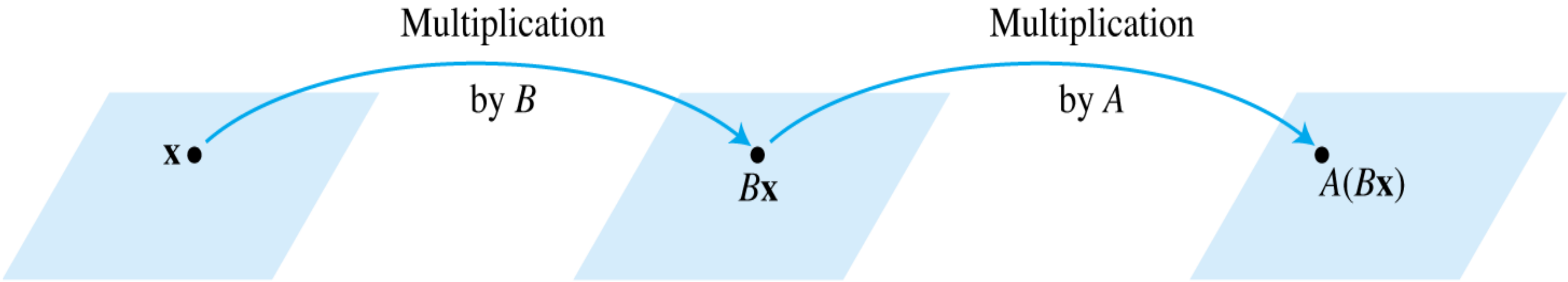
Matrix Algebra

Overview and Learning Outcomes

- Inverse of a matrix
  - Apply properties of matrix inverse
  - Write the elementary matrix corresponding to an ERO
  - Find the inverse of a 3×3 matrix using EROs
  - Prove the invertible matrix theorem
- Matrix factorization
  - Perform LU factorization with and without permutation

2.1 Matrix Multiplication

Previous chapter: Matrix as a tranformation $A\mathbf{x} = \mathbf{b}$

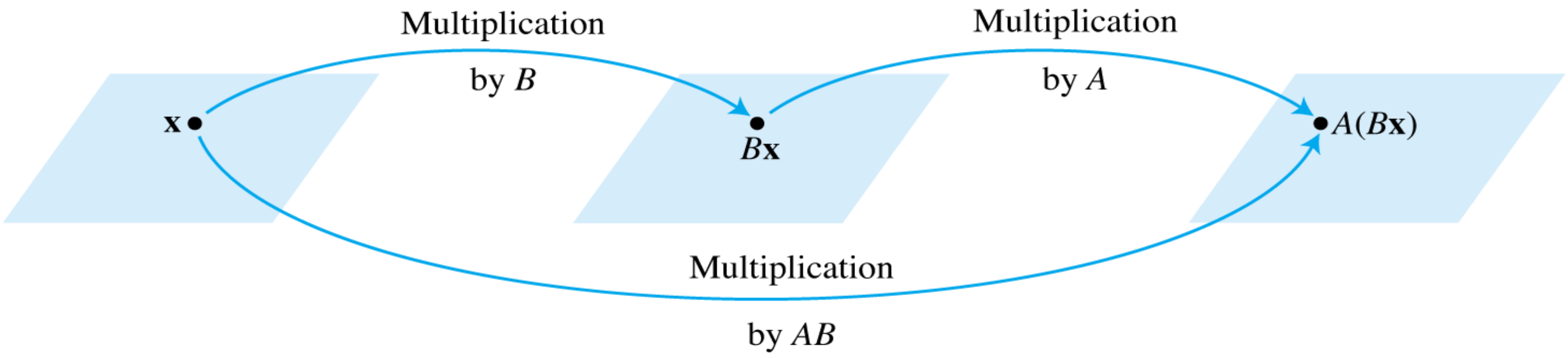


Multiplication by B and then A.

$A(B\mathbf{x})$ is produced from $\mathbf{x}$ by a *composition* of two linear trasnformations

Represent the two transformations as multiplication by a single matrix

Represent the two transformations as multiplication by a single matrix AB



Multiplication by AB.

$$A(B\mathbf{x}) = (AB)\mathbf{x}$$

2.2 Inverse of a Matrix

Theorem 2.1. *If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.*

Proof. Let $\mathbf{b} \in \mathbb{R}^n$.

Solution *exists* : Substitute $A^{-1}\mathbf{b}$ in $A\mathbf{x} = \mathbf{b}$.
LHS $= A\mathbf{x} = A(A^{-1})\mathbf{b} = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b} =$ RHS.

Solution is *unique* : Show that if $\mathbf{u}$ is a solution, it must be $A^{-1}\mathbf{b}$.
If $A\mathbf{u} = \mathbf{b}$, multiply both sides by A^{-1}
$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}$ or $I\mathbf{u} = A^{-1}\mathbf{b}$, i.e., $\mathbf{u} = A^{-1}\mathbf{b}$

$\square$

Theorem 2.2. *Invertible matrices*

1. *If A is an invertible matrix, then A^{-1} is invertible and $(A^{-1})^{-1} = A$*
2. *If A and B are $n \times n$ invertible matrices, then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$*
3. *If A is an invertible matrix, then so is A^T, and $(A^T)^{-1} = (A^{-1})^T$*

Proof

1. Find a matrix C such that $A^{-1}C = I$ and $CA^{-1} = I$. Here, C is simply A. Hence, A^{-1} is invertible and its inverse is A.
2. Find a matrix C such that $(AB)C = I$ and $C(AB) = I$.
If $C = B^{-1}A^{-1}$, then $AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. Similarly show that $(B^{-1}A^{-1})(AB) = I$.
3.

$\blacksquare$

Definition. An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.

Exercise 2.2.1:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \overset{\displaystyle\downarrow E_1}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}} \qquad r3 \leftarrow r3 - 4r1$$

$$E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

$$E_1 A = \begin{bmatrix} a & b & c \\ d & e & f \\ g - 4a & h - 4b & i - 4c \end{bmatrix}$$

$E_2 A = ?$ $\quad E_3 A = ?$

If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

Exercise 2.2.2:

Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

To transform E_1 to I, add $+4$ times row 1 to row 3. The elementary matrix that does this is

$$E_1{}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}.$$

$E_2{}^{-1} =?$ $\quad E_3{}^{-1} =?$

Theorem 2.3. *An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n to A^{-1}.*

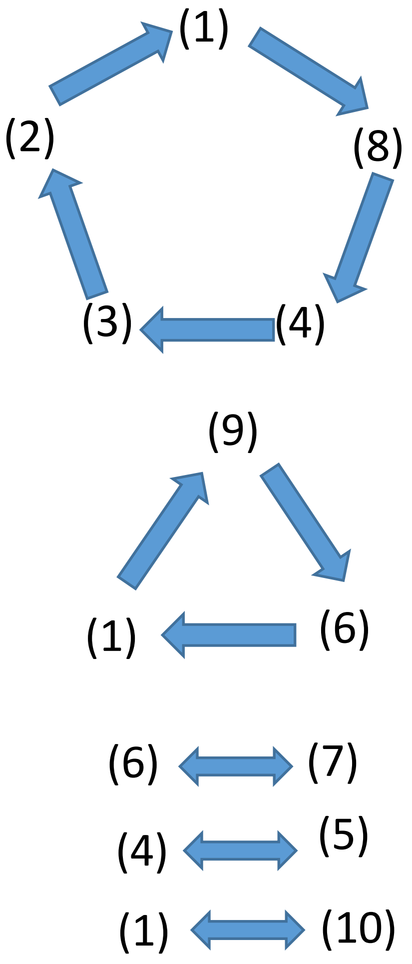
Proof. $\square$

Exercise 2.2.3:

Find the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$.

Theorem 2.4. *The Invertible Matrix Theorem*
Let A be a square $n \times n$ matrix. Then the following statements are equivalent, i.e., for a given A, the statements are either all true or all false.

1. *A is an invertible matrix.*
2. *A is row equivalent to I_n.*
3. *A has n pivot positions.*
4. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
5. *The columns of A form a linearly independent set.*
6. *$A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.*
7. *The columns of A span $\mathbb{R}^n$.*
8. *There is an $n \times n$ matrix C such that $CA=I$.*
9. *There is an $n \times n$ matrix D such that $AD=I$.*
10. *A^T is an invertible matrix.*



To help prove (6) $\Rightarrow$ (1), recall Theorem 1.2 from Chapter 1, slide 27.

Theorem 1.2. *Let A be an $m \times n$ matrix. Then the following statements are logically equivalent, i.e., for a particular A, either they are all true statements or they are all false.*

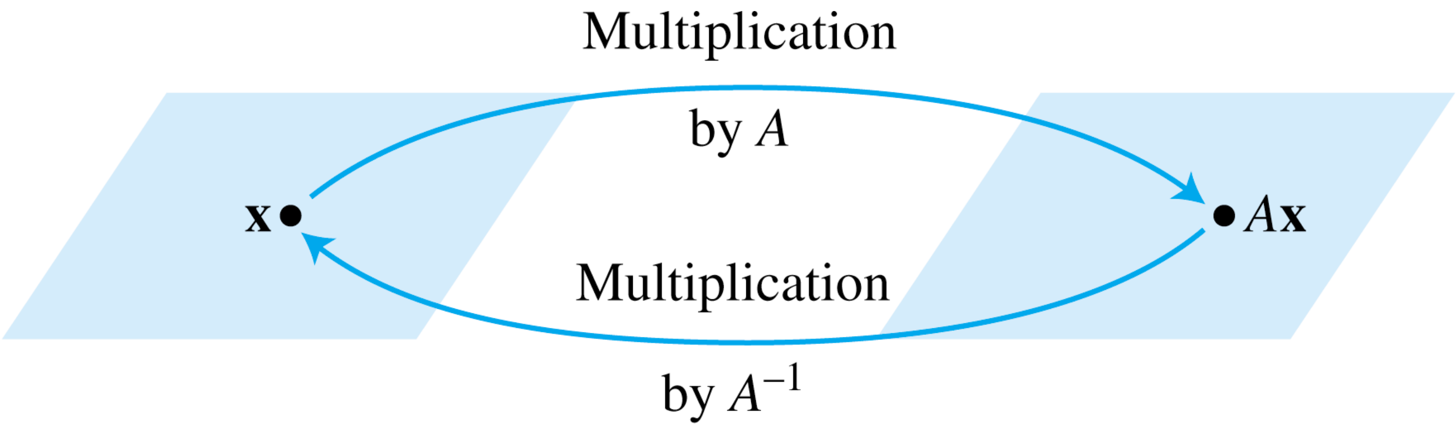
a. *For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*

b. *Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.*

c. *The columns of A span $\mathbb{R}^m$.*

d. *A has a pivot position in every row.*

Invertible Linear Transformations



A^{-1} transforms $A\mathbf{x}$ back to $\mathbf{x}$.

2.3 Matrix Factorizations

- Matrix multiplication $\Rightarrow$ *synthesis* of data
- A expressed as a product of two or more matrices $\Rightarrow$ *analysis* of data

$$\underbrace{X(:,j)}_{j\text{th facial image}} \approx \sum_{k=1}^{r} \underbrace{W(:,k)}_{\text{facial features}} \underbrace{H(k,j)}_{\substack{\text{importance of features}\\ \text{in } j\text{th image}}} = \underbrace{WH(:,j)}_{\substack{\text{approximation}\\ \text{of } j\text{th image}}} .$$

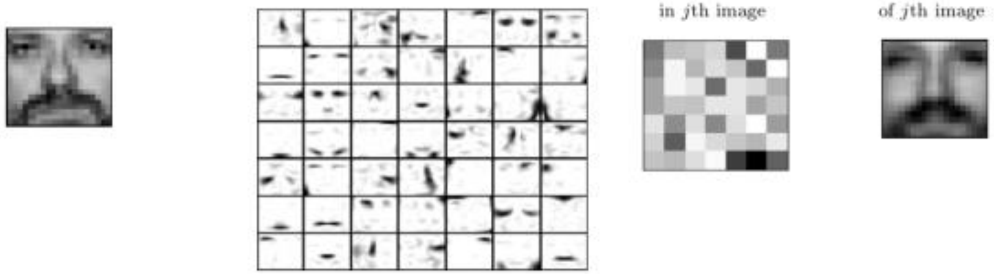


Figure 1: Decomposition of the CBCL face database, MIT Center For Biological and Computation Learning (2429 gray-level 19-by-19 pixels images) using $r = 49$ as in [79].

2.3.1 The LU factorization

- Why?

  Consider solving a sequence of equations $A\mathbf{x} = \mathbf{b_1}, A\mathbf{x} = \mathbf{b_2}, \ldots, A\mathbf{x} = \mathbf{b_p}$

  Inefficient solution: Compute A^{-1} and then $A^{-1}\mathbf{b_1}, \ldots, A^{-1}\mathbf{b_p}$

  Efficient solution: $A_{m\times n} = L_{m\times m}U_{m\times n}$

  Assumption - A can be reduced to echelon form *without row interchanges*

  L: Unit Lower triangular

  U: Upper triangular

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{\substack{U \\ \text{Echelon form}}}$$

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

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

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

‘Forward Substitution’

$L\mathbf{y} = \mathbf{b} \rightarrow$ Solve for $\mathbf{y}$

$U\mathbf{x} = \mathbf{y} \rightarrow$ Solve for $\mathbf{x}$

Easy to solve because L and U are triangular

‘Backward Substitution’

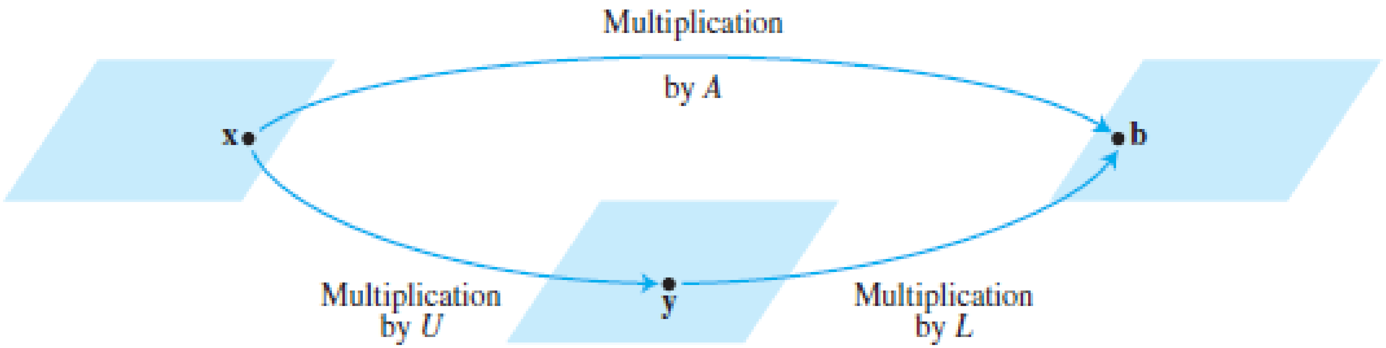


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

Exercise 2.3.1

Solve $A\mathbf{x} = \mathbf{b}$ if $A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix}}_{\mathrm{L}} \underbrace{\begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{\mathrm{U}}$

and $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

$$L\mathbf{y} = \mathbf{b} : [L \quad \mathbf{b}] = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = [I \quad \mathbf{y}]$$

Number of multiplication - addition pairs to reduce L to I } 3 2 1

6 multiplications
6 additions

$$y_1 = -9$$
$$-y_1 + y_2 = 5$$
$$2y_1 - 5y_2 + y_3 = 7$$
$$-3y_1 + 8y_2 + 3y_3 + y_4 = 11$$

$$U\mathbf{x} = \mathbf{y} : \begin{bmatrix} U & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{x} \end{bmatrix}$$

To reduce U to I:
Number of divisions - 4
Number of additions - 6
Number of multiplications - 6

$$\begin{aligned} 3x_1 - 7x_2 - 2x_3 + 2x_4 &= -9 \\ -2x_2 - x_3 + 2x_4 &= -4 \\ -x_3 + x_4 &= 5 \\ -x_4 &= 1 \end{aligned}$$

Through LU factorization : 28 arithmetic operations or "flops" (floating point operations) - excluding cost of factorization
Through row reduction of $\begin{bmatrix} A & \mathbf{b} \end{bmatrix}$ to $\begin{bmatrix} I & \mathbf{x} \end{bmatrix}$: 62 flops

$$\mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}.$$

2.3.2 LU factorization procedure

- Row reduction of A to U produces L without extra work

RECALL: *Assumption* - A can be reduced to echelon form *without row interchanges*

There exist unit lower triangular elementary matrices $E_1, \ldots, E_p$ such that

$$E_p \ldots E_1 A = U$$

$$\Rightarrow A = (E_p \ldots E_1)^{-1} U = LU$$

$$\text{where } L = (E_p \ldots E_1)^{-1}$$

[Products and inverses of unit lower triangular matrices are also unit lower triangular]

Same row operations that reduce A to U also reduce L to I

$$E_p \ldots E_1 L = I$$

Exercise 2.3.2:

Find an LU factorization of $A = \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix}$.

$$E_{21}A = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 4 & -8 & 30 \end{bmatrix}$$

$$E_{31}(E_{21}A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 4 & -8 & 30 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & -4 & 24 \end{bmatrix}$$

$$E_{32}(E_{31}E_{21}A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & -4 & 24 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 4 \end{bmatrix} = U.$$

Since A has 3 rows, L should be 3×3

$$L = \begin{bmatrix} 1 & 0 & 0 \\ ? & 1 & 0 \\ ? & ? & 1 \end{bmatrix}$$

The row operations that create zeros in each column of A will also create zeros in each column of L.

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & -4 & 24 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 4 \end{bmatrix} = U.$$

Circled entries are used to determine the sequence of transformations that transform A to U. At each pivot column, divide the encircled entries by the pivot (first element inside the circle) and place the result into L.

$$L = \begin{bmatrix} 2 & & \\ 6 & -1 & \\ 4 & -4 & 4 \end{bmatrix} \rightarrow L = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 4 & 1 \end{bmatrix}$$

- Row reduction of A to U produces L without extra work

Alternately,

$$L = E_{21}^{-1}E_{31}^{-1}E_{32}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 4 & 1 \end{bmatrix}$$

Just put the nonzero off-diagonal elements of the elementary matrices into the appropriate positions in L.

Exercise 2.3.3 (when below assumption is not valid)

(*Assumption* - A can be reduced to echelon form *without row interchanges*)

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

To switch rows 2 and 3, use **permutation matrix** $P = P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

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

In summary,
For every $n \times n$ matrix A there exists a permutation matrix P, such that PA possesses an LU-factorization, i.e., $PA = LU$, where L is a lower triangular matrix with all diagonal entries equal to 1, and U is an upper triangular matrix.

For an $n \times n$ dense matrix and for n moderately large, say $n \geq 30$,

LU factorization : about $2n^3/3$ flops
Finding A^{-1} : about $2n^3$ flops
Solving $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$: $2n^2$ flops
Multiplication of $\mathbf{b}$ by A^{-1} : about $2n^2$ flops

***** END OF CHAPTER *****