Introduction to Numerical Analysis for Engineers

- Systems of Linear Equations — Mathews
  - Cramer's Rule
  - Gaussian Elimination — 3.3-3.5
    - Numerical implementation
      3.3-3.4
    - Numerical stability
      - Partial Pivoting
      - Equilibration
      - Full Pivoting
    - Multiple right hand sides
    - Computation count
    - LU factorization — 3.5
    - Error Analysis for Linear Systems — 3.4
      - Condition Number
    - Special Matrices
  - Iterative Methods — 3.6
    - Jacobi's method
    - Gauss-Seidel iteration
    - Convergence

Systems of Linear Equations
Gaussian Elimination

Multiple Right-hand Sides

Reduction Step k

$$
\begin{array}{c}
\\ \\ k \rightarrow \\ \\ \\ \\ \\ \\
\end{array}
\underbrace{\left[\begin{array}{ccc|cccc}
\times & & & & & & \\
0 & \times & & & & & \\
\cdot & \cdot & \times & & & & \\
\hline
\cdot & & 0 & \times & \times & \times & \times \\
\cdot & & \cdot & \times & & & \\
\cdot & & \cdot & \times & & & \\
\cdot & & \cdot & \times & & & \\
0 & \cdot & 0 & \times & & &
\end{array}\right]}_{n \;=\; k \,+\, (n-k)}
\cdot \bar{\mathbf{x}} =
\underbrace{\left\{\begin{array}{ccc}
\times & \times & \times \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\hline
\times & \times & \times \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot \\
\times & \times & \times
\end{array}\right\}}_{p}
$$

Computation Count Reduction Step k

$(n-k)(n-k+p)$ Operations

Total Computation Count

$$n \gg 1$$

Reduction

$$N_r = \sum_{k=1}^{n-1} (n-k)(n-k+p) \simeq \frac{1}{3}n^3 + \frac{1}{2}n^2(p-1)$$

Back Substitution

$$N_b = \sum_{k=1}^{n-1} (n-k)p \simeq \frac{1}{2}n^2 p$$

$$\boxed{n \gg 1 \Rightarrow N_r \gg N_b}$$

Reduction for each right-hand side inefficient.
However, RHS may be result of iteration and unknown a priori
(e.g. Euler's method) -> LU Factorization

Systems of Linear Equations
LU Factorization

The coefficient Matrix $\overline{\overline{\mathbf{A}}}$ is decomposed as

$$\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{L}}} \cdot \overline{\overline{\mathbf{U}}}$$

where $\overline{\overline{\mathbf{L}}}$ is a lower triangular matrix
and $\overline{\overline{\mathbf{U}}}$ is an upper triangular matrix

$$\overline{\overline{\mathbf{L}}} = [l_{ij}] = \begin{bmatrix} l_{11} & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ l_{21} & l_{22} & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot \\ \cdot & & \cdot & l_{kk} & \cdot & & \cdot \\ & & & & \cdot & \cdot & 0 \\ l_{n1} & \cdot & \cdot & \cdot & \cdot & l_{n,n-1} & l_{nn} \end{bmatrix}$$

Then the solution is performed in two simple steps

1. $\overline{\overline{\mathbf{L}}}\vec{y} = \vec{b}$ Forward substitution
2. $\overline{\overline{\mathbf{U}}}\vec{x} = \vec{y}$ Back substitution

$$\overline{\overline{\mathbf{U}}} = [u_{ij}] = \begin{bmatrix} u_{11} & u_{12} & \cdot & \cdot & \cdot & \cdot & u_{1n} \\ 0 & u_{22} & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & & & & \\ \cdot & & \cdot & u_{kk} & \cdot & & \\ \cdot & & & & \cdot & \cdot & u_{n-1,n} \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & u_{nn} \end{bmatrix}$$

How to determine $\overline{\overline{\mathbf{L}}}$ and $\overline{\overline{\mathbf{U}}}$?

Systems of Linear Equations
LU Factorization

$$
\overline{\mathbf{A}}^{(i)} = \left[a_{ij}^{(i)}\right] =
\begin{array}{c|c}
 & j \rightarrow \\
\hline
i \downarrow &
\begin{bmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & \cdot & \cdot & \cdot & \cdot & a_{1n}^{(1)} \\
0 & a_{22}^{(2)} & \cdot & \cdot & \cdot & \cdot & a_{2n}^{(2)} \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & 0 & a_{ii}^{(i)} & \cdot & \cdot & \cdot \\
\cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \cdot & 0 & a_{ni}^{(i)} & \cdot & \cdot & a_{nn}^{(i)}
\end{bmatrix}
\end{array}
$$

After reduction step *i-1*:

Above and on diagonal

$$i \leq j$$

Unchanged after step *i-1*

$$a_{ij}^{(n)} = \cdots a_{ij}^{(i)}$$

Below diagonal

$$j < i$$

Become and remain *0* in step *j*

$$a_{ij}^{(n)} = \cdots a_{ij}^{(j+1)} = 0$$

Change in reduction steps *1 - i-1*:

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - m_{ik} a_{kj}^{(k)}, \quad m_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}$$

Total change above diagonal

$$i \leq j \; : \; a_{ij}^{(i)} = a_{ij} - \sum_{k=1}^{i-1} m_{ik} a_{kj}^{(k)}$$

Total change below diagonal

$$i > j \; : \; a_{ij}^{(i)} = 0 = a_{ij} - \sum_{k=1}^{j} m_{ik} a_{kj}^{(k)}$$

Define

$$m_{ii} = 1, \quad i = 1, \ldots n$$

=>

$$i \leq j \; : \; a_{ij} = \sum_{k=1}^{i} m_{ik} a_{kj}^{(k)}$$

$$i > j \; : \; a_{ij} = \sum_{k=1}^{j} m_{ik} a_{kj}^{(k)}$$

$$\boxed{\Rightarrow a_{ij} = \sum_{k=1}^{\min(i,j)} m_{ik} a_{kj}^{(k)}}$$

Systems of Linear Equations
LU Factorization

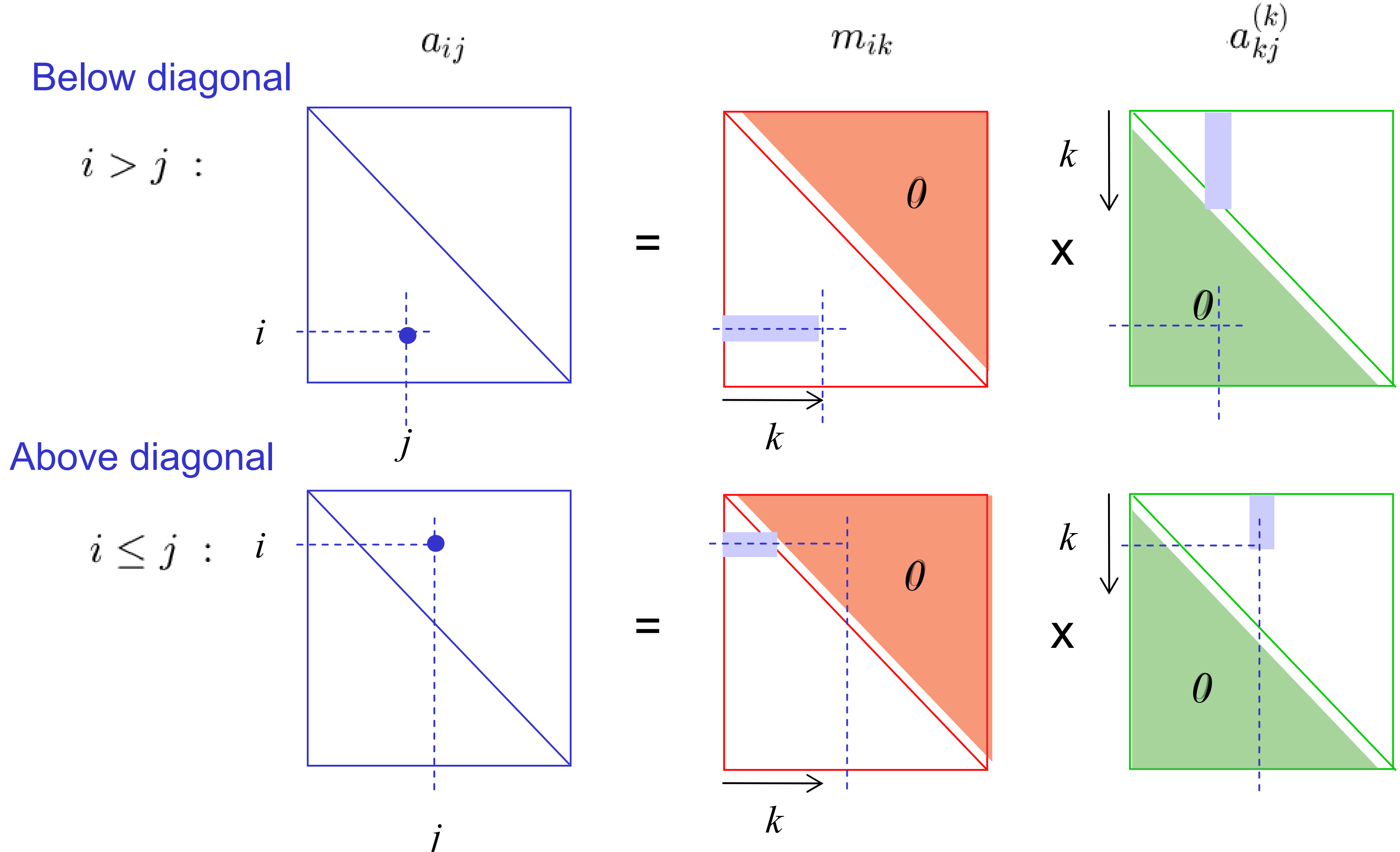
'Matrix product'

Sum stops at diagonal

$$a_{ij} = \sum_{k=1}^{\min(i,j)} m_{ik} a_{kj}^{(k)}$$

Lower triangular

Upper triangular



Systems of Linear Equations
LU Factorization

GE Reduction directly yields LU factorization

$$\overline{\overline{\mathbf{A}}} = \overline{\overline{\mathbf{L}}} \cdot \overline{\overline{\mathbf{U}}}$$

Lower triangular

$$\overline{\overline{\mathbf{L}}} = l_{ij} = \begin{cases} 0 & i < j \\ 1 & i = j \\ m_{ij} & i > j \end{cases}$$

Upper triangular

$$\overline{\overline{\mathbf{U}}} = u_{ij} = \begin{cases} a_{ij}^{(i)} & i \le j \\ 0 & i > j \end{cases}$$

Compact storage

$$\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdot & \cdot & \cdot & \cdot & a_{1n}^{(1)} \\ m_{21} & a_{22}^{(2)} & \cdot & \cdot & \cdot & \cdot & a_{2n}^{(2)} \\ \cdot & \cdot & \cdot & & & & \cdot \\ \cdot & & \cdot & \cdot & & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ m_{n1} & \cdot & \cdot & \cdot & \cdot & m_{n,n-1} & a_{nn}^{(n)} \end{bmatrix}$$

Lower diagonal implied

$$m_{ii} = 1, \quad i = 1, \ldots n$$

Systems of Linear Equations
Pivoting in LU Factorization

Before reduction, step k

$$\left[\begin{array}{ccccccc} a_{11}^{(1)} & a_{12}^{(1)} & \cdot & \cdot & \cdot & \cdot & a_{1n}^{(1)} \\ m_{21} & a_{22}^{(2)} & \cdot & \cdot & \cdot & \cdot & a_{2n}^{(2)} \\ \cdot & \cdot & \cdot & & & & \cdot \\ \hline \cdot & & m_{k,k-1} & a_{kk}^{(k)} & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot \\ m_{n1} & \cdot & m_{n,k-1} & a_{nk}^{(k)} & \cdot & \cdots & a_{nn}^{(n)} \end{array}\right]$$

Pivoting if

$$\left|a_{ik}^{(k)}\right| \gg \left|a_{kk}^{(k)}\right| \;, \quad i > k$$

Interchange rows i and k

$$p_k = i$$

else

$$p_k = k$$

Pivot element vector

$$p_i, \quad i = 1, \ldots n$$

Forward substitution, step k

$$\overline{\overline{\mathbf{L}}}\vec{y} \;=\; \vec{b}$$

Interchange rows i and k

$$\left\{\begin{array}{c} b_1 \\ \cdot \\ \cdot \\ b_k \\ \cdot \\ b_i \\ \cdot \\ b_n \end{array}\right\}$$

$$p_k = i \Rightarrow \left\{\begin{array}{lcl} b_i^{(k)} & = & b_k \\ b_k & = & b_i \\ b_i & = & b_i^{(k)} \end{array}\right.$$

Linear Systems of Equations
Error Analysis

Function of one variable

$$y = f(x)$$

Condition number

$$\left| \frac{f(\overline{x}) - f(x)}{f(x)} \right| = K \left| \frac{\overline{x} - x}{x} \right| \;, \quad \overline{x} = x + \delta x$$

$$\left| \frac{\delta y}{y} \right| = K \left| \frac{\delta x}{x} \right|$$

The condition number K is a measure of the amplification of the relative error by the function f(x)

Linear systems

How is the relative error of $\overline{\mathbf{x}}$ dependent on errors in $\overline{\mathbf{b}}$?

$$\overline{\overline{\mathbf{A}}}\,\overline{\mathbf{x}} = \overline{\mathbf{b}}$$

Example

$$\overline{\overline{\mathbf{A}}} = \begin{bmatrix} 1.0 & 1.0 \\ 1.0 & 1.0001 \end{bmatrix} \;, \quad \det(\overline{\overline{\mathbf{A}}}) = 0.0001$$

$$\overline{\mathbf{b}} = \begin{Bmatrix} 2 \\ 2 \end{Bmatrix} \Rightarrow \overline{\mathbf{x}} = \begin{Bmatrix} 2 \\ 0 \end{Bmatrix}$$

$$\overline{\mathbf{b}} = \begin{Bmatrix} 2 \\ 2.0001 \end{Bmatrix} \Rightarrow \overline{\mathbf{x}} = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

Small changes in $\overline{\mathbf{b}}$ give large changes in $\overline{\mathbf{x}}$

The system is ill-Conditioned

Linear Systems of Equations
Error Analysis

Vector and Matrix Norm

$$||\overline{\mathbf{x}}||_\infty = \max_i |x_i|$$

$$\left\|\overline{\overline{\mathbf{A}}}\right\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

Properties

$$\overline{\overline{\mathbf{A}}} \neq \overline{\overline{\mathbf{0}}} \Rightarrow \left\|\overline{\overline{\mathbf{A}}}\right\| > 0$$

$$\left\|\alpha\overline{\overline{\mathbf{A}}}\right\| = |\alpha| \left\|\overline{\overline{\mathbf{A}}}\right\|$$

$$\left\|\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{B}}}\right\| \leq \left\|\overline{\overline{\mathbf{A}}}\right\| + \left\|\overline{\overline{\mathbf{B}}}\right\|$$

$$\left\|\overline{\overline{\mathbf{A}\mathbf{B}}}\right\| \leq \left\|\overline{\overline{\mathbf{A}}}\right\| \left\|\overline{\overline{\mathbf{B}}}\right\|$$

$$\boxed{\left\|\overline{\overline{\mathbf{A}}}\overline{\mathbf{x}}\right\| \leq \left\|\overline{\overline{\mathbf{A}}}\right\| ||\overline{\mathbf{x}}||}$$

Perturbed Right-hand Side

$$\overline{\overline{\mathbf{A}}}\overline{\mathbf{x}} = \overline{\mathbf{b}}$$

$$\Downarrow$$

$$\overline{\overline{\mathbf{A}}}(\overline{\mathbf{x}} + \delta\overline{\mathbf{x}}) = \overline{\mathbf{b}} + \delta\overline{\mathbf{b}}$$

Subtract original equation

$$\begin{aligned} \overline{\overline{\mathbf{A}}}\delta\overline{\mathbf{x}} &= \delta\overline{\mathbf{b}} \\ \delta\overline{\mathbf{x}} &= \overline{\overline{\mathbf{A}}}^{-1}\delta\overline{\mathbf{b}} \end{aligned}$$

$$\Downarrow$$

$$\left.\begin{aligned} ||\delta\overline{\mathbf{x}}|| &\leq \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| ||\delta\overline{\mathbf{b}}|| \\ ||\overline{\mathbf{b}}|| = \left\|\overline{\overline{\mathbf{A}}}\overline{\mathbf{x}}\right\| &\leq \left\|\overline{\overline{\mathbf{A}}}\right\| ||\overline{\mathbf{x}}|| \end{aligned}\right\} \Rightarrow$$

Relative Error Magnification

$$\frac{||\delta\overline{\mathbf{x}}||}{||\overline{\mathbf{x}}||} \leq \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| \left\|\overline{\overline{\mathbf{A}}}\right\| \frac{||\delta\overline{\mathbf{b}}||}{||\overline{\mathbf{b}}||}$$

Condition Number

$$K(\overline{\overline{\mathbf{A}}}) = \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| \left\|\overline{\overline{\mathbf{A}}}\right\|$$

Linear Systems of Equations
Error Analysis

Vector and Matrix Norm

$$||\overline{\mathbf{x}}||_\infty = \max_i |x_i|$$

$$\left\|\overline{\overline{\mathbf{A}}}\right\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$$

Properties

$$\overline{\overline{\mathbf{A}}} \neq \overline{\overline{\mathbf{0}}} \Rightarrow \left\|\overline{\overline{\mathbf{A}}}\right\| > 0$$

$$\left\|\alpha\overline{\overline{\mathbf{A}}}\right\| = |\alpha| \left\|\overline{\overline{\mathbf{A}}}\right\|$$

$$\left\|\overline{\overline{\mathbf{A}}} + \overline{\overline{\mathbf{B}}}\right\| \leq \left\|\overline{\overline{\mathbf{A}}}\right\| + \left\|\overline{\overline{\mathbf{B}}}\right\|$$

$$\boxed{\begin{aligned} \left\|\overline{\overline{\mathbf{A}}\,\overline{\overline{\mathbf{B}}}}\right\| &\leq \left\|\overline{\overline{\mathbf{A}}}\right\| \left\|\overline{\overline{\mathbf{B}}}\right\| \\ \left\|\overline{\overline{\mathbf{A}}}\overline{\mathbf{x}}\right\| &\leq \left\|\overline{\overline{\mathbf{A}}}\right\| ||\overline{\mathbf{x}}|| \end{aligned}}$$

Perturbed Coefficient Matrix

$$\left(\overline{\overline{\mathbf{A}}} + \delta\overline{\overline{\mathbf{A}}}\right)(\overline{\mathbf{x}} + \delta\overline{\mathbf{x}}) = \overline{\mathbf{b}}$$

Subtract unperturbed equation

$$\overline{\overline{\mathbf{A}}}\delta\overline{\mathbf{x}} + \delta\overline{\overline{\mathbf{A}}}(\overline{\mathbf{x}} + \delta\overline{\mathbf{x}}) = \overline{\mathbf{0}}$$

$$\delta\overline{\mathbf{x}} = -\overline{\overline{\mathbf{A}}}^{-1}\delta\overline{\overline{\mathbf{A}}}(\overline{\mathbf{x}} + \delta\overline{\mathbf{x}}) \simeq -\overline{\overline{\mathbf{A}}}^{-1}\delta\overline{\overline{\mathbf{A}}}\overline{\mathbf{x}}$$

$$||\delta\overline{\mathbf{x}}|| \leq \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| \left\|\delta\overline{\overline{\mathbf{A}}}\right\| ||\overline{\mathbf{x}}||$$

Relative Error Magnification

$$\frac{||\delta\overline{\mathbf{x}}||}{||\overline{\mathbf{x}}||} \leq \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| \left\|\overline{\overline{\mathbf{A}}}\right\| \frac{\left\|\delta\overline{\overline{\mathbf{A}}}\right\|}{\left\|\overline{\overline{\mathbf{A}}}\right\|}$$

Condition Number

$$K(\overline{\overline{\mathbf{A}}}) = \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\| \left\|\overline{\overline{\mathbf{A}}}\right\|$$

Ill-Conditioned System

$$\begin{bmatrix} 1.0 & 1.0 \\ 1.0 & 1.0001 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det(\overline{\overline{\mathbf{A}}}) = 0.0001$$

$$\begin{aligned} a_{11} &= \frac{1.0001}{0.0001} = 10,001 \\ a_{12} &= \frac{-1}{0.0001} = -10,000 \\ a_{21} &= \frac{-1}{0.0001} = -10,000 \\ a_{11} &= \frac{1.0}{0.0001} = 10,000 \end{aligned}$$

$$\left. \begin{aligned} \left\|\overline{\overline{\mathbf{A}}}\right\|_\infty &= 2.0001 \\ \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\|_\infty &= 20,001 \end{aligned} \right\} \Rightarrow K(\overline{\overline{\mathbf{A}}}) \simeq \boxed{40,000}$$

Ill-conditioned system

tbt6.m

```
n=4
a = [ [1.0 1.0]' [1.0 1.0001]']
b= [1 2]'

ai=inv(a);
a_nrm=max( abs(a(1,1)) + abs(a(1,2)) ,
           abs(a(2,1)) + abs(a(2,2)) )
ai_nrm=max( abs(ai(1,1)) + abs(ai(1,2)) ,
           abs(ai(2,1)) + abs(ai(2,2)) )
k=a_nrm*ai_nrm

r=ai * b

x=[0 0];
m21=a(2,1)/a(1,1);
a(2,1)=0;
a(2,2)  = radd(a(2,2),-m21*a(1,2),n);
b(2)    = radd(b(2),-m21*b(1),n);

x(2)    = b(2)/a(2,2);
x(1)    = (radd(b(1), -a(1,2)*x(2),n))/a(1,1);
x'
```

Well-Conditioned System

$$\begin{bmatrix} 0.0001 & 1.0 \\ 1.0 & 1.0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det(\overline{\overline{\mathbf{A}}}) = 0.9999$$

$$a_{11} = \frac{-1}{0.9999} = -1,0001$$

$$a_{12} = \frac{1}{0.9999} = 1.0001$$

$$a_{21} = \frac{1}{0.9999} = 1.0001$$

$$a_{11} = \frac{-0.0001}{0.9999} = -0.0001$$

$$\left.\begin{array}{rcr} \left\|\overline{\overline{\mathbf{A}}}\right\|_\infty & = & 2.0 \\ \left\|\overline{\overline{\mathbf{A}}}^{-1}\right\|_\infty & = & 2.0002 \end{array}\right\} \Rightarrow K(\overline{\overline{\mathbf{A}}}) \simeq \boxed{4}$$

Well-conditioned system

4-digit Arithmetic

tbt7.m

```
n=4
a = [ [0.0001 1.0]' [1.0 1.0]']
b= [1 2]'

ai=inv(a);
a_nrm=max( abs(a(1,1)) + abs(a(1,2)) ,
           abs(a(2,1)) + abs(a(2,2)) )
ai_nrm=max( abs(ai(1,1)) + abs(ai(1,2)) ,
            abs(ai(2,1)) + abs(ai(2,2)) )
k=a_nrm*ai_nrm

r=ai * b

x=[0 0];
m21=a(2,1)/a(1,1);
a(2,1)=0;
a(2,2)  = radd(a(2,2),-m21*a(1,2),n);
b(2)    = radd(b(2),-m21*b(1),n);

x(2)    = b(2)/a(2,2);
x(1)    = (radd(b(1), -a(1,2)*x(2),n))/a(1,1);
x'
```

Algorithmically ill-conditioned