 4 by 4 matrix is

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

The diagonal consisting of the elements a_{11}, a_{22}, a_{33}, and a_{44} is termed the *principal* or *main diagonal* of the matrix.

4.2.1 Special Types of Square Matrix

- A ***symmetric matrix*** is one where $a_{ij} = a_{ji}$ for all i's and j's. For example, $[A] = \begin{bmatrix} 5 & 1 & 2 \\ 1 & 3 & 7 \\ 2 & 7 & 8 \end{bmatrix}$ is a 3 by 3 symmetric matrix.

- A ***diagonal matrix*** is a square matrix where all elements off the main diagonal are equal to zero, as in

$$[A] = \begin{bmatrix} a_{11} & & & \\ & a_{22} & & \\ & & a_{33} & \\ & & & a_{44} \end{bmatrix}$$

Note that where large blocks of elements are zero, they are left blank.

- An ***identity matrix*** is a diagonal matrix where all elements on the main diagonal are equal to 1, as in

$$[A] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

The symbol $[I]$ is used to denote the identity matrix. The identity matrix has properties similar to unity.

- An ***upper triangular matrix*** is one where all the elements below the main diagonal are zero, as in

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ & a_{22} & a_{23} & a_{24} \\ & & a_{33} & a_{34} \\ & & & a_{44} \end{bmatrix}$$

- An ***lower triangular matrix*** is one where all the elements above the main diagonal are zero, as in

$$[A] = \begin{bmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ a_{31} & a_{32} & a_{33} & \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

- A ***banded matrix*** has all elements equal to zero, except for a band centered on the main diagonal:

$$[A] = \begin{bmatrix} a_{11} & a_{12} & & \\ a_{21} & a_{22} & a_{23} & \\ & a_{32} & a_{33} & a_{34} \\ & & a_{43} & a_{44} \end{bmatrix}$$

4.2.2 Matrix Operations

- ***Addition*** of two matrices, say, [A] and [B], is accomplished by adding corresponding terms in each matrix. The elements of the resulting matrix [C] are computed,

$$c_{ij} = a_{ij} + b_{ij} \qquad \text{for } i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, m.$$

- Similarly, the ***subtraction*** of two matrices, say, [E] minus [F], is obtained by subtracting corresponding terms, as in

$$d_{ij} = e_{ij} - f_{ij} \qquad \text{for } i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, m.$$

- Addition and subtraction can be performed only between matrices having the same dimensions.
- Both addition and subtraction are *commutative*:

$$[A] + [B] = [B] + [A]$$

- Addition and subtraction are also *associative*, that is,

$$([A] + [B]) + [C] = [A] + ([B] + [C])$$

- The ***multiplication*** of a matrix [*A*] by a scalar *g* is obtained by multiplying every element of [*A*] by *g*, as in

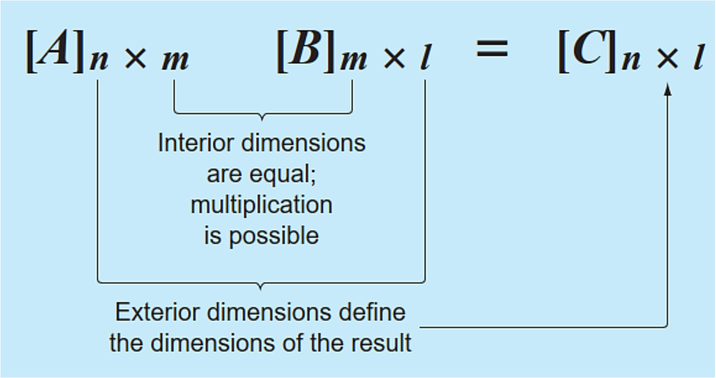
$$[D] = g[A] = \begin{bmatrix} ga_{11} & ga_{12} & \cdots & ga_{1m} \\ ga_{21} & ga_{22} & \cdots & ga_{2m} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ ga_{n1} & ga_{n2} & \cdots & ga_{nm} \end{bmatrix}$$

- The ***product*** of two matrices is represented as [*C*] = [*A*][*B*], where the elements of [*C*] are defined as

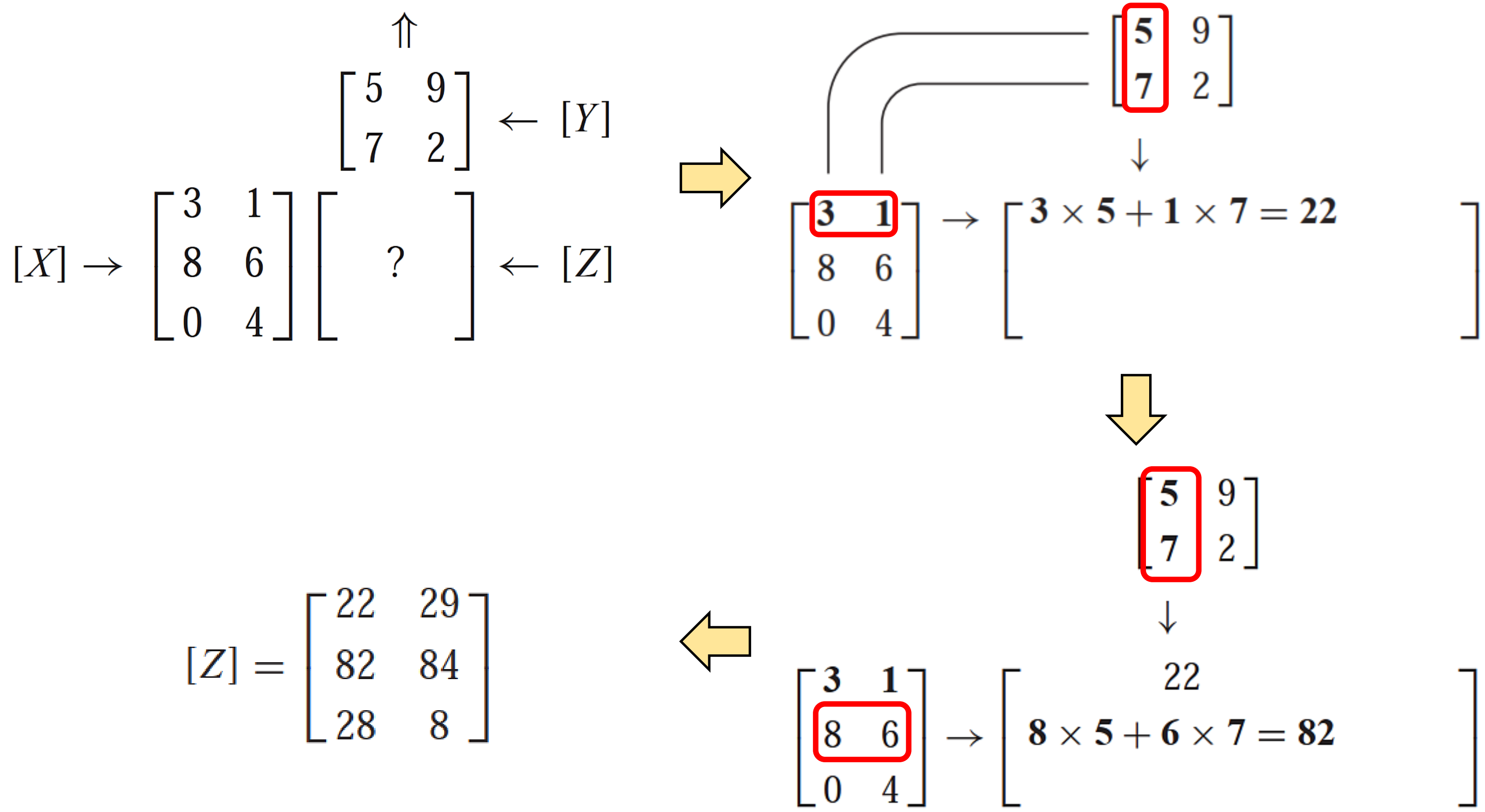
$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

where *n* = the column dimension of [*A*] and the row dimension of [*B*]. That is, the c_{ij} element is obtained by adding the product of individual elements from the *i*th row of the first matrix, in this case [*A*], by the *j*th column of the second matrix [*B*].

- According to this definition, multiplication of two matrices can be performed *only if the first matrix has as many columns as the number of rows in the second matrix.*



- Suppose that we want to multiply $[X]$ by $[Y]$ to yield $[Z]$, $[Z]=[X][Y]=\begin{bmatrix}3 & 1\\ 8 & 6\\ 0 & 4\end{bmatrix}\begin{bmatrix}5 & 9\\ 7 & 2\end{bmatrix}$
- A simple way to visualize the computation of $[Z]$ is to raise $[Y]$, as in



- If the dimensions of the matrices are suitable, matrix multiplication is *associative*,

$$([A][B])[C] = [A]([B][C])$$

- and *distributive*,

$$[A]([B] + [C]) = [A][B] + [A][C]$$

or

$$([A] + [B])[C] = [A][C] + [B][C]$$

- However, multiplication is *not generally commutative*:

$$[A][B] \neq [B][A]$$

- The ***transpose of a matrix*** involves transforming its rows into columns and its columns into rows.

$$[A] = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \Rightarrow [A]^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix}$$

- In other words, the element a_{ij} of the transpose is equal to the a_{ji} element of the original matrix.
- The ***trace*** of a matrix is the sum of the elements on its principal diagonal. It is designated as tr $[A]$ and is computed as

$$\text{tr}\,[A] = \sum_{i=1}^{n} a_{ii}$$

- If a matrix $[A]$ is square and nonsingular, there is another matrix $[A]^{-1}$, called the ***inverse*** of $[A]$, for which

$$[A][A]^{-1} = [A]^{-1}[A] = [I]$$

- The *inverse of a two-dimensional square matrix* can be represented simply by

$$[A]^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- The ***determiant of a matrix*** is equal to the sum of the products of all elements in any row or column by their cofactors.

$$\det(A) = |A| = \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1i} & \cdots & a_{1j} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2i} & \cdots & a_{2j} & \cdots & a_{2N} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3i} & \cdots & a_{3j} & \cdots & a_{3N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ii} & \cdots & a_{ij} & \cdots & a_{iN} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & a_{j3} & \cdots & a_{ji} & \cdots & a_{jj} & \cdots & a_{jN} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{N1} & a_{N2} & a_{N3} & \cdots & a_{Ni} & \cdots & a_{Nj} & \cdots & a_{NN} \end{vmatrix} = \sum_{i=1}^{N} a_{ik} M_{ik} (-1)^{i+k} = \sum_{j=1}^{N} a_{kj} M_{kj} (-1)^{k+j}$$

Cofactor matrix M *is* the matrix composed of multiplication of the minors of A by $(-1)^{i+j}$:

Example: Calculate the determinant and inverse of matrix A.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$$

We need the cofactor matrix C of A to find the inverse and determinant of matrix A:

$$C = \begin{bmatrix} \begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 4 \\ 1 & 3 \end{vmatrix} \\ -\begin{vmatrix} 3 & 3 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} \\ \begin{vmatrix} 3 & 3 \\ 4 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 4 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 7 & -1 & -1 \\ -3 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$A^{-1} = C^T = \begin{bmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\det(A) = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix} = 1\begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} - 1\begin{vmatrix} 3 & 3 \\ 3 & 4 \end{vmatrix} + 1\begin{vmatrix} 3 & 3 \\ 4 & 3 \end{vmatrix} = 7 - 3 - 3 = 1$$

(Using 1st column elements)

- The final matrix manipulation that will have utility in our discussion is ***augmentation***. A matrix is augmented by the addition of a column (or columns) to the original matrix.
- For example, suppose that matrix *A* augmented with the column matrix *B*:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & \vdots & b_{11} \\ a_{21} & a_{22} & \vdots & b_{21} \end{bmatrix}$$

4.3.2 Cramer's Rule

- Determinant of 2 by 2 system: $D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$

minors

- Determinant of 3 by 3 system: $D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$

Cramer's Rule: Each unknown is calculated as a fraction of two determinants. The denominator is the determinant of the system, D. The numerator is the determinant of a modified system obtained by replacing the column of coefficients of the unknown being calculated by the right-hand-side (RHS) vector.

4.3.2 Cramer's Rule

For a *3x3* system:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

$\longrightarrow$ $[A]\{x\} = \{b\}$

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}}{D} \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}}{D} \qquad x_3 = \frac{\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}}{D}$$

Example: Use the Cramer's rule to solve

$$0.3x_1 + 0.52x_2 + x_3 = -0.01$$
$$0.5x_1 + x_2 + 1.9x_3 = 0.67$$
$$0.1x_1 + 0.3x_2 + 0.5x_3 = -0.44$$

The determinant D can be written as

$$D = \begin{vmatrix} 0.3 & 0.52 & 1 \\ 0.5 & 1 & 1.9 \\ 0.1 & 0.3 & 0.5 \end{vmatrix}$$

$$D = \begin{vmatrix} 0.3 & 0.52 & 1 \\ 0.5 & 1 & 1.9 \\ 0.1 & 0.3 & 0.5 \end{vmatrix} = 0.3(-0.07) - 0.52(0.06) + 1(0.05) = -0.0022$$

The minors are

$$A_1 = \begin{vmatrix} 1 & 1.9 \\ 0.3 & 0.5 \end{vmatrix} = 1(0.5) - 1.9(0.3) = -0.07$$

$$A_2 = \begin{vmatrix} 0.5 & 1.9 \\ 0.1 & 0.5 \end{vmatrix} = 0.5(0.5) - 1.9(0.1) = 0.06$$

$$A_3 = \begin{vmatrix} 0.5 & 1 \\ 0.1 & 0.3 \end{vmatrix} = 0.5(0.3) - 1(0.1) = 0.05$$

$$x_1 = \frac{\begin{vmatrix} -0.01 & 0.52 & 1 \\ 0.67 & 1 & 1.9 \\ -0.44 & 0.3 & 0.5 \end{vmatrix}}{-0.0022} = \frac{0.03278}{-0.0022} = -14.9$$

$$x_2 = \frac{\begin{vmatrix} 0.3 & -0.01 & 1 \\ 0.5 & 0.67 & 1.9 \\ 0.1 & -0.44 & 0.5 \end{vmatrix}}{-0.0022} = \frac{0.0649}{-0.0022} = -29.5$$

$$x_3 = \frac{\begin{vmatrix} 0.3 & 0.52 & -0.01 \\ 0.5 & 1 & 0.67 \\ 0.1 & 0.3 & -0.44 \end{vmatrix}}{-0.0022} = \frac{-0.04356}{-0.0022} = 19.8$$