

IYTE EE204 Lecture 1

Linear Algebra Applications

Şevket Gümüştekin
(slides compiled in coop. with Barış Bozkurt)

Solving simultaneous equations

Systems of linear equations are generally expressed in matrix form. Define a matrix **A** of coefficients in the form

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} & \cdots & a_{0,n-1} \\ a_{10} & a_{11} & \cdots & a_{1,n-1} \\ a_{20} & a_{21} & \cdots & a_{2,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1,0} & a_{n-1,1} & \cdots & a_{n-1,n-1} \end{pmatrix}$$

and define two vectors, **x** and **b**, with elements x_0, x_1, \dots, x_{n-1} and b_0, b_1, \dots, b_{n-1} , respectively, then a system of linear equations can be written in the matrix form

$$\mathbf{Ax} = \mathbf{b}$$

Gauss-Jordan elimination

We present the method in two parts. First, we implement a simple (but slightly flawed) version of Gauss-Jordan elimination to show the basic steps of the method. In the next subsection we augment that method by a technique called *partial pivoting* to eliminate the flaw.

To see how these rules apply to a specific situation, consider a matrix \mathbf{A} and a vector \mathbf{b} as follows:

$$\mathbf{A} = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ 4 & 7 & 3 \end{pmatrix}$$

$$\mathbf{b} = \begin{pmatrix} 10 \\ 5 \\ 15 \end{pmatrix}$$

This represents the following system of equations:

$$2x_0 + 3x_1 + 5x_2 = 10$$

$$x_0 + 2x_1 + 4x_2 = 5$$

$$4x_0 + 7x_1 + 3x_2 = 15$$

Gauss-Jordan elimination proceeds by taking half the first equation and subtracting it from the second equation. The goal of these operations is to transform element a_{10} in the coefficient matrix \mathbf{A} to zero. These transformations would yield the equations

$$2x_0 + 3x_1 + 5x_2 = 10$$

$$0x_0 + .5x_1 + 1.5x_2 = 0$$

$$4x_0 + 7x_1 + 3x_2 = 15$$

The next step would be to transform a_{20} to zero by taking twice the first equation and subtracting it from the third. This would reduce the system of equations to

$$2x_0 + 3x_1 + 5x_2 = 10$$

$$0x_0 + .5x_1 + 1.5x_2 = 0$$

$$0x_0 + 1x_1 - 7x_2 = -5$$

The effect of these transformations has been to reduce all off-diagonal coefficients in the first column of \mathbf{A} to zero. We call the element on the diagonal used to transform the off-diagonal elements in the same column the *pivot element*.

The next step in Gauss-Jordan elimination is to use the diagonal element in the second row as a pivot to reduce all off-diagonal terms in the *second* column to zero. Taking 6 times the second row and subtracting it from the first row yields

$$2x_0 + 0x_1 - 4x_2 = 10$$

$$0x_0 + .5x_1 + 1.5x_2 = 0$$

$$0x_0 + 1x_1 - 7x_2 = -5$$

Taking 2 times the second row and subtracting it from the third row produces

$$2x_0 + 0x_1 - 4x_2 = 10$$

$$0x_0 + .5x_1 + 1.5x_2 = 0$$

$$0x_0 + 0x_1 - 10x_2 = -5$$

The last step is to use the third equation to eliminate off-diagonal terms in the final column. Taking .4 times the third equation and subtracting it from the first yields

$$2x_0 + 0x_1 + 0x_2 = 12$$

$$0x_0 + .5x_1 + 1.5x_2 = 0$$

$$0x_0 + 0x_1 - 10x_2 = -5$$

Taking .15 times the third equation and adding it to the second produces

$$2x_0 + 0x_1 + 0x_2 = 12$$

$$0x_0 + .5x_1 + 0x_2 = -.75$$

$$0x_0 + 0x_1 - 10x_2 = -5$$

The final solution is found by dividing the right-hand side by the diagonal coefficients, resulting in the solution $x_0 = 6$, $x_1 = -1.5$, and $x_2 = .5$. In vector notation, this solution is written as $\mathbf{x} = (6, -1.5, .5)$.

This algorithm is most easily implemented in C by creating a “working matrix” that contains all of the coefficients and the right-hand-side values. All transformations are done on the working matrix. This avoids changing values in the original arrays **A** and **b**. The working matrix has n rows and $n + 1$ columns. The first n columns are the matrix **A**, and the last column is **b**. In our simple three equation example, the working matrix would be

$$\begin{pmatrix} 2 & 3 & 5 & 10 \\ 1 & 2 & 4 & 5 \\ 4 & 7 & 3 & 15 \end{pmatrix}$$

The C implementation copies values from **A** and **b** into the working matrix and then does the appropriate Gauss-Jordan elimination. It then computes the solution vector **x**. The function requires that the user define a working array with sufficient space. This working array is one of the arguments to the function. The implementation is as follows:

```
/* Gauss-Jordan without partial pivoting */
#define A(I,J) (*(&a + (I)*n + (J)))
#define WORK(I,J) (*(&work + (I)*(n+1) + (J)))

void simple_gauss(
    double *a,          /* the coefficients in A */
    double b[],         /* the right-hand side b */
    double *work,        /* the working array */
    double x[],          /* the solution array */
    int n)               /* number of equations */

{
    double m;
    int i,j,k;

    /* set up working matrix */
    for(i=0; i<n; i++) {
        for(j=0; j<n; j++)
            WORK(i,j) = A(i,j);
        WORK(i,n) = b[i];
    }
}
```

```
/* loop through the rows of the working matrix */
for(i=0; i<n; i++)
/* perform elimination */
    for(j=0; j<n; j++)
        if(j != i) {
            m = WORK(j,i)/WORK(i,i);
            for(k=i; k<=n; k++)
                WORK(j,k) -= m*WORK(i,k);
        }

/* compute solution as rhs divided by diagonal */

for(i=0; i<n; i++)
    x[i] = WORK(i,n)/WORK(i,i);
}

#define A
#define WORK
```

The function prototype for `simple_gauss()` is

```
void simple_gauss(double *, double *, double *,
                  double *, int);
```

and a code fragment that uses the function to solve our sample three-equation linear system would be

```
/* code using simple_gauss() */
double coeff[3][3] = { {2,3,5}, {1,2,4}, {4,7,3} };
double rhs[3] = {10,5,15};
double place[3][4];
double answer[3];
simple_gauss(&coeff[0][0],rhs,&place[0][0],answer, 3);
printf("Solution is: %lf %lf %lf\n", answer[0],
      answer[1], answer[2]);
```

Note that the two matrices, `coeff` and `place`, are passed to `simple_gauss()` in the form `&coeff[0][0]` and `&place[0][0]` rather than as `coeff` and `mat`. This is because, as discussed in Section 4.6.2, `coeff` and `mat` are not pointers to `doubles`, but are pointers to *arrays of doubles*. The function `simple_gauss()` is written to accept pointers to `doubles`, which are then used as the base addresses of the matrices. Many C compilers will accept either form for the arguments even though the use of the matrix names alone is technically incorrect.

4.8.2 Partial Pivoting

The implementation of `simple_gauss()` unfortunately has a major flaw. *In particular, the method fails if any of the pivot elements is zero.* In this case the algorithm will attempt to divide an element of the coefficient matrix by zero, producing a floating-point error.

In order to avoid division by zero, we need to select a nonzero pivot element at each step. The simplest method is based on the observation that exchanging two rows in the working matrix leaves the system of equations unchanged. Thus, before pivoting on a diagonal element, we can look for a row in the working matrix that would produce a nonzero pivot element. Once we find this row, we can exchange it with the row that would have yielded a zero pivot value. This approach is called *partial pivoting*.

At each step of Gauss-Jordan elimination, all the rows below the one currently being considered are candidates for pivot elements. The rows above the current one already have been used for pivot elements. They cannot be used again because moving them to a different row of the working matrix will place their nonzero diagonal element in a nondiagonal position.

In general, there will be more than one candidate row that yields a nonzero pivot element. The selection of which one to use turns out to be significant. This is because without an appropriate selection, it is possible to show that Gauss-Jordan elimination is numerically unstable. Roundoff errors tend to accumulate in ways that seriously compromise the accuracy of the method.

The basic steps of partial pivoting are as follows:

- **Step 1.** Set up a working matrix as in the simple version of Gauss-Jordan elimination.
- **Step 2.** For each row i (starting with row 0) in the working matrix:
 1. Check rows i through $n - 1$ for the one which yields the pivot element with the largest absolute value. Call that row k .
 2. Swap row i with k .
 3. Perform Gauss-Jordan elimination using element ii of the working matrix as the pivot.
- **Step 3.** For each row i compute solution x_i as element in of the working matrix divided by element ii of the working matrix.

```
/* implementation of Gauss-Jordan elimination
   with partial pivoting */

#include <math.h>
#define TRUE 1

#define FALSE 0
#define A(I,J) (*a + (I)*n + (J))
#define WORK(I,J) (*(work + (I)*(n+1) + (J)))

int gauss_jordan(
    double *a,          /* the coefficients in A */
    double b[],         /* the right-hand side b */
    double *work,        /* the working array */
    double x[],          /* the solution array */
    int n,              /* number of equations */
    double tol)         /* minimum allowed pivot element */

{
    double m, max, temp;
    int i,j,k, swap;
```

```
/* set up working matrix */
for(i=0; i<n; i++) {
    for(j=0; j<n; j++)
        WORK(i,j) = A(i,j);
    WORK(i,n) = b[i];
}

/* loop through the rows of the working matrix */
for(i=0; i<n; i++) {
    max = -1.0;
    for(k=i; k<n; k++)
        if(fabs(WORK(k,i)) > max) {
            max = fabs(WORK(k,i));
            swap = k;
        }
    /* check if pivot element is "large enough" */
    if(max <= tol)
        return(FALSE);
```

```
/* swap rows */
    if(swap != i)
        for(k=i; k<=n; k++) {
temp = WORK(i,k);
        WORK(i,k) = WORK(swap,k);
        WORK(swap,k) = temp;
    }

/* perform elimination */
    for(j=0; j<n; j++)
        if(j != i) {
m = WORK(j,i)/WORK(i,i);
        for(k=i; k<=n; k++)
            WORK(j,k) -= m*WORK(i,k);
    }

/*
 * compute solution as rhs divided by diagonal */
    for(i=0; i<n; i++)
        x[i] = WORK(i,n)/WORK(i,i);
    return(TRUE);
}

#define A
#define WORK
```

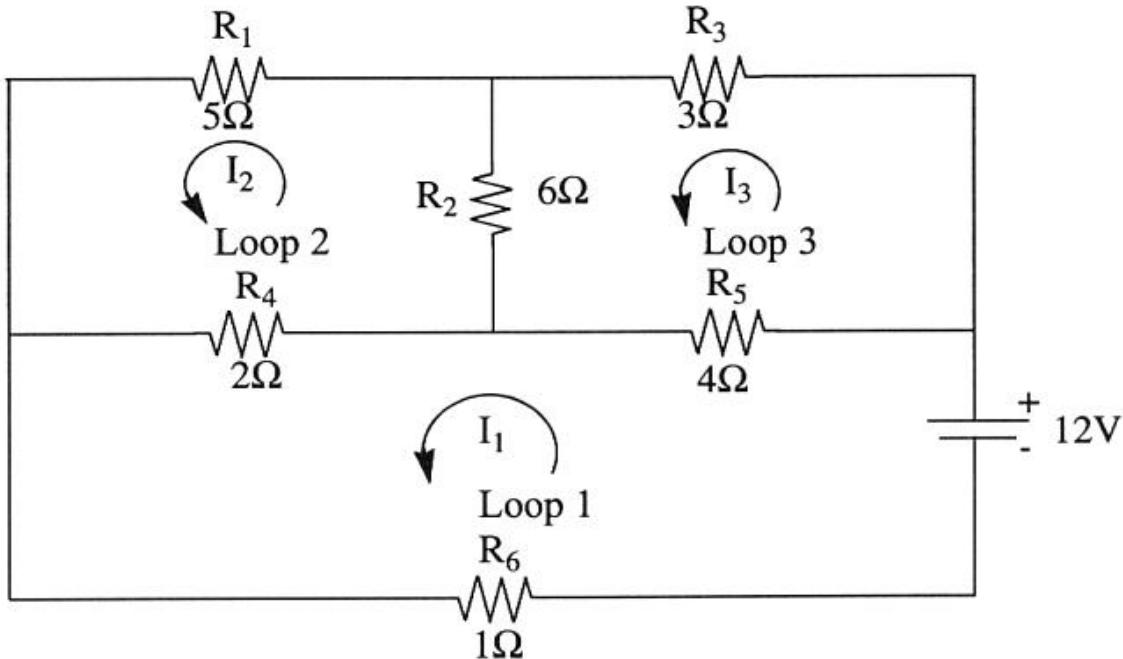
The function `gauss_jordan()` returns an `int` which is `TRUE` if it found a solution without encountering a pivot element smaller than the tolerance `tol`, and returns `FALSE` otherwise. A typical usage for the function is shown in the following `main()` program.

```
/* example program using Gauss-Jordan elimination */
#include <stdio.h>
int gauss_jordan(double *, double *, double *,
                  double *, int, double);

double coeff[3][3] = { {2,3,5}, {1,2,4}, {4,7,3} };
double rhs[3] = {10,5,15};
double place[3][4];

main()
{
    double answer[3];
    if (gauss_jordan(&coeff[0][0], rhs, &place[0][0],
                      answer, 3, .00001))
        printf("Solution is: %lf %lf %lf\n",
               answer[0], answer[1], answer[2]);
    else
        printf("No solution found\n");
}
```

Application: Electrical circuit networks



For loop 1:

$$R_4(I_1 - I_2) + R_5(I_1 - I_3) + R_6I_1 = 12$$

For loop 2:

$$R_1I_2 - R_2(I_2 - I_3) + R_4(I_2 - I_1) = 0$$

For loop 3:

$$R_2(I_3 - I_2) + R_3I_3 - R_5(I_3 - I_1) = 0$$

$$7I_1 - 2I_2 - 4I_3 = 12$$

$$-2I_1 + 13I_2 - 6I_3 = 0$$

$$-4I_1 - 6I_2 + 13I_3 = 0$$

SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS USING GAUSS ELIMINATION

This program uses Gauss Elimination to solve the system $\{A\}\{X\} = \{B\}$, where $\{A\}$ is the matrix of known coefficients, $\{B\}$ is the vector of known constants, and $\{X\}$ is the column matrix of the unknowns.

Number of equations: 3

Enter elements of matrix [A]

A(1,1) = 7
A(1,2) = -2
A(1,3) = -4
A(2,1) = -2
A(2,2) = 13
A(2,3) = -6
A(3,1) = -4
A(3,2) = -6
A(3,3) = 13

Enter elements of [B] vector

B(1) = 12
B(2) = 0
B(3) = 0

SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS

The solution is

x(1) = 2.775652
x(2) = 1.043478
x(3) = 1.335652
Determinant = 575.000000

Iterative methods for solving linear equations

Iterative methods are also available where initial guess is made for solution and then improved in steps. Unlike direct methods, iterative methods may not always yield a solution.

$$\begin{array}{lllllll} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n = b_2 \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot & & \cdot \\ a_{n1}x_1 & + & a_{n2}x_2 & & & + & a_{nn}x_n = b_n \end{array}$$

is rewritten as

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3 - \dots - a_{1n}x_n)/a_{11}$$

$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3 - \dots - a_{2n}x_n)/a_{22}$$

$$x_n = (b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n,n-1}x_{n-1})/a_{nn}$$

An iterative procedure is guaranteed to converge for any initial guess for the solution if the coefficient matrix \mathbf{A} is diagonally dominant. A diagonally dominant matrix is one in which the absolute value of the diagonal element a_{ii} in each row i is larger than the sum of the absolute values of all other elements in that row. That is, for each row $i = 1, 2, \dots, n$

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

An iterative method may still converge even if the coefficient matrix is not diagonally dominant, particularly if the largest elements of the matrix are located on the diagonal. Thus, the requirement of diagonal dominance is a sufficient condition but not a necessary condition and it is possible to find a solution to systems that possess a strong diagonal.

THE JACOBI METHOD

One of the simplest iterative methods is the Jacobi method. As with the other iterative methods, the Jacobi method begins by making an initial guess of the values of $x_1^{(1)}$, $x_2^{(1)}$, ..., $x_n^{(1)}$. A common approach is to use the following first approximation:

$$x_1^{(1)} = b_1/a_{11}$$

$$x_2^{(1)} = b_2/a_{22}$$

.

.

.

$$x_n^{(1)} = b_n/a_{nn}$$

The superscript in the above equations represents the number of iterations performed in approximating the solution.

The second approximation is obtained by substituting these first approximations in Equation 19.14. Thus,

$$\begin{aligned}
x_1^{(2)} &= (b_1 - a_{12}x_2^{(1)} - a_{13}x_3^{(1)} - \dots - a_{1n}x_n^{(1)})/a_{11} \\
x_2^{(2)} &= (b_2 - a_{21}x_1^{(1)} - a_{23}x_3^{(1)} - \dots - a_{2n}x_n^{(1)})/a_{22} \\
&\cdot \\
&\cdot \\
&\cdot \\
x_n^{(2)} &= (b_n - a_{n1}x_1^{(1)} - a_{n2}x_2^{(1)} - \dots - a_{nn-1}x_{n-1}^{(1)})/a_{nn}
\end{aligned} \tag{19.17}$$

The equations used to solve for the x_i 's for the k th iteration can be written as:

$$\begin{aligned}
x_1^{(k+1)} &= (b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \dots - a_{1n}x_n^{(k)})/a_{11} \\
x_2^{(k+1)} &= (b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)})/a_{22} \\
&\cdot \\
&\cdot \\
&\cdot \\
x_n^{(k+1)} &= (b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \dots - a_{nn-1}x_{n-1}^{(k)})/a_{nn}
\end{aligned} \tag{19.18}$$

In general, for a n by n system of equations, these equations can be written as

$$x_i^{(k+1)} = b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j^{(k)} \tag{19.19}$$

The process is continued until the new values $\{x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_n^{(k+1)}\}$ are sufficiently close to the previous values $\{x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}\}$. The test for convergence can be written as

$$\left| x_i^{(k+1)} - x_i^{(k)} \right| > \left| x_i^{(k+1)} \cdot \varepsilon \right| , \quad i = 1, 2, \dots, n \quad (19.20)$$

where ε is the desired tolerance.

Example 19.5 Jacobi Method

PROBLEM STATEMENT: Solve the following set of linear simultaneous equations using the Jacobi method

$$10x_1 + 2x_2 + 3x_3 = 11$$

$$x_1 + 5x_2 + 2x_3 = 20$$

$$3x_1 + 2x_2 + 6x_3 = -12$$

SOLUTION: We begin by first checking the diagonal coefficients to determine if the coefficient matrix is diagonally dominant.

$$|10| > |2| + |3| \rightarrow 10 > 5 \quad \text{O.K.}$$

$$|5| > |1| + |2| \rightarrow 5 > 3 \quad \text{O.K.}$$

$$|6| > |3| + |2| \rightarrow 6 > 5 \quad \text{O.K.}$$

The coefficient matrix is diagonally dominant so the Jacobi iterative procedure will converge.

We compute initial values for $x_1^{(1)}$, $x_2^{(1)}$, and $x_3^{(1)}$ from

$$x_1^{(1)} = \frac{b_2}{a_{11}} = \frac{10}{10} = 1.0$$

$$x_2^{(1)} = \frac{b_2}{a_{22}} = \frac{20}{5} = 4.0$$

$$x_3^{(1)} = \frac{b_3}{a_{33}} = -\frac{12}{6} = -2.0$$

The second approximation to the solution is

$$x_1^{(2)} = \frac{(b_1 - a_{12}x_2^{(1)} - a_{13}x_3^{(1)})}{a_{11}} = \frac{[10 - 2(4) - 3(-2.0)]}{10} = 0.8$$

$$x_2^{(2)} = \frac{(b_2 - a_{21}x_1^{(1)} - a_{23}x_3^{(1)})}{a_{22}} = \frac{[20 - 1(1.0) - 2(-2.0)]}{5} = 4.6$$

$$x_3^{(2)} = \frac{b_3 - a_{31}x_1^{(1)} - a_{32}x_2^{(1)}}{a_{33}} = \frac{-12 - 3(1.0) - 2(4.0)}{6} = -3.833$$

Substituting the above values, we obtain the third approximation from

$$x_1^{(3)} = \frac{b_1 - a_{12}x_2^{(2)} - a_{13}x_3^{(2)}}{a_{11}} = \frac{10 - 2(4.6) - 3(-3.833)}{10} = 1.230$$

$$x_2^{(3)} = \frac{b_2 - a_{21}x_1^{(2)} - a_{23}x_3^{(2)}}{a_{22}} = \frac{20 - 1(0.8) - 2(-3.833)}{5} = 5.373$$

$$x_3^{(3)} = \frac{b_3 - a_{31}x_1^{(2)} - a_{32}x_2^{(2)}}{a_{33}} = \frac{-12 - 3(0.8) - 2(4.6)}{6} = -3.933$$

Table 19.4 shows the results for the first seven iterations. It is left for the reader to verify that the solution is $x_1 = 1.231$, $x_2 = 5.538$, and $x_3 = -4.462$.

Table 19.4: Results of first seven iteration for the Jacobi method

Variable	Iteration Number						
	1	2	3	4	5	6	7
x_1	1.000	0.800	1.230	1.105	1.256	1.190	1.247
x_2	4.000	4.600	5.373	5.373	5.480	5.480	5.552
x_3	-2.000	-3.833	-3.933	-4.406	-4.328	-4.475	4.422

C program for the Jacobi method

```
#include <stdio.h>
#include <math.h>

#define MAXSIZE 20                      /* maximum size of matrix */
#define FALSE 0
#define TRUE 1
/* function prototype */
int jacobi(double a[][MAXSIZE],double b[],int n,int max_iter,
           double tol,double x[]);
```

```
void main(void)
{
    double a[MAXSIZE][MAXSIZE], b[MAXSIZE], x[MAXSIZE];
    double tol;
    int i, j, n, return_val;
    int max_iter;

    printf("\n \t SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS");
    printf("\n \t USING THE JACOBI METHOD\n\n ");
    printf("\n      This program uses the JAcobi Method to solve the");
    printf("\n      system [A]{X} = {B}, where [A] is the matrix of");
    printf("\n      known coefficients, {B} is the vector of known
          constants,");
    printf("\n      and {X} is the column matrix of the unknowns.");

    /* get number of equations */
    n = 0;
    while (n <= 0 || n > MAXSIZE)
    {
        printf("\n Number of equations: ");
        scanf("%d", &n);
    }
    /* Read matrix [A] */
    printf("\n Enter elements of matrix [A] \n");
    for (i=0; i < n; ++i)
        for (j=0; j < n; ++j)
        {
            printf(" A(%d,%d) = ", i+1, j+1);
            scanf("%lf", &a[i][j]);
        }
    /* Read {B} vector */
    printf("\n Enter elements of [B] vector \n");
    for (i = 0; i < n; ++i)
    {
        printf(" B(%d) = ", i+1);
        scanf("%lf", &b[i]);
    }
}
```

```
/* read maximum number of iterations */
printf("\n Maximum number of iterations: ");
scanf("%d", &max_iter);

/* read tolerance */
printf("\n Enter desired tolerance ");
scanf("%d", &tol);

/* call Jacobi function */
return_val= jacobi(a,b,n,max_iter,tol,x);

/* print results      */
if (return_val == 0)
{
    printf("\n \t SOLUTION OF SIMULTANEOUS LINEAR EQUATIONS\n");
    printf("\n \t The solution is");
}
else
{
    printf("\n\n\t Did not converge after %d iterations",max_iter);
    printf("\n Current values of x{} are ");
}
for (i=0; i < n; ++i)
    printf("\n \t x(%d) = %lf",i+1,x[i]);
}
```

```
/*
 *-----*
 * Function: jacobi() *
 */
/*
 * This function solves the system of equations *
 * [A]{x} = {B} using the Jacobi iterative method. *
 */
/*
 * Input Parameters: *
 */
/*      n          - number of equations *
/*      a[n][n]     - coefficient matrix *
/*      b[n]        - right-hand side vector *
/*      max_iter    - maximum number of iterations *
*/
/*
 * Output Parameters: *
*/
/*      x[n]        - solution vector *
*/
/*
 * Returns *
*/
/*      0 if successful *
/*      1 if did not converge after max_iter iterations *
*/
/*-----*/
int jacobi(double a[][MAXSIZE],double b[],int n,int max_iter,
           double tol,double x[])
{
    int num_iter=0;
    int tol_exceeded = TRUE;
    int i,j;
    double x_old[MAXSIZE];           /* previous solution */
    double sum;
```

```
for (i=0; i < n; ++i)
    x[i] = b[i]/a[i][i];

while (tol_exceeded && num_iter < max_iter)
{
    /* save old values of x{} */
    for (i=0; i < n; ++i)
        x_old[i] = x[i];

    /* compute new values of x{} */

    for (i = 0; i < n; ++i)
    {
        sum = b[i];
        for (j = 0; j <n; ++j)
            if (i != j )
                sum -= a[i][j] * x_old[j];

        x[i] = sum/a[i][i];
    }

    /* check if solution is within desired tolerance */
    tol_exceeded = FALSE;
    for (i = 0; i < n ; ++i)
        if ( fabs(x[i] - x_old[i]) > fabs(x_old[i] * tol) )
            tol_exceeded = TRUE;

    ++num_iter;
}

return(tol_exceeded);
}
```

Matrix Inversion

Note that, using the inverse of a square matrix A, one can get:

$$AA^{-1} = I$$

Gauss-Jordan elimination can be used for finding the inverse of a matrix by restructuring the above equation:

$$\begin{aligned} Ax^0 &= I^0 \\ Ax^1 &= I^1 \\ Ax^2 &= I^2 \\ &\vdots \\ Ax^{n-1} &= I^{n-1} \end{aligned}$$

Where x^i and I^i are the respective columns of inverse matrix A^{-1} and identity matrix I .

Matrix Decomposition: Case 1, LU Decomposition (adopted from slides by Nishith Kumar)

Let A be a $m \times m$ with nonsingular square matrix. Then there exists two matrices L and U such that, where L is a lower triangular matrix and U is an upper triangular matrix.

$$A = LU$$
$$L = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mm} \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ 0 & u_{22} & \cdots & u_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{mm} \end{bmatrix}$$

Uses: The LU decomposition is most commonly used in the solution of systems of simultaneous linear equations. We can also find determinant easily by using LU decomposition (Product of the diagonal element of upper and lower triangular matrix).

Case 1, LU Decomposition: Calculation of L & U

$$A = \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix}$$

Now reducing the first column we have

$$\begin{bmatrix} 6 & -2 & 2 \\ 0 & -4 & 2 \\ 0 & -12 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 6 & -2 & 2 \\ 0 & -4 & 2 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix}$$

Case 1, LU Decomposition: Calculation of L & U

Now

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

Therefore,

$$A = \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1/2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 6 & -2 & 2 \\ 0 & -4 & 2 \\ 0 & 0 & -5 \end{bmatrix} = LU$$

If A is a Non singular matrix then for each L (lower triangular matrix) the upper triangular matrix is unique but an LU decomposition is not unique. There can be more than one such LU decomposition for a matrix. Such as

$$A = \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 12 & 1 & 0 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2/6 & 2/6 \\ 0 & -4 & 2 \\ 0 & 0 & -5 \end{bmatrix} = LU$$

Case 1, LU Decomposition: Solving System of Linear Equations using LU Decomposition

Suppose we would like to solve a $m \times m$ system $AX = b$. Then we can find a LU-decomposition for A, then to solve $AX = b$, it is enough to solve the systems

$$\left. \begin{array}{l} LY = b \\ UX = Y \end{array} \right\}$$

Thus the system $LY = b$ can be solved by the method of forward substitution and the system $UX = Y$ can be solved by the method of backward substitution. To illustrate, we give some examples

Consider the given system $AX = b$, where

$$A = \begin{bmatrix} 6 & -2 & 2 \\ 12 & -8 & 6 \\ 3 & -13 & 2 \end{bmatrix} \quad \text{and} \quad b = \begin{pmatrix} 8 \\ 14 \\ -17 \end{pmatrix}$$

Case 1, LU Decomposition: Solving System of Linear Equations using LU Decomposition

We have seen $A = LU$, where

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1/2 & 3 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 6 & -2 & 2 \\ 0 & -4 & 2 \\ 0 & 0 & -5 \end{bmatrix}$$

Thus, to solve $AX = \mathbf{b}$, we first solve $LY = \mathbf{b}$ by forward substitution

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1/2 & 3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 14 \\ -17 \end{bmatrix} \quad \text{then} \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -2 \\ -15 \end{bmatrix}$$

Now, we solve $UX = Y$ by backward substitution

$$\begin{bmatrix} 6 & -2 & 2 \\ 0 & -4 & 2 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -2 \\ -15 \end{bmatrix} \quad \text{then} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Case 1, LU Decomposition code by Pramesh Pudasaini (from <http://www.codewithc.com>)

```
C Program for LU Factorization (LU Decomposition)
1 #include<stdio.h>
2 #include<conio.h>
3 void main()
4 {
5     float A[20][20] = {0}, L[20][20] = {0}, U[20][20];
6     float B[20] = {0}, X[20] = {0}, Y[20] = {0};
7     int i,j,k,n;
8     printf("Enter the order of square matrix: ");
9     scanf("%d",&n);
10    printf("\nEnter matrix element:\n");
11    for(i=0; i<n; i++)
12    {
13        for(j=0; j<n; j++)
14        {
15            printf("Enter A[%d][%d] element: ", i, j);
16            scanf("%f", &A[i][j]);
17        }
18    }
19    printf("\nEnter the constant terms: \n");
20    for(i=0; i<n; i++)
21    {
22        printf("B[%d]", i);
23        scanf("%f", &B[i]);
24    }
```

```
25     for(j=0; j<n; j++)
26     {
27         for(i=0; i<n; i++)
28         {
29             if(i<=j)
30             {
31                 U[i][j]=A[i][j];
32                 for(k=0; k<i-1; k++)
33                     U[i][j]-=L[i][k]*U[k][j];
34                 if(i==j)
35                     L[i][j]=1;
36                 else
37                     L[i][j]=0;
38             }
39             else
40             {
41                 L[i][j]=A[i][j];
42                 for(k=0; k<=j-1; k++)
43                     L[i][j]-=L[i][k]*U[k][j];
44                 L[i][j]/=U[j][j];
45                 U[i][j]=0;
46             }
47         }
48     }
```

```
49     printf("[L]: \n");
50     for(i=0; i<n; i++)
51     {
52         for(j=0; j<n; j++)
53             printf("%9.3f", L[i][j]);
54         printf("\n");
55     }
56     printf("\n\n[U]: \n");
57     for(i=0; i<n; i++)
58     {
59         for(j=0; j<n; j++)
60             printf("%9.3f", U[i][j]);
61         printf("\n");
62     }
63     for(i=0; i<n; i++)
64     {
65         Y[i]=B[i];
66         for(j=0; j<i; j++)
67         {
68             Y[i]-=L[i][j]*Y[j];
69         }
70     }
```

```
71     printf("\n\n[Y]: \n");
72     for(i=0; i<n; i++)
73     {
74         printf("%9.3f", Y[i]);
75     }
76     for(i=n-1; i>=0; i--)
77     {
78         X[i]= Y[i];
79         for(j=i+1; j<n; j++)
80         {
81             X[i]-=U[i][j]*X[j];
82         }
83         X[i]/=U[i][i];
84     }
85     printf("\n\n[X]: \n");
86     for(i=0; i<n; i++)
87     {
88         printf("%9.3f", X[i]);
89     }
90     getch();
91 }
```

Case 2, QR Decomposition

If A is a $m \times n$ matrix with linearly independent columns, then A can be decomposed as , $A=QR$ where Q is a $m \times n$ matrix whose columns form an orthonormal basis for the column space of A and R is an nonsingular upper triangular matrix.

Uses: QR decomposition is widely used in computer codes to find the eigenvalues of a matrix, to solve linear systems, and to find least squares approximations.

Case 2, QR Decomposition

Suppose $A = [u_1 \mid u_2 \mid \dots \mid u_n]$ and $\text{rank}(A) = n$.

Apply the Gram-Schmidt process to $\{u_1, u_2, \dots, u_n\}$ and the orthogonal vectors v_1, v_2, \dots, v_n are

$$v_i = u_i - \frac{\langle u_i, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle u_i, v_2 \rangle}{\|v_2\|^2} v_2 - \dots - \frac{\langle u_i, v_{i-1} \rangle}{\|v_{i-1}\|^2} v_{i-1}$$

Let $q_i = \frac{v_i}{\|v_i\|}$ for $i=1,2,\dots,n$. Thus q_1, q_2, \dots, q_n form a orthonormal basis for the column space of A .

Case 2, QR Decomposition

Let $Q = [q_1 \ q_2 \ \dots \ q_n]$, so Q is a $m \times n$ matrix whose columns form an orthonormal basis for the column space of A .

Now,

$$A = [u_1 \ u_2 \ \dots \ u_n] = [q_1 \ q_2 \ \dots \ q_n] \begin{bmatrix} \|v_1\| & \langle u_2, q_1 \rangle & \langle u_3, q_1 \rangle & \dots & \langle u_n, q_1 \rangle \\ 0 & \|v_2\| & \langle u_3, q_2 \rangle & \dots & \langle u_n, q_2 \rangle \\ 0 & 0 & \|v_3\| & \dots & \langle u_n, q_3 \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \|v_n\| \end{bmatrix}$$

i.e., $A = QR$.

where,

$$R = \begin{bmatrix} \|v_1\| & \langle u_2, q_1 \rangle & \langle u_3, q_1 \rangle & \dots & \langle u_n, q_1 \rangle \\ 0 & \|v_2\| & \langle u_3, q_2 \rangle & \dots & \langle u_n, q_2 \rangle \\ 0 & 0 & \|v_3\| & \dots & \langle u_n, q_3 \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \|v_n\| \end{bmatrix}$$

Thus A can be decomposed as $A = QR$, where R is an upper triangular and nonsingular matrix. 42

Case 2, QR Decomposition

Example: Find the QR decomposition of

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

Case 2, QR Decomposition : Calculation of Q & R

Applying Gram-Schmidt process of computing QR decomposition

1st Step: $r_{11} = \| \alpha_1 \| = \sqrt{3}$

$$q_1 = \frac{1}{\| \alpha_1 \|} \alpha_1 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{pmatrix}$$

2nd Step: $r_{12} = q_1^T \alpha_2 = -2/\sqrt{3}$

3rd Step:

$$\hat{q}_2 = \alpha_2 - q_1 q_1^T \alpha_2 = \alpha_2 - q_1 r_{12} = \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \end{pmatrix} - (-2/\sqrt{3}) \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 2/3 \\ -1/3 \\ 0 \end{pmatrix}$$

$$r_{22} = \| \hat{q}_2 \| = \sqrt{2/3}$$

$$q_2 = \frac{1}{\| \hat{q}_2 \|} \hat{q}_2 = \begin{pmatrix} -1/\sqrt{6} \\ \sqrt{2/3} \\ -1/\sqrt{6} \\ 0 \end{pmatrix}$$

Case 2, QR Decomposition : Calculation of Q & R

4th Step:

$$r_{13} = q_1^T a_3 = -1/\sqrt{3}$$

5th Step:

$$r_{23} = q_2^T a_3 = 1/\sqrt{6}$$

6th Step:

$$\hat{q}_3 = a_3 - q_1 q_1^T a_3 - q_2 q_2^T a_3 = a_3 - r_{13} q_1 - r_{23} q_2 = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \\ -1 \end{pmatrix}$$

$$r_{33} = \| \hat{q}_3 \| = \sqrt{6}/2$$

$$q_3 = \frac{1}{\| \hat{q}_3 \|} \hat{q}_3 = \begin{pmatrix} -1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{pmatrix}$$

Case 2, QR Decomposition : Calculation of Q & R

Therefore, $A=QR$

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

**Other Methods exist for Matrix Decompositions such as
Cholesky (similar to LU but more efficient),
Singular Value Decomposition SVD (widely used in many
signal processing applications),
etc.**