Linear Algebraic Equations (Chapters 9,10,11,12)

General form of a system of linear algebraic equations

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

which can be rewritten as

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

or

$$\mathbf{AX} = \mathbf{b}$$

Example:

$$
\begin{aligned}
2x_1 + x_2 &= 5 \\
x_1 + 2x_2 &= 4
\end{aligned}
$$

can be rewritten as

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

where $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$.

Outline:

- Graphical method
- Cramer's rule
- Gauss elimination
- LU decomposition
- Cholesky decomposition
- Gauss-Seidel iteration
- Error analysis

1 Graphical Method

The simplest method to solve a set of two linear equations is to use the graphical method. For

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{1}$$

$$a_{21}x_1 + a_{22}x_2 = b_2 \tag{2}$$

From (1) we have

$$x_2 = -\frac{a_{11}}{a_{12}}x_1 + \frac{b_1}{a_{12}} \tag{3}$$

From (2) we have

$$x_2 = -\frac{a_{21}}{a_{22}}x_1 + \frac{b_2}{a_{22}} \tag{4}$$

where $-\frac{a_{11}}{a_{12}}$ and $-\frac{a_{21}}{a_{22}}$ are slopes of the lines and $\frac{b_1}{a_{12}}$ and $\frac{b_2}{a_{22}}$ are intercepts.
Example:

$$2x_1 + x_2 = 5 \;\longrightarrow\; x_2 = -2x_1 + 5$$

$$x_1 + 2x_2 = 4 \;\longrightarrow\; x_2 = -\frac{1}{2}x_1 + 2$$

Comments:

- Not precise; and not practical for 3-dimensions and above.

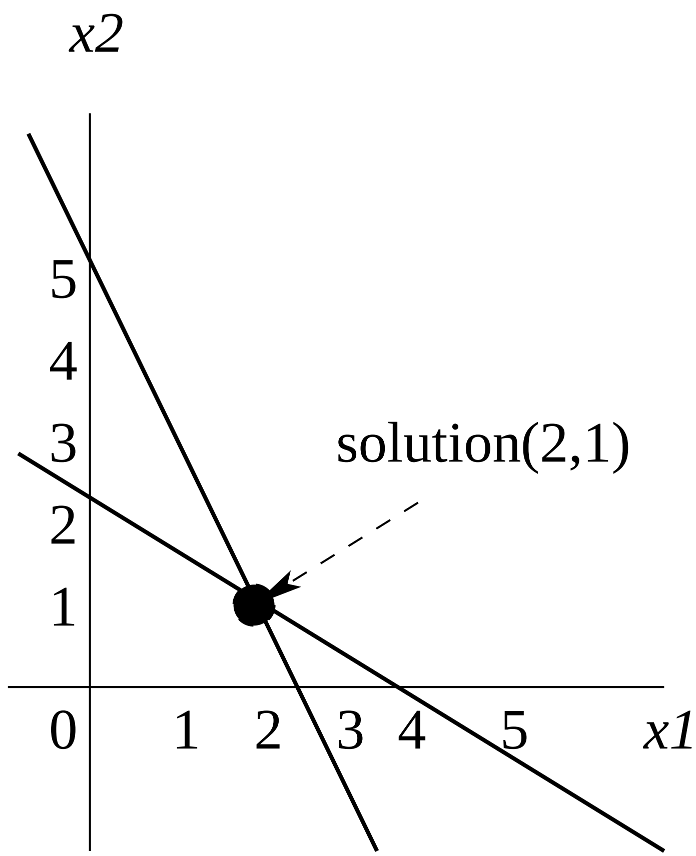

Figure 1: Example of using graphical method

2 Cramer's Rule

$\mathbf{AX} = \mathrm{b}.$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad d = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Cramer's rule uses the determinant to solve a set of linear equations.

For 3-dimensional case:

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

Solutions:

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

2-dimensional case:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Solutions:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

Example:

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

$$x_1 = \frac{\begin{vmatrix} 5 & 1 \\ 4 & 2 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}} = \frac{5 \times 2 - 1 \times 4}{2 \times 2 - 1 \times 1} = \frac{6}{3} = 2,$$

$$x_2 = \frac{\begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix}}{\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix}} = \frac{2 \times 4 - 1 \times 5}{2 \times 2 - 1 \times 1} = \frac{3}{3} = 1,$$

Comment: Cramer's rule is not feasible for larger values of n because of the difficulty in evaluating the determinants.

3 Gauss Elimination

Example:

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{5}$$

$$a_{21}x_1 + a_{22}x_2 = b_2 \tag{6}$$

(5)$\times a_{21}$:

$$a_{11}a_{21}x_1 + a_{12}a_{21}x_2 = b_1a_{21} \tag{7}$$

(6)$\times a_{11}$:

$$a_{11}a_{21}x_1 + a_{11}a_{22}x_2 = b_2a_{11} \tag{8}$$

(7)-(8)

$$(a_{12}a_{21} - a_{11}a_{22})x_2 = b_1a_{21} - b_2a_{11} \tag{9}$$

$$x_2 = \frac{b_1a_{21} - b_2a_{11}}{a_{12}a_{21} - a_{11}a_{22}} \tag{10}$$

Substituting back to (7),

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{12}a_{21} - a_{11}a_{22}} \tag{11}$$

Gauss elimination steps:

- Forward elimination
  - n unknowns: $n-1$ rounds of elimination
    The first round is to eliminate x_1 from equations (2) to (n)
    The second round is to eliminate x_2 from equations (3) to (n)
    $\ldots$
    The $(n-1)$th round is to eliminate x_{n-1} from equation (n)

- Back substitution
  First find x_n from the nth equation
  then find x_{n-1} from the $(n-1)$th equation
  $\ldots$
  then find x_2 from equation (2)
  finally find x_1 from equation (1)

<u>Forward elimination</u>

Original set of equations:

$$\begin{array}{lllllll} a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3+ & \ldots+ & a_{1,n-1}x_{n-1}+ & a_{1n}x_n = & b_1 \quad (1) \\ a_{21}x_1+ & a_{22}x_2+ & a_{23}x_3+ & \ldots+ & a_{2,n-1}x_{n-1}+ & a_{2n}x_n = & b_2 \quad (2) \\ & & & \ldots & & & \\ a_{n1}x_1+ & a_{n2}x_2+ & a_{n3}x_3+ & \ldots+ & a_{n,n-1}x_{n-1}+ & a_{nn}x_n = & b_n \quad (n) \end{array} \tag{12}$$

The first round of elimination: $(i) - (1) \times \frac{a_{i1}}{a_{11}}$, where (i) is from (2) to (n). Then the new equation (i) becomes

$$a'_{i2}x_2 + a'_{i3}x_3 + \ldots + a'_{i,n-1}x_{n-1} + a'_{in}x_n = b'_i,$$

where

$$a'_{ij} = a_{ij} - a_{1j} \times \frac{a_{i1}}{a_{11}}$$

$$b'_i = b_i - b_1 \times \frac{a_{i1}}{a_{11}}$$

for $i = 2, 3, \ldots, n$, $j = 2, 3, \ldots, n$.

Pivot element: a_{11}.

The full set of new equations after the first round of elimination is

$$\begin{array}{lllllll} a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3+ & \ldots+ & a_{1,n-1}x_{n-1}+ & a_{1n}x_n = & b_1 \quad (1) \\ & a'_{22}x_2+ & a'_{23}x_3+ & \ldots+ & a'_{2,n-1}x_{n-1}+ & a'_{2n}x_n = & b'_2 \quad (2') \\ & & & \ldots & & & \\ & a'_{n2}x_2+ & a'_{n3}x_3+ & \ldots+ & a'_{n,n-1}x_{n-1}+ & a'_{nn}x_n = & b'_n \quad (n') \end{array} \tag{13}$$

In general, the kth round of elimination eliminates x_k from the $(k+1)$th equation to the nth equation. That is,

$$(i^{(k-1)}) - (k^{(k-1)}) \times \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

where (i) is from $(k+1)$ to (n). Then the new equation (i) (or equation $(i^{(k)})$) becomes

$$a_{i,k+1}^{(k)} x_{k+1} + \ldots + a_{i,n-1}^{(k)} x_{n-1} + a_{i,n}^{(k)} x_n = b_i^{(k)},$$

where

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{kj}^{(k-1)} \times \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

$$b_i^{(k)} = b_i^{(k-1)} - b_k^{(k-1)} \times \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

for $i = k+1, k+2, \ldots, n$, $j = k+1, k+2, \ldots, n$.
Pivot element: $a_{kk}^{(k-1)}$.

After the $(n-1)th$ round of elimination:

$$\begin{array}{llllll} a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3+ & \ldots+ & a_{1,n-1}x_{n-1}+ & a_{1n}x_n = & b_1 & (1) \\ & a_{22}^{'}x_2+ & a_{23}^{'}x_3+ & \ldots+ & a_{2,n-1}^{'}x_{n-1}+ & a_{2n}^{'}x_n = & b_2^{'} & (2^{'}) \\ & & & \ldots & & & & \\ & & & & & a_{nn}^{(n-1)}x_n = & b_n^{(n-1)} & (n^{(n-1)}) \end{array} \quad (14)$$

<u>Back substitution</u>

From equation $(n^{(n-1)})$, we have $x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}}$.

From equation $((n-1)^{(n-2)})$, we have $x_{n-1} = \frac{b_{n-1}^{(n-2)} - a_{n-1,n}^{(n-2)}x_n}{a_{n-1,n-1}^{(n-2)}}$.

$\ldots$

In general,

$$x_i = \frac{1}{a_{ii}^{(i-1)}}\left[b_i^{(i-1)} - a_{i,i+1}^{(i-1)}x_{i+1} - \cdots - a_{i,n-1}^{(i-1)}x_{n-1} - a_{in}^{(i-1)}x_n\right]$$

for $i = n-1, n-2, \ldots, 1$.

Comment: Most operations are for eliminations. As n increases, computational load increases.

Example:

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 3 & (1) \\ 2x_1 + x_2 + 2x_3 - x_4 = 7 & (2) \\ 2x_1 - x_2 + x_3 + 2x_4 = -1 & (3) \\ x_1 - 2x_2 + x_3 - 2x_4 = 0 & (4) \end{cases}$$

Solution:

$(2) - (1) \times \frac{a_{21}}{a_{11}}$, and $\frac{a_{21}}{a_{11}} = 2$,

$$-3x_2 + 4x_3 - 3x_4 = 1 \;\; (2')$$

$(3) - (1) \times \frac{a_{31}}{a_{11}}$, and $\frac{a_{31}}{a_{11}} = 2$,

$$-5x_2 + 3x_3 = -7 \;\; (3')$$

$(4) - (1) \times \frac{a_{41}}{a_{11}}$, and $\frac{a_{41}}{a_{11}} = 1$,

$$-4x_2 + 2x_3 - 3x_4 = -3 \;\; (4')$$

$$\begin{cases} x_1 & +2x_2 & -x_3 & +x_4 & = 3 & (1) \\ & -3x_2 & +4x_3 & -3x_4 & = 1 & (2') \\ & -5x_2 & +3x_3 & & = -7 & (3') \\ & -4x_2 & +2x_3 & -3x_4 & = -3 & (4') \end{cases}$$

$(3^{'}) - (2^{'}) \times \frac{a^{'}_{32}}{a^{'}_{22}}$ and $(4^{'}) - (2^{'}) \times \frac{a^{'}_{42}}{a^{'}_{22}}$

$$\begin{cases} x_1 & +2x_2 & -x_3 & +x_4 & = 3 & (1) \\ & -3x_2 & +4x_3 & -3x_4 & = 1 & (2^{'}) \\ & & -\frac{11}{3}x_3 & +5x_4 & = -\frac{26}{3} & (3^{''}) \\ & & -\frac{10}{3}x_3 & +x_4 & = -\frac{13}{3} & (4^{''}) \end{cases}$$

$(4^{''}) - (3^{''}) \times \frac{a^{''}_{43}}{a^{''}_{33}}$

$$\begin{cases} x_1 & +2x_2 & -x_3 & +x_4 & = 3 & (1) \\ & -3x_2 & +4x_3 & -3x_4 & = 1 & (2^{'}) \\ & & -\frac{11}{3}x_3 & +5x_4 & = -\frac{26}{3} & (3^{''}) \\ & & & -\frac{39}{11}x_4 & = \frac{39}{11} & (4^{'''}) \end{cases}$$

From $(4^{'''})$, $x_4 = -1$.
From $(3^{''})$, $x_3 = -\frac{3}{11}(-\frac{26}{3} - 5x_4) = 1$.
From $(2^{'})$, $x_2 = -\frac{1}{3}(1 - 4x_3 + 3x_4) = 2$.
From (1), $x_1 = 3 - 2x_2 + x_3 - x_4 = 1$.

Example:

$$\begin{cases} 0.1x_2 + 0.2x_3 = 1.1 & (1) \\ 5x_1 + x_2 + 3x_3 = 25 & (2) \\ x_1 + 2x_2 + x_3 = 12 & (3) \end{cases}$$

Cannot do elimination since $a_{11} = 0$. Exchange positions of equations (1) and (2):

$$\begin{cases} 5x_1 + x_2 + 3x_3 = 25 & (1) \\ 0.1x_2 + 0.2x_3 = 1.1 & (2) \\ x_1 + 2x_2 + x_3 = 12 & (3) \end{cases}$$

$(3) - (1) \times \frac{a_{31}}{a_{11}}, \frac{a_{31}}{a_{11}} = \frac{1}{5},$

$$\begin{cases} 5x_1 + x_2 + 3x_3 = 25 & (1) \\ 0.1x_2 + 0.2x_3 = 1.1 & (2) \\ \frac{9}{5}x_2 + \frac{2}{5}x_3 = 7 & (3^{'}) \end{cases}$$

$(3^{'}) - (2) \times \frac{a_{32}^{'}}{a_{22}}, \frac{a_{32}^{'}}{a_{22}} = \frac{1.8}{0.1} = 18,$

$$\begin{cases} 5x_1 + x_2 + 3x_3 = 25 & (1) \\ 0.1x_2 + 0.2x_3 = 1.1 & (2) \\ -3.2x_3 = -12.8 & (3^{''}) \end{cases}$$

$x_3 = 4$, $x_2 = 2$, $x_1 = 2$.

Pivoting: switching rows so that the pivot element in each round of elimination is non-zero (maximum).

Pivoting results in better results when $a_{ii} \approx 0$, since it avoids division by small numbers during elimination.

4 LU Decomposition

In Gauss elimination,

- more than 90% operations are for elimination,
- both A and b are modified during the elimination process,
- to solve $AX = b$ and $AY = b^{'}$, the same elimination process has to be repeated for A.

LU decomposition records the elimination process information, so that it can be used later.

Consider $AX = b$. After Gauss elimination we have

$$
\begin{array}{cccccccc}
a_{11}x_1+ & a_{12}x_2+ & a_{13}x_3+ & \ldots+ & a_{1,n-1}x_{n-1}+ & a_{1n}x_n = & b_1 & (1) \\
 & a^{'}_{22}x_2+ & a^{'}_{23}x_3+ & \ldots+ & a^{'}_{2,n-1}x_{n-1}+ & a^{'}_{2n}x_n = & b^{'}_2 & (2^{'}) \\
 & & & \ldots & & & & \\
 & & & & & a^{(n-1)}_{nn}x_n = & b^{(n-1)}_n & (n^{(n-1)})
\end{array}
\tag{15}
$$

which can be written as

$$UX = d \quad (*)$$

where

$$U = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} \end{bmatrix}, \quad d = \begin{bmatrix} b_1 \\ b'_2 \\ \vdots \\ b_n^{(n-1)} \end{bmatrix}$$

Premultiplying (*) by matrix L, (L is an $n \times n$ matrix)

$$LUX = Ld$$

Comparing with $AX = b$, we have

$$LU = A \text{ and } Ld = b$$

where L is defined as a special lower triangular matrix carrying the elimination information as

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{a_{21}}{a_{11}} & 1 & 0 & \cdots & 0 \\ \frac{a_{31}}{a_{11}} & \frac{a'_{32}}{a'_{22}} & 1 & \cdots & 0 \\ \cdots & \cdots & & \ddots & \\ \frac{a_{n1}}{a_{11}} & \frac{a'_{n2}}{a'_{22}} & \cdots & & 1 \end{bmatrix}, \quad \text{or} \quad l_{ij} = \begin{cases} 0, & i < j \\ 1, & i = j \\ \frac{a_{ij}^{(j-1)}}{a_{jj}^{(j-1)}}, & i > j \end{cases}$$

Solving $AX = b$ using LU decomposition

- Decomposition
  Do Gauss elimination to find L (lower triangular matrix) and U (upper triangular matrix) so that $A = LU$
- Substitution
  - Forward substitution
    From $Ld = b$ to find d

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{a_{21}}{a_{11}} & 1 & 0 & \cdots & 0 \\ \frac{a_{31}}{a_{11}} & \frac{a'_{32}}{a'_{22}} & 1 & \cdots & 0 \\ \cdots & \cdots & & \ddots & \\ \frac{a_{n1}}{a_{11}} & \frac{a'_{n2}}{a'_{22}} & \cdots & & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Find $d_1 \rightarrow d_2 \rightarrow \ldots \rightarrow d_{n-1} \rightarrow d_n$.
The ith row:

$$b_i = \sum_{k=1}^{n} l_{ik} d_k = \sum_{k=1}^{i} l_{ik} d_k = d_i + \sum_{k=1}^{i-1} l_{ik} d_k$$

Then

$$d_i = b_i - \sum_{k=1}^{i-1} l_{ik} d_k$$

– Backward substitution
From $UX = d$ to find X

$$U = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{'} & \dots & a_{2n}^{'} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Find $x_n \to x_{n-1} \to \dots \to x_2 \to x_1$.
The ith row:

$$d_i = \sum_{j=1}^{n} u_{ij} x_j = \sum_{j=i}^{n} u_{ij} x_j = u_{ii} x_i + \sum_{j=i+1}^{n} u_{ij} x_j$$

Then

$$x_i = \frac{d_i - \sum_{j=i+1}^{n} u_{ij} x_j}{u_{ii}}$$

Example: $A = [a_{ij}]_{3\times 3}$, find A^{-1} so that $A^{-1}A = I$.

Let

$$A^{-1} = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}$$

$$A \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$AX = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad AY = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad AZ = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

- Gauss elimination, $A = LU$, find L and U.
- $b = [1\ 0\ 0]'$. From $Ld = b$ find d; and from $UX = d$ find X
- $b = [0\ 1\ 0]'$. From $Ld = b$ find d; and from $UY = d$ find Y
- $b = [0\ 0\ 1]'$. From $Ld = b$ find d; and from $UZ = d$ find Z

Example:

$$
\begin{aligned}
2x_1 - 2x_2 \quad + 4x_4 &= 2 \\
3x_1 - 3x_2 \quad - x_4 &= -18 \\
-x_1 + 6x_2 + 5x_3 - 7x_4 &= -26 \\
-5x_1 + x_2 \quad + 6x_4 &= 7
\end{aligned}
$$

Solution:

Gauss elimination

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

Row (2) - row (1) $\times\frac{3}{2}$, row (3) - row (1) $\times\frac{-1}{2}$, and row (4) - row (1) $\times\frac{-5}{2}$,

$$
\begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 0 & 0 & -7 \\ 0 & 5 & 5 & -5 \\ 0 & -4 & 0 & 16 \end{bmatrix}
$$

Exchange positions of row (2) and row (3):

$$\begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 5 & 5 & -5 \\ 0 & 0 & 0 & -7 \\ 0 & -4 & 0 & 16 \end{bmatrix}$$

row (4) - row (2) $\times \frac{-4}{5}$,

$$\begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 5 & 5 & -5 \\ 0 & 0 & 0 & -7 \\ 0 & 0 & 4 & 12 \end{bmatrix}$$

Exchange positions of row (3) and row (4):

$$\begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 5 & 5 & -5 \\ 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & -7 \end{bmatrix} = U,$$

$L = ?$

$LU = PA$, where P is the permutation matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \overset{(2) \leftrightarrow (3)}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \overset{(3) \leftrightarrow (4)}{\longrightarrow} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} = P$$

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

and

$$LU = \begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 5 & 5 & -5 \\ 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & -7 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 0 & 4 \\ -1 & 6 & 5 & -7 \\ -5 & 1 & 0 & 6 \\ 3 & -3 & 0 & -1 \end{bmatrix},$$

Row 2 and column 1: $l_{21} \times 2 = -1$, $l_{21} = -\frac{1}{2}$.
Row 3 and column 1: $l_{31} \times 2 = -5$, $l_{31} = -\frac{5}{2}$
Row 4 and column 1: $l_{41} \times 2 = 3$, $l_{41} = \frac{3}{2}$
Row 3 and column 2: $l_{31} \times (-2) + l_{32} \times 5 = 1$, $l_{32} = -\frac{4}{5}$
Row 4 and column 2: $l_{41} \times (-2) + l_{42} \times 5 = -3$, $l_{42} = 0$

Row 4 and column 3: $l_{41} \times (0) + l_{42} \times 5 + l_{43} \times 4 = 0$, $l_{43} = 0$

Then L is

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

$Ld = b$, $d = ?$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ -5/2 & -4/5 & 1 & 0 \\ 3/2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -26 \\ 7 \\ -18 \end{bmatrix}$$

$d_1 = 2$,
$-\frac{1}{2}d_1 + d_2 = -26$, $d_2 = -25$,
$-\frac{5}{2} - \frac{4}{5}d_2 + d_3 = 7$, $d_3 = -8$
$\frac{3}{2}d_1 + d_4 = -18$, $d_4 = -21$.

$UX = d,\ X = ?$

$$\begin{bmatrix} 2 & -2 & 0 & 4 \\ 0 & 5 & 5 & -5 \\ 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -25 \\ -8 \\ -21 \end{bmatrix}$$

$-7x_4 = -21,\ x_4 = 3$

$4x_3 + 12x_4 = -8,\ x_3 = -11$

$5x_2 + 5x_3 - 5x_4 = -25,\ x_2 = 9$

$2x_1 - 2x_2 + 4x_4 = 2,\ x_1 = 4.$

5 Cholesky Decomposition

Cholesky decomposition is another (efficient) way to implement LU decomposition for symmetric matrices.

Consider $AX = b$, $A = [a_{ij}]_{n\times n}$, and $a_{ij} = a_{ji}$ ($A^{'} = A$).

Chokesky decomposition: $A = LL^{'}$, where

$$L = \begin{bmatrix} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & & \\ \dots & & \ddots & 0 \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{bmatrix}$$

Let l_{ki} be the kth row and ith column entry of L, then

$$l_{ki} = \begin{cases} 0, & k < i \\ \sqrt{a_{ii} - \sum_{j=1}^{i-1} l_{kj}^2}, & k = i \\ \frac{1}{l_{ii}}\left(a_{ki} - \sum_{j=1}^{i-1} l_{ij}l_{kj}\right), & i < k \end{cases}$$

Orders for finding l_{ki}'s

- Row by row

  1). l_{11}

  2). $l_{21} \rightarrow l_{22}$

  3). $l_{31} \rightarrow l_{32} \rightarrow l_{33}$

  ...). $\ldots$

  n). $l_{n1} \rightarrow l_{n2} \rightarrow \ldots \rightarrow l_{nn}$

- Column by column

  1). $l_{11} \rightarrow l_{21} \rightarrow l_{31} \rightarrow \ldots \rightarrow l_{n1}$

  2). $l_{22} \rightarrow l_{32} \rightarrow \ldots \rightarrow l_{n2}$

  ...). $\ldots$

  n-1). $l_{n-1,n-1} \rightarrow l_{n,n-1}$

  n). l_{nn}

Using Cholesky decomposition to solve $AX = b$, where $A = A^{'}$

- Find L and $L^{'}$, $A = LL^{'}$
- Forward substitution
  $Ld = b$, find d
- Back substitution
  $L^{'}X = d$, find X

Comments:

- Cholesky decomposition fails when

$$a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2 < 0$$

- Sufficient condition:
  When A is a positive definite matrix, $a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2 \geq 0$.

Example: In the above figure, $R_1 = R_2 = R_3 = R_4 = 5$, $R_5 = R_6 = R_7 = R_8 = 2$, $V_1 = V_2 = 5$, find i_1, i_2, i_3 and i_4.

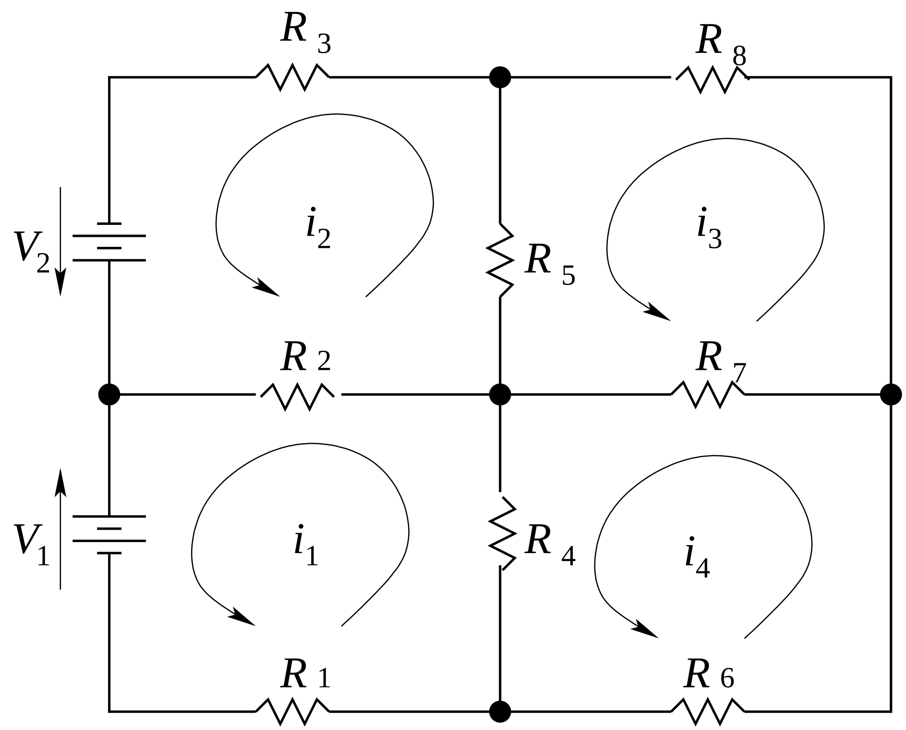

Figure 2: Example

Solution: Using Kirchoff law,

$$(i_2 - i_1)R_2 + (i_4 - i_1)R_4 - i_1 R_1 = V_1, \quad (1)$$
$$(i_2 - i_1)R_2 + (i_2 - i_3)R_5 + i_2 R_3 = V_2, \quad (2)$$
$$(i_4 - i_3)R_7 + (i_2 - i_3)R_5 - i_3 R_8 = 0, \quad (3)$$
$$(i_4 - i_3)R_7 + (i_4 - i_1)R_4 + i_4 R_6 = 0, \quad (4)$$

Rewrite the equations,

$$\begin{aligned}
(R_1+R_2+R_4)i_1 - R_2 i_2 - R_4 i_4 &= -V_1 \\
-R_2 i_1 + (R_2+R_3+R_5)i_2 - R_5 i_3 &= V_2 \\
-R_5 i_2 + (R_5+R_7+R_8)i_3 - R_7 i_4 &= 0 \\
-R_4 i_1 - R_7 i_3 + (R_4+R_6+R_7)i_4 &= 0
\end{aligned}$$

$$A = \begin{bmatrix} R_1+R_2+R_4 & -R_2 & 0 & -R_4 \\ -R_2 & R_2+R_3+R_5 & -R_5 & 0 \\ 0 & -R_5 & R_5+R_7+R_8 & -R_7 \\ -R_4 & 0 & -R_7 & R_4+R_6+R_7 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & -5 & 0 & -5 \\ -5 & 12 & -2 & 0 \\ 0 & -2 & 6 & -2 \\ -5 & 0 & -2 & 9 \end{bmatrix}, \qquad b = \begin{bmatrix} -5 \\ 5 \\ 0 \\ 0 \end{bmatrix}$$

Decompose $A = LL'$:

$l_{11} = \sqrt{a_{11}} = \sqrt{15} = 3.873$

$l_{21} = \frac{1}{l_{11}} a_{21} = -1.291$

$l_{31} = \frac{1}{l_{11}} a_{31} = 0$

$l_{41} = \frac{1}{l_{11}} a_{41} = -1.291$

$$l_{22} = \sqrt{a_{22} - l_{21}^2} = 3.215$$
$$l_{32} = \frac{1}{l_{22}}(a_{32} - l_{21}l_{31}) = -0.622$$
$$l_{42} = \frac{1}{l_{22}}(a_{42} - l_{21}l_{41}) = -0.518$$
$$l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = 2.369$$
$$l_{43} = \frac{1}{l_{33}}(a_{43} - l_{31}l_{41} - l_{32}l_{42}) = -0.980$$
$$l_{44} = \sqrt{a_{44} - l_{41}^2 - l_{42}^2 - l_{43}^2} = 2.471.$$

$$L = \begin{bmatrix} 3.873 & 0 & 0 & 0 \\ -1.291 & 3.215 & 0 & 0 \\ 0 & -0.622 & 2.369 & 0 \\ -1.291 & -0.518 & -0.980 & 2.471 \end{bmatrix}$$

$$Ld = b, \rightarrow, d$$
$$L^{'}X = d, \rightarrow, X$$

6 Gauss-Seidel Iteration

Example:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \quad (1) \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \quad (2) \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \quad (3)
\end{aligned}
$$

From (1), $x_1 = \frac{b_1 - a_{12}x_2 - a_{13}x_3}{a_{11}} \quad (4)$
From (2), $x_2 = \frac{b_2 - a_{21}x_1 - a_{23}x_3}{a_{22}} \quad (5)$
From (3), $x_3 = \frac{b_3 - a_{31}x_1 - a_{32}x_2}{a_{33}} \quad (6)$

Steps:

1. Initial guess x_2, x_3
2. Update x_1 using (4)
3. Update x_2 using (5)
4. Update x_3 using (6)
5. If $\epsilon_i < \epsilon_{threshold}$ for all $i = 1, 2, 3$, end; otherwise, repeat step 2.

Comment: The Gauss-Seidel method does not always converge.

Example: (a).

$$11x_1 + 9x_2 = 99 \;\; (v)$$
$$11x_1 + 13x_2 = 286 \;\; (u)$$

From (v), $x_1 = \frac{99-9x_2}{11}$

From (u), $x_2 = \frac{286-11x_1}{13}$

$x_1 = 0 \;\rightarrow\; x_2 = \frac{286-11x_1}{13} \;\rightarrow x_1 = \frac{99-9x_2}{11}$

The Gauss-Seidel method converges.

(b).

$$11x_1 + 13x_2 = 286 \;\; (u)$$
$$11x_1 + 9x_2 = 99 \;\; (v)$$

From (u), $x_1 = \frac{286-13x_2}{11}$

From (v), $x_2 = \frac{99-11x_1}{9}$

$x_1 = 0 \;\rightarrow\; x_2 = \frac{99-11x_1}{9} \;\rightarrow x_1 = \frac{286-13x_2}{11}$

The Gauss-Seidel method diverges.

Sufficient (NOT necessary) condition: If $|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|$ for all i, the Gauss-Seidel approach converges. That is, the diagonal coefficient in each equation must be larger than the sum of the absolute values of all other coefficients in the equation.

$$
\begin{aligned}
|a_{11}| &> |a_{12}| + |a_{13}| + \ldots + |a_{1n}| \\
|a_{22}| &> |a_{21}| + |a_{23}| + \ldots + |a_{2n}| \\
&\ldots \\
|a_{nn}| &> |a_{n1}| + |a_{n2}| + \ldots + |a_{n,n-1}|
\end{aligned}
$$

7 Error Analysis for Solving a Set of Linear Equations

Consider $AX = b$,

- When $|A| \neq 0$, A is non-singular, there is a unique solution.
- When $|A| = 0$, A is singular, there is no solution or an infinite number of solutions.
- When $|A| \approx 0$, the solution is sensitive to numerical errors.

"$A_{n\times n}$ is non-singular" is equivalent to

- A has an inverse. Then $X = A^{-1}b$.
- $|A| \neq 0$
- A has full rank, or $\mathrm{rank}(A) = n$.
- All n rows in A are linear independent, and all n columns in A are linear independent.
- For any $Z_{n\times 1} \neq 0$, $AZ \neq 0$.

If $A_{n\times n}$ is singular, then

- $|A| = 0$

- A does have an inverse
- $\text{rank}(A) < n$
- There exists $Z_{n\times 1} \neq 0$, so that $AZ = 0$.
- For $AX = b$, either there is no solution or there is an infinite number of solutions.
  Proof: If A is singular, then there exists $Z_{n\times 1} \neq 0$ so that $AZ = 0$. If there is X_1 so that $AX_1 = b$, then $A(X_1 + \gamma Z) = AX_1 + \gamma AZ = b$, or $X_1 + \gamma Z$ is a solution for $AX = b$. Since γ can be any scaler, $AX = b$ has an infinite number of solutions.

Example:

$$2x_1 + 3x_2 = 4$$
$$4x_1 + 6x_2 = 8$$

$A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$, $|A| = 0$. $X_1 = [2 \quad 0]'$ is one solution.

Find Z so that $AZ = 0$.

$$\begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

then $Z = [-3 \quad 2]'$ and $X = X_1 + \gamma Z = [2 - 3\gamma \quad 2\gamma]'$.

Example:

$$2x_1 + 3x_2 = 4 \quad (1)$$
$$4x_1 + 6x_2 = 7 \quad (2)$$

$|A| = 0$, no solution.

Linear dependent

Consider n vectors, $V_1, V_2, \ldots, V_n$,

- If there exist $\alpha_1, \alpha_2, \ldots, \alpha_n$ (not all zeros), such that
$$\alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n = 0$$
then $V_1, V_2, \ldots, V_n$ are linear dependent. That is, at least one vector can be derived linearly from others.
- If $V_1, V_2, \ldots, V_n$ are linear independent and
$$\alpha_1 V_1 + \alpha_2 V_2 + \ldots + \alpha_n V_n = 0,$$
then $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

Example: $A_{3\times 3}$, $AZ = 0$, $|A| = 0$. There exists $Z \neq 0$ so that $AZ = 0$.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} z_1 + \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} z_2 + \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} z_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3 columns vectors in A are linear dependent.

<u>Vector Norms</u>

Consider vector $X_{n\times 1} = [x_1, x_2, \ldots, x_n]'$

The p-norm of X is defined as

$$||X||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

where p is an integer.

1-norm: $||X||_1 = \sum_{i=1}^{n} |x_i|$

2-norm: $||X||_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{\frac{1}{2}}$

Special, ∞-norm: $||X||_\infty = \max_{1\leq i\leq n} |x_i|$

Example: $X = [1.6 \;\; 1.2]'$.

$||X||_1 = |-1.6| + |1.2| = 2.8$

$||X||_2 = \sqrt{|-1.6|^2 + |1.2|^2} = 2$

$||X||_\infty = \max\{|-1.6|, |1.2|\} = 1.6$

Properties:

- If $X \neq 0$, $||X|| > 0$.
- For any X, $||X||_1 \geq ||X||_2 \geq ||X||_\infty$.
  Special: $X = [x_1 \quad x_2]'$. $||X||_1 = |x_1| + |x_2|$, $||X||_2 = \sqrt{|x_1|^2 + |x_2|^2}$, $||X||_\infty = \max\{|x_1|, |x_2|\}$.
- $||\gamma X|| = |\gamma| \cdot ||X||$
- $||X + Y|| \leq ||X|| + ||Y||$

Matrix Norms

p-norm of matrix A is defined as

$$||A||_p = \max_{X \neq 0} \frac{||AX||_p}{||X||_p}$$

$||A||_p$ represents the maximum ratio that the p-norm of vector X can be changed after multiplying by A.

Special:
$||A||_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$, column-sum norm

$||A||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$, row-sum norm

Example:

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{bmatrix}$$

$||A||_1 = \max_j \sum_{i=1}^{3} |a_{ij}| = \max\{2+1+3, 1+0+1, 1+1+4\} = 6$
$||A||_\infty = \max_i \sum_{j=1}^{n} |a_{ij}| = \max\{2+1+1, 1+0+1, 3+1+4\} = 8.$

Properties:

- If $A \neq 0_{n\times n}$, then $||A|| > 0$.
- $||\gamma A|| = |\gamma| \cdot ||A||$, γ is any scalar.
  $||\gamma A|| = \max_{X\neq 0} \frac{||\gamma AX||}{||X||} = \max_{X\neq 0} \frac{|\gamma|\cdot||AX||}{||X||} = |\gamma| \cdot ||A||.$
- $||AX|| \leq ||A|| \cdot ||X||$ for any $X \neq 0$.
  $||A|| = \max_{X\neq 0} \frac{||AX||}{||X||} \quad \rightarrow ||A|| \geq \frac{||AX||}{||X||} \quad \rightarrow ||AX|| \leq ||A|| \cdot ||X||.$

Matrix Condition Number

The condition number of matrix A is defined as

$$\text{cond}(A) = ||A|| \cdot ||A^{-1}||$$

When A is singular, A^{-1} does not exist, and $\mathrm{cond}(A) = \infty$.

Typically, consider 1-norm and ∞-norm.

Example:

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 3 & -1 & 4 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} 0.5 & 1.5 & -0.5 \\ -0.5 & 2.5 & -0.5 \\ -0.5 & -0.5 & 0.5 \end{bmatrix}$$

$||A^{-1}||_1 = \max_j \sum_{i=1}^{3} |\hat{a}_{ij}| = \max\{0.5{+}0.5{+}0.5, 1.5{+}2.5{+}0.5, 0.5{+}0.5{+}0.5\} = 4.5$

$||A^{-1}||_\infty = \max_i \sum_{j=1}^{3} |\hat{a}_{ij}| = \max\{0.5{+}1.5{+}0.5, 0.5{+}2.5{+}0.5, 0.5{+}0.5{+}0.5\} = 3.5$.

$\mathrm{cond}_1(A) = ||A||_1 \cdot ||A^{-1}||_1 = 6 \times 4.5 = 27$

$\mathrm{cond}_\infty(A) = ||A||_\infty \cdot ||A_\infty^{-1}||_1 = 8 \times 3.5 = 28$

Condition number and eigenvalues:

X and λ are eigenvector and corresponding eigenvalue of A

- $AX = \lambda X$, $||AX|| = |\lambda| \cdot ||X||$, $|\lambda| = \frac{||AX||}{||X||}$, and $|\lambda|_{\mathrm{max}} = \max_X \frac{||AX||}{||X||}$.

- $X = \lambda A^{-1}X$, $|A| \neq 0$, then $\lambda^{-1}X = A^{-1}X$, $|\lambda^{-1}| \cdot ||X|| = ||A^{-1}X||$, $|\lambda^{-1}| = \frac{||A^{-1}X||}{||X||}$, $|\lambda^{-1}|_{\max} = \max_X \frac{||A^{-1}X||}{||X||} = ||A^{-1}||$.

$$\begin{aligned}
\text{cond}(A) &= ||A|| \cdot ||A^{-1}|| = \max_{X \neq 0} \frac{||AX||}{||X||} \cdot \max_{X \neq 0} \frac{||A^{-1}X||}{||X||} \\
&= \frac{|\lambda|_{\max}}{|\lambda|_{\min}} = \frac{\max_{X \neq 0} \frac{||AX||}{||X||}}{\min_{X \neq 0} \frac{||AX||}{||X||}}
\end{aligned}$$

Comment: Condition number of matrix A is the ratio of the maximum change and the minimum change to vector norm when multiplying A to a vector.

Example:

1) $A_1 = \begin{bmatrix} 0.87 & 0.5 \\ -0.5 & 0.87 \end{bmatrix}$, $X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\text{cond}(A_1) = 1$.

$$A_1 X_1 = \begin{bmatrix} 0.87 & 0.5 \\ -0.5 & 0.87 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.87 \\ -0.5 \end{bmatrix} = Y_1$$

$$A_1 X_2 = \begin{bmatrix} 0.87 & 0.5 \\ -0.5 & 0.87 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.87 \end{bmatrix} = Y_2$$

$2) A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}$, $\text{cond}(A_2) = 4.$

$$A_2X_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = Y_1$$

$$A_2X_2 = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix} = Y_2$$

$3) A_3 = \begin{bmatrix} 1.73 & 0.25 \\ -1 & 0.43 \end{bmatrix}$, $\text{cond}(A_3) = 4.$

$4) A_4 = \begin{bmatrix} 1.52 & 0.91 \\ 0.47 & 0.94 \end{bmatrix}$, $\text{cond}(A_4) = 4.$

Comments:

- A matrix with a large condition number is nearly singular, whereas a matrix with a condition number close to 1 is far from singular.
- $\text{cond}(A) = \text{cond}(A^{-1})$ for $|A| \neq 0$.
- If A is close to singular, A^{-1} is also close to singular.

<u>Error Bounds and Sensitivity in Solving $AX = b$</u>

Sensitivity: If there is a small disturbance in b, e.g., truncation errors, how much solution X is affected?

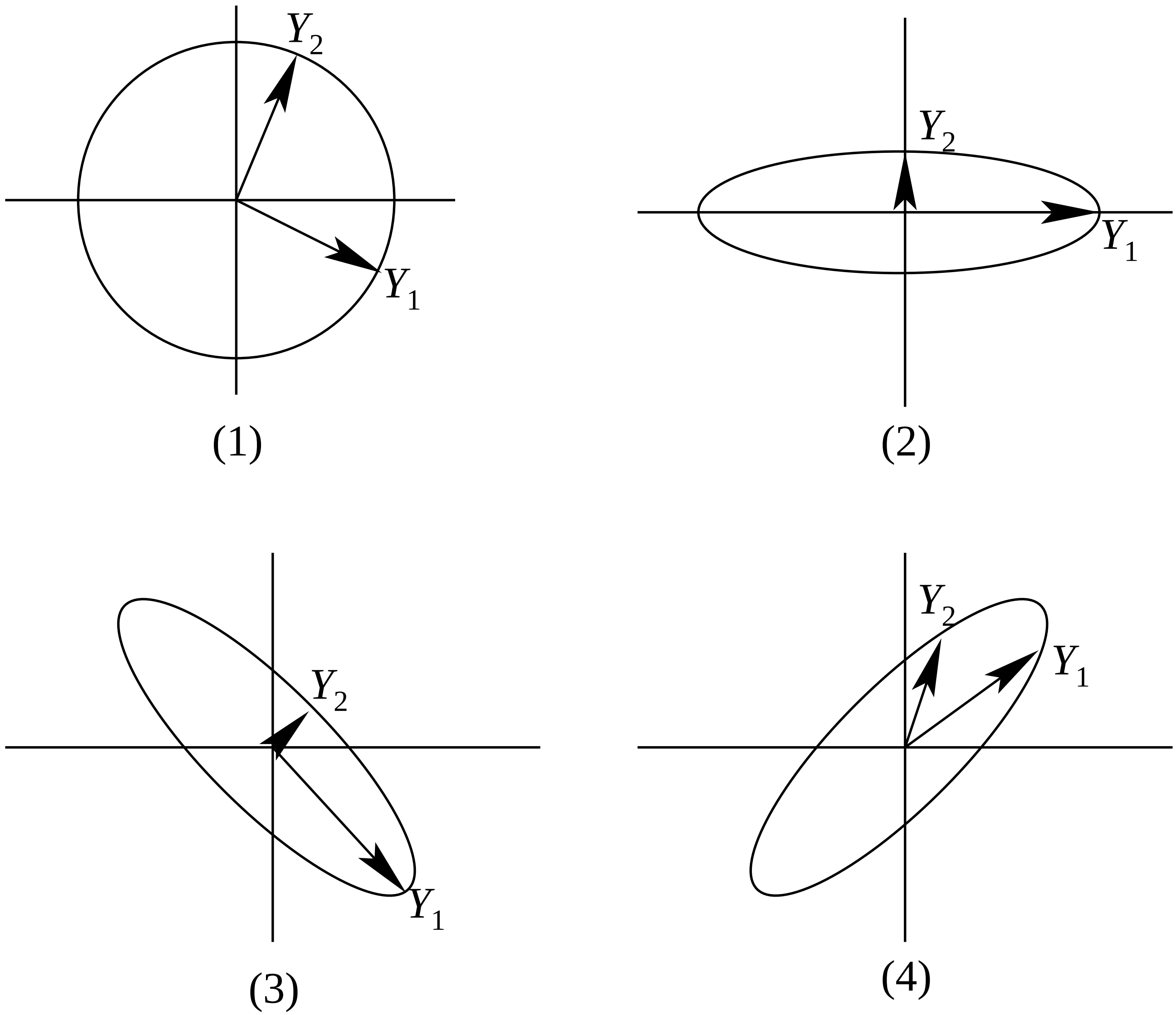

Figure 3: Distortion of a circle into an ellipse (by multiplying a matrix)

$AX = b \quad \rightarrow A(X + \Delta X) = b + \Delta b$
$A\Delta X = \Delta b \quad \rightarrow \Delta X = A^{-1}\Delta b$
$||\Delta X|| = ||A^{-1}\Delta b|| \leq ||A^{-1}|| \cdot ||\Delta b||$
$||AX|| = ||b|| \leq ||A|| \cdot ||X||$, or $||X|| \geq \frac{||b||}{||A||}$
$\frac{||\Delta X||}{||X||} \leq ||A^{-1}|| \cdot ||\Delta b|| \cdot \frac{||A||}{||b||} = ||A|| \cdot ||A^{-1}|| \cdot \frac{||\Delta b||}{||b||}$
$\frac{||\Delta X||}{||X||} \leq \mathrm{cond}(A)\frac{||\Delta b||}{||b||}$
As $\mathrm{cond}(A)$ increases, the effect of change in b will be high in solution — more sensitive to disturbance.
Example, if $\frac{||\Delta b||}{||b||} = 10^{-4}$, $\mathrm{cond}(A) = 10^4$, then $\frac{||\Delta X||}{||X||} \leq 1$.
A ill-conditioned System is a system where a small change in coefficients can result in large changes in solution.

Example:

(1)

$$x_1 + 2x_2 = 10 \quad (1)$$
$$1.1x_1 + 2x_2 = 10.4 \quad (2)$$

Using Cramer's rule, $x_1 = \dfrac{\begin{vmatrix} 10 & 2 \\ 10.4 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1.1 & 2 \end{vmatrix}} = \frac{10\times2-10.4\times2}{1\times2-1.1\times2} = 4$

$$x_2 = \dfrac{\begin{vmatrix} 1 & 10 \\ 1.1 & 10.4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1.1 & 2 \end{vmatrix}} = \frac{1\times10.4-1.1\times10}{1\times2-1.1\times2} = 3$$

(2)

$$x_1 + 2x_2 = 10 \quad (1)$$
$$1.05x_1 + 2x_2 = 10.4 \quad (2)$$

Using Cramer's rule, $x_1 = \dfrac{\begin{vmatrix} 10 & 2 \\ 10.4 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1.05 & 2 \end{vmatrix}} = \frac{10\times2-10.4\times2}{1\times2-1.05\times2} = 8$

$$x_2 = \dfrac{\begin{vmatrix} 1 & 10 \\ 1.05 & 10.4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 \\ 1.05 & 2 \end{vmatrix}} = \frac{1\times10.4-1.05\times10}{1\times2-1.05\times2} = 1$$

8 Singular Value Decomposition

Eigen values and eigenvectors

For $A_{n\times n}$ and $X_{n\times 1}(\neq 0)$, if

$$AX = \lambda X \qquad (*)$$

then λ is called an eigenvalue of A, and X is the corresponding eigenvector.

How to find λ and X in $(*)$?

(1) AX=λX

$\Rightarrow (A - \lambda I)X = 0 \qquad \Rightarrow |A - \lambda I| = 0$

There are n roots for $|A - \lambda I| = 0$: $\lambda_1, \lambda_2, \ldots, \lambda_n$.

(2) Let X_i be the corresponding eigenvector to λ_i, then

$AX_i = \lambda_i X_i \qquad \Rightarrow (A - \lambda_i I)X_i = 0$

X_i has an infinity number of solutions. (why?)

Example:

$$A = \begin{bmatrix} 12 & 6 & -6 \\ 6 & 16 & 2 \\ -6 & 2 & 16 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 12 & 6 & -6 \\ 6 & 16 & 2 \\ -6 & 2 & 16 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

$$|A - \lambda I| = \begin{vmatrix} 12-\lambda & 6 & -6 \\ 6 & 16-\lambda & 2 \\ -6 & 2 & 16-\lambda \end{vmatrix} = -\lambda^3 + 44\lambda^2 - 564\lambda + 1728 = 0$$

$$\lambda_1 = 4.4560, \quad \lambda_2 = 18.00, \quad \lambda_3 = 21.544$$

Find eigenvector corresponding to λ_1: $(A - \lambda_1 I)X_1 = 0$

$$\begin{bmatrix} 12-4.4560 & 6 & -6 \\ 6 & 16-4.4560 & 2 \\ -6 & 2 & 16-4.4560 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 7.544 & 6 & -6 \\ 6 & 11.544 & 2 \\ -6 & 2 & 11.544 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} (2) - (1) \times \frac{6}{7.544} \\ (3) - (1) \times \frac{-6}{7.544} \end{matrix} \begin{bmatrix} 7.544 & 6 & -6 \\ 0 & 6.772 & 6.772 \\ 0 & 6.772 & 6.772 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} \\ \\ (3) - (2) \end{matrix} \begin{bmatrix} 7.544 & 6 & -6 \\ 0 & 6.772 & 6.772 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} \\ (2)/6.772 \end{matrix} \begin{bmatrix} 7.544 & 6 & -6 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{matrix} (1) - (2) \times 6 \\ \\ \end{matrix} \begin{bmatrix} 7.544 & 0 & -12 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad (1)/7.544 \begin{bmatrix} 1 & 0 & -1.59 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Let $x_{11} = 1, x_{31} = \frac{1}{1.59} = 0.6287, x_{21} = -0.6287$

$$X_1 = \begin{bmatrix} 1 \\ -0.6287 \\ 0.6287 \end{bmatrix}, \|X_1\|_2 = \sqrt{x_{11}^2 + x_{21}^2 + x_{31}^2} = 1.3381$$

Normalized eigenvector: $V_1 = \frac{X_1}{\|X_1\|_2} = [0.7473 \quad -0.4698 \quad 0.4698]'$, $\|V_1\| = 1.$

Find eigenvector corresponding to λ_2, $(A - \lambda_2 I)X_2 = 0$

$$X_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad V_2 = \frac{X_2}{\|X_2\|_2} = \begin{bmatrix} 0 \\ 0.7071 \\ 0.7071 \end{bmatrix}$$

Find eigenvector corresponding to λ_3, $(A - \lambda_3 I)X_3 = 0$

$$X_3 = \begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix} = \begin{bmatrix} 1 \\ -0.7955 \\ 0.7955 \end{bmatrix}, \quad V_3 = \frac{X_3}{\|X_3\|_2} = \begin{bmatrix} 0.6644 \\ 0.5285 \\ -0.5285 \end{bmatrix}$$

$$V = [V_1 \;\; V_2 \;\; V_3] = \begin{bmatrix} 0.7473 & 0 & 0.6644 \\ -0.4698 & 0.7071 & 0.5285 \\ 0.4698 & 0.7071 & -0.5285 \end{bmatrix}$$

$$V^{-1} = V' = \begin{bmatrix} 0.7473 & -0.4698 & 0.4698 \\ 0 & 0.7071 & 0.7071 \\ 0.6644 & 0.5282 & -0.5285 \end{bmatrix}$$

• Orthonormal:

$$< V_i, V_j >= V_i^{'} V_j = \begin{cases} V_{i1}V_{j1} + V_{i2}V_{j2} + \cdots + V_{in}V_{jn} = 0, & i \neq j \\ V_{i1}^2 + V_{i2}^2 + \cdots + V_{in}^2 = 1, & i = j \end{cases}$$

Find eigen-decomposition of A:

$$AX_i = \lambda X_i \quad \Rightarrow \quad AV_i = \lambda_i V_i, \quad \text{for } i = 1, 2, \ldots, n$$

$$\Rightarrow A[V_1 \;\; V_2 \;\; \cdots \;\; V_n] = [V_1 \;\; V_2 \;\; \cdots \;\; V_n] \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix}$$

Define $D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = \text{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then $AV = VD$, and $A = VDV^{-1} = VDV^{'}$ (why ?)

SVD

SVD is a general way for eigen-decomposition.

$$A_{m\times n} = U_{m\times m} \times S_{m\times n} \times V^{'}_{n\times n}$$

- $U_{m\times m}$ and $V_{n\times n}$ are orthonormal matrices.
- $S_{m\times n}$ is a diagonal matrix
  $S_{ij} = \begin{cases} 0, & for\ i \neq j \\ S_i, & for\ i = j \end{cases}$
  S_i is a singular values of A, $S_1 \geq S_2 \geq S_3 \cdots$
- $U = [U_1\ \ U_2\ \ \cdots\ \ U_m], \quad V = [V_1\ \ V_2\ \ \cdots\ \ V_n]$
  U_i : left singular vector corresponding to S_i
  V_i : right singular vector corresponding to S_i

How to find $S_i,\ U_i$ and V_i?

$$A_{m\times n} = U\ S\ V' \quad \Rightarrow (A')_{n\times m} = V\ S\ U'$$

$$\Rightarrow \begin{cases} (AA')_{m\times m} = U\ S\ V'\ V\ S\ U' = U\ S^2\ U', & U' = U^{-1} \\ (A'A)_{n\times n} = V\ S\ U'\ U\ S\ V' = V\ S^2\ V', & V' = V^{-1} \end{cases}$$

- The singular value of A is the square root of the eigenvalue of (AA') or $(A'A)$.
- The left singular vector of A (U_i)is the eigenvector of (AA').
- The right singular vector of A (V_i)is the eigenvector of $(A'A)$.

<u>Applications of SVD</u>

- Euclidean norm (2-norm)

$$\|A\|_2 = \max_{x\neq 0} \frac{\|AX\|_2}{\|X\|_2} = \lambda_{\max}$$

- Condition number: cond(A) = $\frac{\lambda_{\max}}{\lambda_{\min}}$

- Determinant

$$|A| = \prod_{i=1}^{n} \lambda_i, \quad A_{n\times n}$$

$$A = VDV^{-1} \quad \Rightarrow |A| = |VDV^{-1}| = |V| \cdot |D| \cdot |V^{-1}| = |D| = \lambda_1\lambda_2\cdots\lambda_n$$

- The rank of A is the number of non-zero eigenvalues.
- Approximate A to a lower rank (for image compression)

$$A_{m\times n} = U_{m\times m}\, S_{m\times n}\, V'_{n\times n}, \quad r = rank(A), \quad r \le \min(m,n)$$

$$A = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ & \cdots & & \\ u_{m1} & u_{m2} & \cdots & u_{mm} \end{bmatrix} \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ 0 & \ddots & \vdots & 0 \\ \vdots & \cdots & S_r & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix} V', \quad S_1 \ge S_2 \ge \cdots \ge S_r$$

$$= \begin{bmatrix} S_1 u_{11} & S_2 u_{12} & \cdots & S_r u_{1r} & 0 & \cdots & 0 \\ S_1 u_{21} & S_2 u_{22} & \cdots & S_r u_{2r} & 0 & \cdots & 0 \\ & \cdots & & & & & \\ S_1 u_{m1} & S_2 u_{m2} & \cdots & S_r u_{mr} & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} & \cdots & v_{n1} \\ v_{12} & v_{22} & \cdots & v_{n2} \\ & \cdots & & \\ v_{1n} & v_{2n} & \cdots & v_{nn} \end{bmatrix}$$

Then

$$\begin{aligned} A_{ij} &= \textstyle\sum_{k=1}^{r} S_k u_{ik} v_{jk} \\ A &= \textstyle\sum_{k=1}^{r} S_k U_k V_k' \end{aligned}$$

where $U_k = [u_{1k}\ u_{2k}\ \cdots\ u_{mk}]'$, $V_k' = [v_{1k}\ v_{2k}\ \cdots\ v_{nk}]$
when $r_1 \le r$,

$$A^* = \textstyle\sum_{k=1}^{r_1} S_k U_k V_k'$$

Instead of storing the $n \times n$ elements of A, S_k, U_k, and V_k, for $k = 1, 2, \ldots, r_1$

are stored, and A^* is the compressed version of A.
Compression rate:

$$\frac{r_1 + r_1 \times m + r_1 \times n}{n \times m}$$