Lecture 5: System of Equations

Abhinav Jha

Indian Institute of Technology, Gandhinagar

23^{rd} March 2026

Outline

$\underline{A}\underline{x} = \underline{b} \implies (LU)\underline{x} = \underline{b} \implies \underbrace{L\underline{y} = \underline{b} \text{ and } U\underline{x} = \underline{y}}$

L: Lower triangular
U: Upper triangular → Backward Substitution $O(n^2)$ process

Starting point: Gaussian elimination

$$[A, b] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right] \rightsquigarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & \bar{b}_1 \\ 0 & \bar{a}_{22} & \cdots & \bar{a}_{2n} & \bar{b}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \bar{a}_{nn} & \bar{b}_n \end{array}\right]$$

A $\rightsquigarrow$ U

$$A^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}$$

1st Step :

$$\rightsquigarrow A^{(2)} = \begin{bmatrix} a_{11} & \bar{a}_{12}^{(2)} & \cdots & \bar{a}_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & \bar{a}_{32}^{(2)} & \cdots & \bar{a}_{3n}^{(2)} \\ \vdots & & & \\ 0 & \bar{a}_{n2}^{(2)} & \cdots & \bar{a}_{nn}^{(2)} \end{bmatrix} \quad (*)$$

$$R_j \longmapsto R_j - m_{j1} R_1 \quad \text{where} \quad m_{j1} = \frac{a_{j1}^{(1)}}{a_{11}^{(1)}}, \; j = 2, \ldots, n$$

$$M^{(1)} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21} & 1 & 0 & \cdots & 0 \\ -m_{31} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix} \rightarrow A^{(2)} = M^{(1)} A^{(1)}$$

$$(*) = (\#)$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -m_{21} & 1 & 0 & \cdots & 0 \\ -m_{31} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -m_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}$$

$$-m_{21} a_{11}^{(1)} + a_{21}^{(1)} = -\frac{a_{21}^{(1)}}{\cancel{a_{11}^{(1)}}} \cancel{a_{11}^{(1)}} + a_{21}^{(1)} = 0$$

$$(\#) \quad = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & & & a_{2n}^{(2)} \\ 0 & \bar{a}_{32}^{(2)} & \ddots & \vdots \\ \vdots & & & \vdots \\ 0 & a_{n2}^{(2)} & & \end{bmatrix}$$

$$-m_{21} a_{12}^{(1)} + a_{22}^{(1)} = a_{22}^{(1)} - \frac{a_{21}^{(1)}}{a_{11}^{(1)}} a_{12}^{(1)}$$

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

$R_j \to R_j - m_{j2} R_2$ where $m_{j2} = \dfrac{a_{j2}^{(2)}}{a_{22}^{(2)}}$ for $j = 3, \ldots, n$

$$M^{(2)} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & -m_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -m_{n2} & 0 & \cdots & 1 \end{bmatrix}$$

$$A^{(3)} = M^{(2)} A^{(2)}$$
$$= M^{(2)} (M^{(1)} A^{(1)}) = M^{(2)} M^{(1)} A$$
$$\vdots$$
$$A^{(n)} = M^{(n-1)} M^{(n-2)} \cdots M^{(2)} M^{(1)} A$$

$$U = \underbrace{(M^{(n-1)} \cdots M^{(2)} M^{(1)})}_{\text{If Invertible}} A \Rightarrow A = LU$$

$A = LU$ where $L = (M^{(n-1)} M^{(n-2)} \dots M^{(1)})^{-1}$

Theorem: Doolittle LU

If Gaussian elimination can be performed without using row-pivoting then we can write $A = LU$ where L and U are defined as early.

Remark: If U has unit diagonal then it is Crout LU factorisation

Example:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 3 & 2 & 2 \end{bmatrix} \quad \therefore \quad L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} & l_{21}u_{13} + u_{23} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + u_{33} \end{bmatrix}$$

1st 2nd 3rd 4th

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

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

Qus:

$$A = LU$$

$$L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Imp: $\det(A) = \det(LU)$ ☆☆

$$= \det(L)\det(U)$$

$$= \det(U)$$

Theorem: If A has LU decomposition, then it is not unique.

Proof: Let $A = LU$ and D is invertible diagonal matrix

$$= L(DD^{-1})U$$

$$= (LD)(D^{-1}U)$$

$$= \tilde{L}\,\tilde{U}$$

Definition (**Diagonally Dominant Matrix**)

A matrix $\mathbf{A} = (a_{ij})$ is said to be *diagonally dominant* if

$$|a_{ii}| \geq \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}|, \qquad i = 1, 2, \ldots, n.$$

If the inequality is strict for every i, then $\mathbf{A}$ is called *strictly diagonally dominant*.

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -3 & 2 \\ 1 & 2 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

Theorem

A strictly diagonally dominant matrix $\mathbf{A}$ *is nonsingular. Moreover, Gaussian elimination applied to a system* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *can be carried out without row or column interchanges, and the computation is stable with respect to the growth of round-off errors.*

Definition (**Positive Definite Matrix**)

A matrix $\mathbf{A}$ is said to be *positive definite* if

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} > 0 \quad \text{for all } \mathbf{x} \neq \mathbf{0}.$$

If $\mathbf{A}$ is symmetric and positive definite, it is called a *symmetric positive definite (SPD)* matrix.

Example: Let $A_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$; $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \rightarrow x^T A_2 x = x_1^2 - x_2^2$

Let $\underline{x} \in \mathbb{R}^2$, $\underline{x} = (x_1, x_2)$ $\therefore \underline{x}^T A \underline{x} = [x_1 \quad x_2] \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 \quad x_2] \begin{bmatrix} 2x_1 - x_2 \\ -x_1 + 2x_2 \end{bmatrix}$

$= x_1^2 + x_2^2 - (x_1 - x_2)^2 > 0$

Definition (**Leading Principal Submatrix**)

A *leading principal submatrix* of a matrix $\mathbf{A}$ is a matrix of the form

$$\mathbf{A}_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix},$$

for some $k = 1, 2, \ldots, n$.

$$A = \begin{array}{c} \begin{array}{cccc} A_1 & A_2 & A_3 & \phantom{a_{1n}} \end{array} \\ \begin{bmatrix} \boxed{a_{11}} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \end{array}$$

Theorem (Necessary and Sufficient Condition for SPD)

A symmetric matrix $\boldsymbol{A}$ *is symmetric positive definite if and only if all its leading principal submatrices have positive determinants.*

Example: $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

$A_1 = [2]$ $\therefore \det(A_1) > 0$

$A_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$, $\det(A_2) = 3 > 0$

$A_3 = A$; $\det(A_3) = 4 > 0 \Rightarrow A$ is spd.

Theorem

A symmetric matrix $\mathbf{A}$ is symmetric positive definite if and only if Gaussian elimination without row interchanges can be applied to the linear system $\mathbf{Ax} = \mathbf{b}$ with all pivot elements positive. Moreover, in this case, the computations are stable with respect to the growth of round-off errors.

Corollary (LDL$^\top$ Factorisation)

A matrix $\mathbf{A}$ is symmetric positive definite if and only if it admits a factorisation of the form $\mathbf{A} = \mathbf{LDL}^\top$, where $\mathbf{L}$ is a unit lower triangular matrix and $\mathbf{D}$ is a diagonal matrix with positive diagonal entries.

Corollary

Let $\mathbf{A}$ be a symmetric matrix for which Gaussian elimination can be applied without row interchanges. Then $\mathbf{A}$ admits a factorisation of the form $\mathbf{A} = \mathbf{L}\mathbf{D}\mathbf{L}^\top$, where $\mathbf{L}$ is a unit lower triangular matrix and $\mathbf{D}$ is a diagonal matrix whose diagonal entries are the pivots $a_{11}^{(1)}, a_{22}^{(2)}, \ldots, a_{nn}^{(n)}$.

Example

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}; \quad L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \quad D = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix}$$

$$LDL^T = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} 1 & l_{21} & l_{31} \\ 0 & 1 & l_{32} \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} d_{11} & d_{11} l_{21} & d_{11} l_{31} \\ d_{11} l_{21} & d_{22} + d_{11} l_{21}^2 & d_{22} l_{32} + d_{11} l_{21} l_{31} \\ d_{11} l_{31} & d_{22} l_{32} + d_{11} l_{21} l_{31} & d_{11} l_{31}^2 + d_{22} l_{32}^2 + d_{33} \end{bmatrix}$$

Corollary (Cholesky Decomposition)

A matrix $\boldsymbol{A}$ is symmetric positive definite if and only if it can be factored in the form

$$\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top,$$

where $\boldsymbol{L}$ is a lower triangular matrix with strictly positive diagonal entries.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} ; L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$$

$$A = LL^T$$

$$= \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

$$= \begin{bmatrix} l_{11}^2 & l_{11} l_{21} & l_{11} \cdot l_{31} \\ \times & l_{21}^2 + l_{22}^2 & l_{21} l_{31} + l_{22} l_{32} \\ \times & \times & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}$$

Computational Complexity

Method	Applicability	Advantages	Computational M/D	Computational A/S
LU	General square matrix	Works for anything but needs pivoting	$\frac{n^3}{3} - \frac{n}{3}$	$\frac{n^3}{3} - \frac{n^2}{2} + \frac{n}{6}$
LDL^T	Symmetric matrix	More stable than LU, reduce storage, No pivot	$\frac{n^3}{6} + n^2 - \frac{7n}{6}$	$\frac{n^3}{6} - \frac{n}{6}$
Cholesky	SPD	Fastest and most efficient	$\frac{n^3}{6} + \frac{n^2}{2} - \frac{2n}{3}$	$\frac{n^3}{6} - \frac{n}{6}$

Definition

A *vector norm* on $\mathbb{R}^n$ is a function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ satisfying the following properties:

1. $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
2. $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
3. $\|\alpha\mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$ for all $\alpha \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.
4. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Definition

The ℓ_2 and ℓ_∞ norms of a vector $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ are defined as

$$\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}, \qquad \|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Definition

A sequence $\{\mathbf{x}^{(k)}\}_{k=1}^{\infty}$ in $\mathbb{R}^n$ is said to converge to $\mathbf{x}$ with respect to a norm $\|\cdot\|$ if for every $\varepsilon > 0$ there exists an integer $N(\varepsilon)$ such that

$$\|\mathbf{x}^{(k)} - \mathbf{x}\| < \varepsilon \quad \text{for all } k \ge N(\varepsilon).$$

Theorem

If $\|\cdot\|$ is a vector norm on $\mathbb{R}^n$, then

$$\|\mathbf{A}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|$$

defines a matrix norm.

Corollary

For any vector $\mathbf{y} \neq \mathbf{0}$, any matrix $\mathbf{A}$, and any induced matrix norm $\|\cdot\|$, we have

$$\|\mathbf{A}\mathbf{y}\| \leq \|\mathbf{A}\| \, \|\mathbf{y}\|.$$

Iterative Methods