# Ch06-3 Linear Systems of Equations, Matrix Factorization

Dr. Feras Fraige

# Outline

- 1 Computation Cost Rationale & Basic Solution Strategy
- 2 Constructing the Matrix Factorization
- 3 Example: *LU Factorization of a 4 × 4 Matrix*
- 4 The LU Factorization Algorithm
- 5 Permutation Matrices for Row Interchanges

#### Background

- Gaussian elimination is the principal tool in the direct solution of linear systems of equations.
- We will now see that the steps used to solve a system of the form
   Ax = b can be used to factor a matrix.
- The factorization is particularly useful when it has the form A = LU, where L is lower triangular and U is upper triangular.
- Although not all matrices have this type of representation, many do that occur frequently in the application of numerical techniques.

#### **Computational Cost Considerations**

- Gaussian elimination applied to an arbitrary linear system Ax = b requires  $O(n^3/3)$  arithmetic operations to determine x.
- However, to solve a linear system that involves an upper-triangular system requires only backward substitution, which takes O(n²) operations.
- The number of operations required to solve a lower-triangular systems is similar.

#### Solution Strategy

Suppose that A has been factored into the triangular form A = LU, where L is lower triangular and U is upper triangular. Then we can solve for **x** more easily by using a two-step process:

- First we let  $\mathbf{y} = U\mathbf{x}$  and solve the lower triangular system  $L\mathbf{y} = \mathbf{b}$  for  $\mathbf{y}$ . Since L is triangular, determining  $\mathbf{y}$  from this equation requires only  $O(n^2)$  operations.
- Once **y** is known, the upper triangular system U**x** = **y** requires only an additional  $O(n^2)$  operations to determine the solution **x**.

Solving a linear system  $A\mathbf{x} = \mathbf{b}$  in factored form means that the number of operations needed to solve the system  $A\mathbf{x} = \mathbf{b}$  is reduced from  $O(n^3/3)$  to  $O(2n^2)$ .

#### Constructing L & U

- First, suppose that Gaussian elimination can be performed on the system Ax = b without row interchanges.
- With the notation used earlier, this is equivalent to having nonzero pivot elements  $a_{ii}^{(i)}$ , for each i = 1, 2, ..., n.
- The first step in the Gaussian elimination process consists of performing, for each j = 2, 3, ..., n, the operations

$$(E_j - m_{j,1}E_1) \rightarrow (E_j), \text{ where } m_{j,1} = \frac{a_{j1}^{(1)}}{a_{11}^{(1)}}$$

 These operations transform the system into one in which all the entries in the first column below the diagonal are zero.

# Matrix Factorization: Constructing L & U (Cont'd)

The system of operations in

$$(E_j - m_{j,1}E_1) \rightarrow (E_j), \text{ where } m_{j,1} = \frac{a_{j1}^{(1)}}{a_{11}^{(1)}}$$

can be viewed in another way. It is simultaneously accomplished by multiplying the original matrix A on the left by the matrix

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

This is called the first Gaussian transformation matrix.

#### Constructing L & U (Cont'd)

We denote the product of this matrix with A<sup>(1)</sup> ≡ A by A<sup>(2)</sup> and with b by b<sup>(2)</sup>, so

$$A^{(2)}\mathbf{x} = M^{(1)}A\mathbf{x} = M^{(1)}\mathbf{b} = \mathbf{b}^{(2)}$$

 In a similar manner we construct M<sup>(2)</sup>, the identity matrix with the entries below the diagonal in the second column replaced by the negatives of the multipliers

$$m_{j,2} = \frac{a_{j2}^{(2)}}{a_{22}^{(2)}}.$$

#### Constructing *L* & *U* (Cont'd)

• The product of  $M^{(2)}$  with  $A^{(2)}$  has zeros below the diagonal in the first two columns, and we let

$$A^{(3)}\mathbf{x} = M^{(2)}A^{(2)}\mathbf{x} = M^{(2)}M^{(1)}A\mathbf{x} = M^{(2)}M^{(1)}\mathbf{b} = \mathbf{b}^{(3)}$$

#### Constructing L & U (Cont'd)

In general, with  $A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$  already formed, multiply by the **k**th Gaussian transformation matrix

$$M^{(k)} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & 0 & \ddots & \ddots & & \vdots \\ \vdots & \vdots & -m_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \cdots & 0 & -m_{n,k} & 0 & \cdots & 0 & 1 \end{bmatrix}$$

#### Constructing *L* & *U* (Cont'd)

to obtain

$$A^{(k+1)}\mathbf{x} = M^{(k)}A^{(k)}\mathbf{x}$$

$$= M^{(k)}\cdots M^{(1)}A\mathbf{x}$$

$$= M^{(k)}\mathbf{b}^{(k)}$$

$$= \mathbf{b}^{(k+1)}$$

$$= M^{(k)}\cdots M^{(1)}\mathbf{b}$$

#### Constructing L & U (Cont'd)

The process ends with the formation of  $A^{(n)}\mathbf{x} = \mathbf{b}^{(n)}$ , where  $A^{(n)}$  is the upper triangular matrix

$$A^{(n)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n}^{(n-1)} \\ 0 & \cdots & \cdots & 0 & a_{n,n}^{(n)} \end{bmatrix}$$

given by

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

#### Constructing *L* & *U* (Cont'd)

- This process forms the  $U = A^{(n)}$  portion of the matrix factorization A = LU.
- To determine the complementary lower triangular matrix L, first recall the multiplication of  $A^{(k)}\mathbf{x} = \mathbf{b}^{(k)}$  by the Gaussian transformation of  $M^{(k)}$  used to obtain:

$$A^{(k+1)}\mathbf{x} = M^{(k)}A^{(k)}\mathbf{x} = M^{(k)}\mathbf{b}^{(k)} = \mathbf{b}^{(k+1)},$$

where  $M^{(k)}$  generates the row operations

$$(E_j - m_{j,k}E_k) \rightarrow (E_j), \quad \text{for } j = k+1,\ldots,n.$$

#### Constructing L & U (Cont'd)

To reverse the effects of this transformation and return to  $A^{(k)}$  requires that the operations  $(E_j + m_{j,k}E_k) \rightarrow (E_j)$  be performed for each j = k + 1, ..., n. This is equivalent to multiplying by  $[M^{(k)}]^{-1}$ :

$$L^{(k)} = \left[ M^{(k)} \right]^{-1} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & m_{k+1,k} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \cdots & 0 & m_{n,k} & 0 & \cdots & 0 & 1 \end{bmatrix}$$

#### Constructing L & U (Cont'd)

The lower-triangular matrix L in the factorization of A, then, is the product of the matrices  $L^{(k)}$ :

$$L = L^{(1)}L^{(2)}\cdots L^{(n-1)} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ m_{21} & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ m_{n1} & \cdots & \cdots & m_{n,n-1} & 1 \end{bmatrix}$$

since the product of L with the upper-triangular matrix  $U = M^{(n-1)} \cdots M^{(2)} M^{(1)} A$  gives

#### Constructing L & U (Cont'd)

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

$$= [M^{(1)}]^{-1}[M^{(2)}]^{-1} \cdots [M^{(n-2)}]^{-1}[M^{(n-1)}]^{-1} \\ \cdot M^{(n-1)}M^{(n-2)} \cdots M^{(2)}M^{(1)} A$$

$$= A$$

We now state a theorem which follows from these observations.

#### Theorem

If Gaussian elimination can be performed on the linear system  $A\mathbf{x} = \mathbf{b}$  without row interchanges, then the matrix A can be factored into the product of a lower-triangular matrix L and an upper-triangular matrix U, that is, A = LU, where  $m_{ji} = a_{ji}^{(i)}/a_{ji}^{(i)}$ ,

$$U = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n}^{(n-1)} \\ 0 & \cdots & 0 & a_{n,n}^{(n)} \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ m_{21} & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ m_{n1} & \cdots & \cdots & m_{n,n-1} & 1 \end{bmatrix}$$

#### Example

(a) Determine the LU factorization for matrix A in the linear system  $A\mathbf{x} = \mathbf{b}$ , where

$$A = \begin{bmatrix} 1 & 1 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 3 & -1 & -1 & 2 \\ -1 & 2 & 3 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ -3 \\ 4 \end{bmatrix}$$

(b) Then use the factorization to solve the system

$$x_1 + x_2 + 3x_4 = 8$$
  
 $2x_1 + x_2 - x_3 + x_4 = 7$   
 $3x_1 - x_2 - x_3 + 2x_4 = 14$   
 $-x_1 + 2x_2 + 3x_3 - x_4 = -7$ 

#### Part (a) Solution (1/2)

The original system was considered under Gaussian Elimination where we saw that the sequence of operations

$$(E_2-2E_1) o (E_2) \hspace{1cm} (E_3-3E_1) o (E_3) \ (E_4-(-1)E_1) o (E_4) \hspace{1cm} (E_3-4E_2) o (E_3) \ (E_4-(-3)E_2) o (E_4)$$

converts the system to the triangular system

$$x_1 + x_2 + 3x_4 = 4$$
 $-x_2 - x_3 - 5x_4 = -7$ 
 $3x_3 + 13x_4 = 13$ 
 $-13x_4 = -13$ 

#### Part (a) Solution (2/2)

The multipliers  $m_{ij}$  and the upper triangular matrix produce the factorization

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

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix}$$

$$= LU$$

#### Part (b) Solution (1/3)

To solve

$$A\mathbf{x} = LU\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$
$$= \begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix}$$

we first introduce the substitution  $\mathbf{y} = U\mathbf{x}$ . Then  $\mathbf{b} = L(U\mathbf{x}) = L\mathbf{y}$ .

#### Part (b) Solution (2/3)

First, solve  $L\mathbf{y} = \mathbf{b}$  (where  $\mathbf{y} = U\mathbf{x}$ :

$$L\mathbf{y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 14 \\ -7 \end{bmatrix}.$$

This system is solved for **y** by a simple forward-substitution process:

$$y_1 = 8$$
  
 $2y_1 + y_2 = 7 \Rightarrow y_2 = 7 - 2y_1 = -9$   
 $3y_1 + 4y_2 + y_3 = 14 \Rightarrow y_3 = 14 - 3y_1 - 4y_2 = 26$   
 $-y_1 - 3y_2 + y_4 = -7 \Rightarrow y_4 = -7 + y_1 + 3y_2 = -26$ 

#### Part (b) Solution (3/3)

We then solve  $U\mathbf{x} = \mathbf{y}$  for  $\mathbf{x}$ , the solution of the original system; that is,

$$\begin{bmatrix} 1 & 1 & 0 & 3 \\ 0 & -1 & -1 & -5 \\ 0 & 0 & 3 & 13 \\ 0 & 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8 \\ -9 \\ 26 \\ -26 \end{bmatrix}$$

Using backward substitution we obtain  $x_4 = 2$ ,  $x_3 = 0$ ,  $x_2 = -1$ ,  $x_1 = 3$ .

# Discuss algorithm

#### Using the LU Factorization to solve $A\mathbf{x} = \mathbf{b}$

Once the matrix factorization is complete, the solution to a linear system of the form

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

is found by first letting

$$\mathbf{y} = U\mathbf{x}$$

and solving

$$Ly = b$$

for y.

#### Using the *LU* Factorization (Cont'd)

• Since *L* is lower triangular, we have  $y_1 = \frac{b_1}{l_{11}}$  and, for each i = 2, 3, ..., n,

$$y_i = \frac{1}{I_{ii}} \left[ b_i - \sum_{j=1}^{i-1} I_{ij} y_j \right]$$

 After y is found by this forward-substitution process, the upper-triangular system Ux = y is solved for x by backward substitution using the equations

$$x_n = \frac{y_n}{u_{nn}}$$
 and  $x_i = \frac{1}{u_{ii}} \left[ y_i - \sum_{j=i+1}^n u_{ij} x_j \right]$ 

#### Limitations of the *LU* Factorization Algorithm

- We assumed that Ax = b can be solved using Gaussian elimination without row interchanges.
- From a practical standpoint, this factorization is useful only when row interchanges are not required to control round-off error.
- We will now consider the modifications that must be made when row interchanges are required.

We begin with the introduction of a class of matrices that are used to rearrange, or permute, rows of a given matrix.

#### Permutation Matrix

An  $n \times n$  permutation matrix  $P = [p_{ij}]$  is a matrix obtained by rearranging the rows of  $I_n$ , the identity matrix. This gives a matrix with precisely one nonzero entry in each row and in each column, and each nonzero entry is a 1.

#### Example

The matrix

$$P = \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right]$$

is a  $3 \times 3$  permutation matrix. For any  $3 \times 3$  matrix A, multiplying on the left by P has the effect of interchanging the second and third rows of A:

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Similarly, multiplying A on the right by P interchanges the second and third columns of A.

#### Two useful properties of permutation matrices (1/2)

Suppose  $k_1, \ldots, k_n$  is a permutation of the integers  $1, \ldots, n$  and the permutation matrix  $P = (p_{ij})$  is defined by

$$p_{ij} = \begin{cases} 1, & \text{if } j = k_i \\ 0, & \text{otherwise} \end{cases}$$

#### Two useful properties of permutation matrices (2/2)

#### Then

PA permutes the rows of A; that is,

$$PA = \begin{bmatrix} a_{k_11} & a_{k_12} & \cdots & a_{k_1n} \\ a_{k_21} & a_{k_22} & \cdots & a_{k_2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k_n1} & a_{k_n2} & \cdots & a_{k_nn} \end{bmatrix}$$

•  $P^{-1}$  exists and  $P^{-1} = P^t$ .

#### Permutation Matrices & Gaussian Elimination

- Earlier, we saw that for any nonsingular matrix A, the linear system Ax = b can be solved by Gaussian elimination, with the possibility of row interchanges.
- If we knew the row interchanges that were required to solve the system by Gaussian elimination, we could arrange the original equations in an order that would ensure that no row interchanges are needed.
- Hence there is a rearrangement of the equations in the system that permits Gaussian elimination to proceed without row interchanges.

#### Permutation Matrices & Gaussian Elimination (Cont'd)

 This implies that for any nonsingular matrix A, a permutation matrix P exists for which the system

$$PAx = Pb$$

can be solved without row interchanges. As a consequence, this matrix PA can be factored into PA = LU, where L is lower triangular and U is upper triangular.

• Because  $P^{-1} = P^t$ , this produces the factorization

$$A = P^{-1}LU = (P^tL)U.$$

The matrix U is still upper triangular, but P<sup>t</sup>L is not lower triangular unless P = I.

#### Example

Determine a factorization in the form  $A = (P^t L)U$  for the matrix

$$A = \left[ \begin{array}{rrrr} 0 & 0 & -1 & 1 \\ 1 & 1 & -1 & 2 \\ -1 & -1 & 2 & 0 \\ 1 & 2 & 0 & 2 \end{array} \right]$$

#### Note

The matrix A cannot have an LU factorization because  $a_{11} = 0$ .

#### Solution (1/4)

However, using the row interchange  $(E_1) \leftrightarrow (E_2)$ , followed by  $(E_3 + E_1) \rightarrow (E_3)$  and  $(E_4 - E_1) \rightarrow (E_4)$ , produces

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

Then, the row interchange  $(E_2) \leftrightarrow (E_4)$ , followed by  $(E_4 + E_3) \rightarrow (E_4)$ , gives the matrix

$$U = \left[ \begin{array}{cccc} 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{array} \right]$$

#### Solution (2/4)

The permutation matrix associated with the row interchanges  $(E_1) \leftrightarrow (E_2)$  and  $(E_2) \leftrightarrow (E_4)$  is

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

and

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

#### Solution (3/4)

- Gaussian elimination is performed on PA using the same operations as on A, except without the row interchanges.
- That is,  $(E_2 E_1) \rightarrow (E_2)$ ,  $(E_3 + E_1) \rightarrow (E_3)$ , followed by  $(E_4 + E_3) \rightarrow (E_4)$ .
- The nonzero multipliers for PA are consequently,

$$m_{21} = 1$$
,  $m_{31} = -1$ , and  $m_{43} = -1$ ,

and the LU factorization of PA is

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

#### Solution (4/4)

Multiplying by  $P^{-1} = P^t$  produces the factorization

$$A = P^{-1}(LU) = P^{t}(LU) = (P^{t}L)U$$

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