Linear Algebra

Chapter 3: Matrices

University of Seoul
School of Computer Science
Minho Kim

Table of contents

Outline

Matrices in Action

- Matrices as **functions on vectors**.
- Matrices **transform** a vector into another vector. (Problem 1)
- Matrices transform a parallelogram to another one. (Problem 2)
- What happens if we apply successive transformations? (Problem 4)
- Can we concatenate two successive transformations? Is it commutative? (Problem 5-7)

Outline

Matrices

Definition

A **matrix** is a rectangular array of numbers called the **entries**, or **elements**, of the matrix.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = [a_{ij}] = [a_{ij}]_{m \times n} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_n] = \begin{bmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_m \end{bmatrix}$$

where

$$\boldsymbol{u}_j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad \text{and} \quad \boldsymbol{v}_i = [a_{i1} \cdots a_{in}]$$

A matrix can be considered as

- "a row vector of column vectors" or
- "a column vector of row vectors"

Special Matrices

- **Square matrix**

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

- **Diagonal matrix**

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

- **Scalar matrix**

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

- **Identity matrix**

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

Two matrices are **equal** if

- they have the same size *and*
- their corresponding entries are equal.

Matrix Operations

- Addition

$$A + B = [a_{ij} + b_{ij}]$$

- Scalar multiplication

$$cA = c[a_{ij}] = [ca_{ij}]$$

- Difference

$$A - B = A + (-B)$$

Matrix Multiplication

Definition

If A is an $m \times n$ matrix and B is an $n \times r$ matrix, then the **product** $C = AB$ is an $m \times r$ matrix. The (i, j) entry of the product is computed as follows:

$$c_{ij} = a_{i1}b_{1j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

- The (i, j) entry is the dot product of the ith row vector of A and the jth column vector of B.

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_i \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_m \end{bmatrix} \left[\begin{array}{c:c:c:c:c} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_j & \cdots & \boldsymbol{b}_r \end{array} \right] = \begin{bmatrix} \boldsymbol{a}_1 \cdot \boldsymbol{b}_1 & \cdots & \boldsymbol{a}_1 \cdot \boldsymbol{b}_j & \cdots & \boldsymbol{a}_1 \cdot \boldsymbol{b}_r \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_i \cdot \boldsymbol{b}_1 & \cdots & \boldsymbol{a}_i \cdot \boldsymbol{b}_j & \cdots & \boldsymbol{a}_i \cdot \boldsymbol{b}_r \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{a}_m \cdot \boldsymbol{b}_1 & \cdots & \boldsymbol{a}_m \cdot \boldsymbol{b}_j & \cdots & \boldsymbol{a}_m \cdot \boldsymbol{b}_r \end{bmatrix}$$

Matrices and Linear Systems

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & + & 3x_3 & = & 5 \\ -x_1 & + & 3x_2 & + & x_3 & = & 1 \\ 2x_1 & - & x_2 & + & 4x_3 & = & 14 \end{array} \rightarrow \begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 14 \end{bmatrix}$$

If we consider the matrix as a row vector of column vectors,

$$\left[\begin{array}{r:r:r} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_1 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}$$

Picking Columns or Rows

Theorem 3.1

Let A be an $m \times n$ matrix, $\boldsymbol{e}_i$ a $1 \times m$ standard unitvector, and $\boldsymbol{e}_j$ an $n \times 1$ standard unitvector. Then

a. $\boldsymbol{e}_i A$ is the ith row of A and

b. $A\boldsymbol{e}_j$ is the jth column of A.

$$
\begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix}
\begin{bmatrix} \boldsymbol{a}_1 \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_i \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_m \end{bmatrix} = \boldsymbol{a}_i
$$

$$
\left[\begin{array}{c:c:c:c:c} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_j & \cdots & \boldsymbol{a}_n \end{array}\right]
\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \boldsymbol{a}_j
$$

Partitioned Matrices

$$\begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix} = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & B \\ O & C \end{bmatrix}$$

- Matrices composed of **submatrices**
- **Partitioned** into **blocks**

Submatrices in GNU Octave

```
M=[1,2,3;
   4,5,6;
   7,8,9]
```

- `M(2,:)=[4,5,6]`
- ```
  M(:,1)=[1;
          4;
          7]
  ```
- ```
  M(2:3,1:2)=[4,5;
              7,8]
  ```

Different Views on Matrix Multiplications

- **Outer product expansion** $A \in \mathbb{R}^{m\times n}$ as a row vector of column vectors and $B \in \mathbb{R}^{n\times r}$ as a column vector of row vectors:

$$AB = \left[\begin{array}{c:c:c} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{array}\right] \left[\begin{array}{c} \boldsymbol{b}_1 \\ \hdashline \vdots \\ \hdashline \boldsymbol{b}_n \end{array}\right] = \boldsymbol{a}_1\boldsymbol{b}_1 + \cdots + \boldsymbol{a}_n\boldsymbol{b}_n$$

$\rightarrow \boldsymbol{a}_k\boldsymbol{b}_k \in \mathbb{R}^{m\times r}$

- $A \in \mathbb{R}^{m\times n}$ as a column vector of row vectors and $B \in \mathbb{R}^{n\times r}$ as a row vector of column vectors:

$$AB = \left[\begin{array}{c} \boldsymbol{a}_1 \\ \hdashline \vdots \\ \hdashline \boldsymbol{a}_m \end{array}\right] \left[\begin{array}{c:c:c} \boldsymbol{b}_1 & \cdots & \boldsymbol{b}_r \end{array}\right] = \begin{bmatrix} \boldsymbol{a}_1\boldsymbol{b}_1 & \cdots & \boldsymbol{a}_1\boldsymbol{b}_j & \cdots & \boldsymbol{a}_1\boldsymbol{b}_r \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{a}_i\boldsymbol{b}_1 & \cdots & \boldsymbol{a}_i\boldsymbol{b}_j & \cdots & \boldsymbol{a}_i\boldsymbol{b}_r \\ \vdots & & \vdots & & \vdots \\ \boldsymbol{a}_m\boldsymbol{b}_1 & \cdots & \boldsymbol{a}_m\boldsymbol{b}_j & \cdots & \boldsymbol{a}_m\boldsymbol{b}_r \end{bmatrix}$$

$\rightarrow \boldsymbol{a}_i\boldsymbol{b}_j \in \mathbb{R}$

Block Multiplication

$$\left[\begin{array}{ccc:cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hdashline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] \left[\begin{array}{cc:cc:c} 4 & 3 & 1 & 2 & 1 \\ -1 & 2 & 2 & 1 & 1 \\ 1 & -5 & 3 & 3 & 1 \\ \hdashline 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \end{array}\right]$$

$$= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

$$= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix}$$

Matrix Powers

For a square matrix $A \in \mathbb{R}^{n \times n}$,

$$A^k = AA \cdots A$$

For nonnegative integers r and s,

- $A^r A^s = A^{r+s}$
- $(A^r)^s = A^{rs}$

Transpose

Definition: Transpose

The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T obtained by interchanging the rows and columns of A. That is, the ith column of A^T is the ith row of A for all i.

- $(A^T)_{ij} = A_{ji}$ for all i and j.
- For column vectors $\boldsymbol{u}$ and $\boldsymbol{v}$,

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v}.$$

Definition: Symmetric matrix

A square matrix A is **symmetric** if $A^T = A$ –that is, if A is equal to its own transpose.

Outline

Properties of Addition and Scalar Multiplication

Theorem 3.2: Algebraic Properties of Matrix Addition and Scalar Multiplication

Let A, B, and C be matrices of the same size and let c and d be scalars. Then

a. $A + B = B + A$ (commutativity)
b. $(A + B) + C = A + (B + A)$ (associativity)
c. $A + O = A$ (O is the identity element of the addition operator)
d. $A + (-A) = O$ ($-A$ is the inverse element of A w.r.t. the addition operator)
e. $c(A + B) = cA + cB$ (distributivity)
f. $(c + d)A = cA + dA$ (distributivity)
g. $c(dA) = (cd)A$
h. $1A = A$

Linear Combination of Matrices

$$c_1 A_1 + c_2 A_2 + \cdots + c_k A_k$$

"The matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a linear combination of the matrices $\begin{bmatrix} b_{11} & b_{12} \\ b21 & b_{22} \end{bmatrix}$ and $\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$."

$\Leftrightarrow$ "The vector $\begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} b_{11} \\ b_{12} \\ b_{21} \\ b_{22} \end{bmatrix}$ and $\begin{bmatrix} c_{11} \\ c_{12} \\ c_{21} \\ c_{22} \end{bmatrix}$."

Linear Combination of Matrices (cont'd)

- **Span** of a set of matrices (Example 3.17)
- The matrices $A_1, A_2, \cdots, A_k$ of the same size are **linearly independent** if the only solution of the equation

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = O$$

is the trivial one: $c_1 = c_2 = \cdots = c_k = 0$.

Properties of Matrix Multiplication

Theorem 3.3: Properties of Matrix Multiplication

Let A, B, and C be matrices (whose size are such that the indicated operations can be performed) and let k be a scalar. Then

a. $A(BC) = (AB)C$ (associativity)

b. $A(B + C) = AB + AC$ (left distributivity)

c. $(A + B)C = AC + BC$ (right distributivity)

d. $k(AB) = (kA)B = A(kB)$

e. $I_m A = A = AI_n$ if $A \in \mathbb{R}^{m \times n}$ (multiplicative identity)

- If $A^2 = O$ then $A = O$? $\rightarrow$ Example 3.19
- $(A + B)^2 = A^2 + 2AB + B^2$? $\rightarrow$ Example 3.20

Properties of the Transpose

Theorem 3.4: Properties of the Transpose

Let A and B be matrices (whose size are such that the indicated operations can be performed) and let k be a scalar. Then

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. $(kA)^T = k(A^T)$

d. $(AB)^T = B^T A^T$

e. $(A^r)^T = (A^T)^r$ for all nonnegative integers r

- $(A_1 A_2 \cdots A_k)^T =?$ $\rightarrow$ Exercise 33

Theorem 3.5

a. If A is a square matrix, then $A + A^T$ is a symmetric matrix.

b. For any matrix A, (*not necessarily square matrix*) AA^T and $A^T A$ are symmetric matrices.

$\rightarrow$ Prove them!

Outline

Solving an Equation

$$\begin{aligned} a+x=b \quad &\Rightarrow -a+(a+x)=-a+(b) \quad &&\Rightarrow (-a+a)+x=b-a \\ &\Rightarrow 0+x=b-a &&\Rightarrow x=b-a \end{aligned}$$

$$ax=b \Rightarrow \frac{1}{a}(ax)=\frac{1}{a}(b) \Rightarrow \left(\frac{1}{a}(a)\right)x=\frac{b}{a} \Rightarrow 1\cdot x=\frac{b}{a} \Rightarrow x=\frac{b}{a}$$

How to solve the equation "$a \star x = b$"?

1. Find the **inverse element** of a, say a', with respect to the (binary) operator $\star$ to get the **identity element** of $\star$, say I, on the left-hand side.
   $a' \star (a \star x) = a' \star b \Rightarrow I \star x = a' \star b$
2. Now we have only x on the left-hand side therefore solve the equation.
   $x = a' \star b$

Solving the Linear System $A\boldsymbol{x} = \boldsymbol{b}$

$A\boldsymbol{x} = \boldsymbol{b} \Rightarrow A'(A\boldsymbol{x}) = A'\boldsymbol{b} \Rightarrow (A'A)\boldsymbol{x} = A'\boldsymbol{b} \Rightarrow I\boldsymbol{x} = A'\boldsymbol{b} \Rightarrow \boldsymbol{x} = A'\boldsymbol{b}$

Two questions:

- *When* can we find such a matrix A'?
- *How* can we compute A'?

Definition: Inverse Matrix

If A is an $n \times n$ matrix, an **inverse** of A is an $n \times n$ matrix A' with the property that

$$AA' = I \quad \text{and} \quad A'A = I$$

where $I = I_n$ is the $n \times n$ identity matrix. If such an A' exists, then A is called **invertible**.

- $AA' = A'A = I \rightarrow A$ and A' are square matrices
- A non-square matrix may or may not have a left-inverse or a right-inverse.

Inverse Matrix

Questions:

- How can we know when a matrix has an inverse?
- If a matrix does have an inverse, how can we find it?
- Can a matrix have more than one matrix?

Theorem 3.6

If A is an invertible matrix, then its inverse is unique.

- "THE" inverse $\rightarrow A^{-1}$

Solving a Linear System using the Inverse Matrix

Theorem 3.7

If A is an invertible $n \times n$ matrix, then the system of linear equations given by $A\boldsymbol{x} = \boldsymbol{b}$ has the unique solution $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ for any $\boldsymbol{b} \in \mathbb{R}^n$.

- “Existence” and “uniqueness”

Inverse Matrix of a 2×2 Matrix

Theorem 3.8

1. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then A is invertible if $ad - bc \neq 0$, in which case

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

2. If $ad - bc = 0$, then A is not invertible.

Properties of Invertible Matrices

Theorem 3.9

If A is an invertible matrix

a. then A^{-1} is invertible and $(A^{-1})^{-1} = A$

b. and c is a nonzero scalar, then cA is an invertible matrix and $(cA)^{-1} = \frac{1}{c}A^{-1}$

c. and B is an invertible matrix of the same size, then AB is invertible and (socks-and-shoes rule) $(AB)^{-1} = B^{-1}A^{-1}$

d. then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$

e. then A^n is invertible for all nonnegative integers n and $(A^n)^{-1} = (A^{-1})^n$

- $(A_1 A_2 \cdots A_n)^{-1} =$?
- $(A + B)^{-1} = A^{-1} + B^{-1}$? $\rightarrow$ Exercise 19
- $A^{-n} := (A^{-1})^n = (A^n)^{-1}$

Elementary Matrices

Example:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 5 & 7 \\ -1 & 0 \\ 8 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 7 \\ 8 & 3 \\ -1 & 0 \end{bmatrix}$$

$\longrightarrow$ Row-interchanging by multiplying an matrix.

Definition

An **elementary matrix** is any matrix that can be obtained by performing an elementary row operation on an identity matrix.

- $R_i \leftrightarrow R_j$
- kR_i
- $R_i + kR_j$

Elementary Matrices (cont'd)

Theorem 3.10

Let E be the elementary matrix obtained by performing an elementary row operation on I_n. If the same elementary row operation is performed on an $n \times r$ matrix A, the result is the same as the matrix EA.

- Applying elementary row operations E_1, E_2 and E_3, in this order, to a matrix A is the same as applying the operations to I first and then applying the resulting matrix:

$$E_3(E_2(E_1A)) = (E_3E_2E_1I)A$$

- "Elementary row operations are *reversible*."
  $\Rightarrow$ "Elementary matrices are *invertible*."

Theorem 3.11

Each elementary matrix is invertible, and its inverse is an elementary matrix of the same type.

The Fundamental Theorem of Invertible Matrices

- What does it mean that "a matrix is invertible"?

Theorem 3.12: The Fundamental Theorem of Invertible Matrices: Version 1

Let A be an $n \times n$ matrix. The following statements are equivalent:

a. A is invertible.
b. $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for every $\boldsymbol{b} \in \mathbb{R}^n$.
c. $A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution.
   $\rightarrow$ Columns of A are linearly independent.
d. The reduced row echelon form of A is I_n.
e. A is a product of elementary matrices.

The Fundamental Theorem of Invertible Matrices (cont'd)

The power of the "Fundamental Theorem":

Theorem 3.13

Let A be a square matrix. If B is a square matrix such that either $AB = I$ of $BA = I$, then A is invertible and $B = A^{-1}$.

Theorem 3.14

Let A be a square matrix. If a sequence of elementary row operations reduces A to I, then the same sequence of elementary row operations transforms I into A^{-1}.

Computing the Inverse of an $n \times n$ Matrix

Elementary row operations to yield

$$[A|I] \longrightarrow [I|A^{-1}]$$

Several views:

1. Gauss-Jordan elimination performed on an $n \times 2n$ augmented matrix.
2. Solving the matrix equation $AX = I_n$ for an $n \times n$ matrix X.
3. Solving n linear systems

$$A\boldsymbol{x}_1 = \boldsymbol{e}_1, A\boldsymbol{x}_2 = \boldsymbol{e}_2, \cdots, A\boldsymbol{x}_n = \boldsymbol{e}_n$$

$$\rightarrow [A|\boldsymbol{e}_1\ \boldsymbol{e}_2 \cdots \boldsymbol{e}_n] = [A|I_n]$$

- If A cannot be reduced to I, then A is not invertible.

Outline

Matrix Factorization/Decomposition

- Integer/prime factorization

$$20 = 2 \cdot 3 \cdot 5$$

- Polynomial factorization

$$2x^2 + 7x + 3 = (2x + 1)(x + 3)$$

- Matrix factorization: Representation of a matrix as a product of two or more other matrices

$$\begin{bmatrix} 3 & -1 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}$$

  - LU factorization → Sec 3.4
  - QR factorization → Sec 5.3
  - SVD (Singular Value Decomposition) → Sec 7.4

Revisiting Gaussian Elimination

Example 3.33

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

$$\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} =: U$$

$$A \to E_3 E_2 E_1 A = U \to A = (E_1^{-1} E_2^{-1} E_3^{-1})U$$

$$E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix} =: L$$

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

LU Factorization

Example 3.33

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

$$A = L \quad U$$

L: unit lower triangular matrix

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ * & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & 1 & 0 \\ * & * & \cdots & * & 1 \end{bmatrix}$$

U: upper triangular matrix (p.160)

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

Definition

Let A be a square matrix. A factorization of A as $A = LU$, where L is unit lower triangular and U is upper triangular, is called an LU **factorization** of A.

LU Factorization (cont'd)

Questions:

- Does an LU factorization always exist?
- How can we find the LU factorization of a matrix?
- Is it unique?
- Why is it useful?

Theorem 3.15

If A is a square matrix that can be reduced to row echelon form without using any row interchanges, then A has an LU factorization.

$\rightarrow$ Why? $\rightarrow$ See the remarks on p.179-180.

Solving a Linear System Using LU Factorization

For the linear system

$$A\boldsymbol{x} = \boldsymbol{b},$$

if A has an LU factorization $A = LU$, we can solve the linear system as follows:

1. Solve $L\boldsymbol{y} = \boldsymbol{b}$ for $\boldsymbol{y}$, where $\boldsymbol{y} := U\boldsymbol{x}$, by *forward substitution*.
2. Solve $\boldsymbol{y} = U\boldsymbol{x}$ for $\boldsymbol{x}$ by *back substitution*.

$\rightarrow$ Example 3.34 (p.180)

How to Find $A = LU$? – Without Any Row Interchange

Example 3.35

1. $R_2 - 2R_1 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

2. $R_3 - 1R_1 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

3. $R_4 - (-3)R_1 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & * & 1 & 0 \\ -3 & * & * & 1 \end{bmatrix}$$

4. $R_3 - \frac{1}{2}R_2 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & * & * & 1 \end{bmatrix}$$

5. $R_4 - 4R_2 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & * & 1 \end{bmatrix}$$

6. $R_4 - (-1)R_3 \rightarrow$
$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix}$$

- The order is important! (See remark on p.183)
  $\rightarrow$ from top to bottom, column by column from left to right
- Does this always work?

How to Find $A = LU$? (cont'd)

Theorem 3.16

If A is an invertible matrix that has an LU factorization, then L and U are invertible.

- What if we need row exchange in during Gauss elimination?

Example (p.184)

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & 3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix} = U = PEA$$

Let's exchange the 2nd and 3rd rows first!

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 4 \\ 3 & 6 & 2 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix} = U = EPA$$

$P^T LU$ Factorization – With Row Interchange

Permutation matrix

- Product of row interchange matrices
- Constructed by permutating the rows of an identity matrix
  $\rightarrow$ related to "picking a row of a matrix"

With the **permutation matrix** P,

$$EPA = U \rightarrow A = (EP)^{-1}U = P^{-1}E^{-1}U = P^{-1}LU$$

Theorem 3.17

If P is a permutation matrix, then $P^{-1} = P^T$.

- $A = P^{-1}LU = P^T LU$

Definition: $P^T LU$ Factorization

Let A be a square matrix. A factorization of A as $A = P^T LU$, where P is a permutation matrix, L is unit lower triangular, and U is upper triangular, is called a $P^T LU$ **factorization** of A.

$P^T LU$ Factorization (cont'd)

Theorem 3.18

Every square matrix has a $P^T LU$ factorization.

- Is it unique?
- How about the zero matrix?

Outline

Geometric Views

- How can we generalize **lines** and **planes** through the origin to higher dimensions?
  $\longrightarrow$ Subspace (of a vector space, see Chapter 6)
- How can we generalize **direction vectors** to higher dimensions?
  $\longrightarrow$ Basis
- How can we generalize the concept of "the **dimension** of a subspace"?

$\longrightarrow$ More in Chapter 6

Subspaces

- The set of vectors in $\mathbb{R}^2$ are **closed** under (i) addition and (ii) scalar multiplication.
- How about the vectors in a plane (through the origin) in $\mathbb{R}^3$?
  $\rightarrow$ Yes!
  - the vectors are 3-dimensional vectors
  - the plane is 2-dimensional
- How can we describe the plane then?

Subspaces (cont'd)

Definition

A **subspace** of $\mathbb{R}^n$ is any collection S of vectors in $\mathbb{R}^n$ such that

1. The zero vector $\mathbf{0}$ is in S.
2. If $\boldsymbol{u}$ and $\boldsymbol{v}$ are in S, then $\boldsymbol{u} + \boldsymbol{v}$ is in S. (S is **closed under addition**.)
3. If $\boldsymbol{u}$ is in S and c is a scalar, then $c\boldsymbol{u}$ is in S. (S is **closed under scalar multiplication**.)

Conditions 2&3
$\rightarrow$ S is **closed under linear combinations**:
If $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ are in S and $c_1, c_2, \cdots, c_k$ are scalars, then $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$ is in S.

Subspaces and Spanning Sets

Are the followings subspaces?

- A plane through the origin in $\mathbb{R}^3$? $\rightarrow$ Example 3.37
- A line through the origin in $\mathbb{R}^2$?
- A line through the origin in $\mathbb{R}^3$?
- $\{\mathbf{0}\}$?

$\rightarrow$ The dimension of the vectors does not matter!

- $\mathbb{R}^2$ is the spanning set of two linearly independent vectors (Sec 2.3)
- $\mathbb{R}^2$ *looks the same* as a plane through the origin

$\rightarrow$ A plane through the origin is the spanning set of two linearly independent vectors.

Theorem 3.19

Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k$ be vectors in $\mathbb{R}^n$. Then $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k)$ is a subspace of $\mathbb{R}^n$.

$\rightarrow$ $\mathrm{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k)$ is **the subspace spanned by** $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$.

Subspaces Associated with Matrices: Row Spaces and Column Spaces

- For a matrix $A \in \mathbb{R}^{m \times n}$ and a column vector $\boldsymbol{x} \in \mathbb{R}^n$,

$$A\boldsymbol{x}$$

  can be viewed as a linear combination of the columns of A.
- How about

$$\boldsymbol{x}A$$

  with a row vector $\boldsymbol{x} \in \mathbb{R}^m$ and a matrix $A \in \mathbb{R}^{m \times n}$?

Subspaces Associated with Matrices: Row Spaces and Column Spaces (cont'd)

Definition

Let $A \in \mathbb{R}^{m \times n}$.

1. The **row space** of A is the subspace $\mathrm{row}(A)$ of $\mathbb{R}^n$ spanned by the rows of A.
2. The **column space** of A is the subspace $\mathrm{col}(A)$ of $\mathbb{R}^m$ spanned by the columns of A.

$\rightarrow$ Example 3.41

- Do the elementary row operations change the row space of a matrix?

Theorem 3.20

Let B be any matrix that is row equivalent to (See the definition on p.72) a matrix A. Then $\mathrm{row}(B) = \mathrm{row}(A)$.

- How about the column spaces?
  $\mathrm{col}(B) \neq \mathrm{col}(A)$! (See the warning on p.199.)

Subspaces Associated with Matrices: Null Spaces

- Is the set of solutions of a linear system a subspace?

Theorem 3.21

Let A be an $m \times n$ matrix and let N be the set of solutions of the homogeneous linear systems $A\boldsymbol{x} = \mathbf{0}$. Then N is a subspace of $\mathbb{R}^n$.

- What does it called?

Definition: Null Space

Let A be an $m \times n$ matrix. The **null space** of A is the subspace of $\mathbb{R}^n$ consisting of solutions of the homogeneous linear system $A\boldsymbol{x} = \mathbf{0}$. It is denoted by $\mathrm{null}(A)$.

Solutions of a Linear System

See p.61

Theorem 3.22

Let A be a matrix whose entries are real numbers. For any system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$, exactly one of the following is true:

1. There is no solution.
2. There is a unique solution.
3. There are infinitely many solution.

$\rightarrow$ Can be proved using the fact that the null space of a matrix is a subspace.

Basis

- Which vectors do we need to generate a line or a plane (through the origin), respectively?
- How can we generalize this fact?

Definition: Basis

A **basis** for a subspace S of $\mathbb{R}^n$ is a set of vectors in S that

1. spans S and
2. is linearly independent.

Example: $\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n \in \mathbb{R}^n \rightarrow$ **standard basis**

- For a subspace, how many bases are there?

Finding a Basis for a Subspace

How to find a basis for $\mathrm{row}(A)$, $\mathrm{col}(A)$, and $\mathrm{null}(A)$, respectively?
$\rightarrow$ Example 3.45, Example 3.47, Example 3.48

Procedure to find bases for $\mathrm{row}(A)$, $\mathrm{col}(A)$, and $\mathrm{null}(A)$

1. Find the reduced row echelon form R of A.
2. Use the nonzero row vectors of R (containing the leading 1s) to form a basis for $\mathrm{row}(A)$.
3. Use the column vectors of A that correspond to the columns of R containing the leading 1s (the pivot columns) to form a basis for $\mathrm{col}(A)$.
4. Solve for the leading variables of $R\boldsymbol{x} = \mathbf{0}$ in terms of the free variables, set the free variables equal to parameters, substitute back into $\boldsymbol{x}$, and write the result as a linear combination of f vectors (where f is the number of free variables). These f vectors form a basis for $\mathrm{null}(A)$.

- (Non-reduced) row echelon form is enough for $\mathrm{row}(A)$ and $\mathrm{col}(A)$. (p.200)

Dimension

- How many vectors do we need for a basis?

Theorem 3.23: The Basis Theorem

Let S be a subspace of $\mathbb{R}^n$. Then any two bases for S have the same number of vectors.

- What does the number called?

Definition: Dimension

if S is a subspace of $\mathbb{R}^n$, then the number of vectors in a basis for S is called the **dimension** of S, denoted $\dim S$.

- $\dim\{\mathbf{0}\} =?$
- $\dim \mathbb{R}^n =?$

Rank

- $\dim(\text{row}(A)) =?$ $\dim(\text{col}(A)) =?$ $\dim(\text{null}(A)) =?$ (Example 3.50)

Theorem 3.24

The row and column spaces of a matrix A have the same dimension.

- What does $\dim(\text{row}(A))$ or $= \dim(\text{col}(A))$ called?

Definition: Rank

The **rank** of a matrix A is the dimension of its row and column spaces and is denoted by $\text{rank}(A)$.

- Is this definition equivalent to the one on p.75? Why?
- What is the relation between $\text{rank}(A)$ and $\text{rank}(A^T)$?

Theorem 3.25

For any matrix A,

$$\text{rank}(A^T) = \text{rank}(A)$$

Nullity

- $\dim(\text{null}(A)) =?$

Definition: Nullity

The **nullity** of a matrix A is the dimension of its null space and is denoted by $\text{nullity}(A)$.

- $\text{nullity}(A)$
- Dimension of the solution space of $A\boldsymbol{x} = \mathbf{0}$
- Number of free variables in the solution of $A\boldsymbol{x} = \mathbf{0}$

All the above are the same. Why?

- See Theorem 2.2 on p.75
  $\rightarrow$ What is the relation between $\text{rank}(A)$ and $\text{nullity}(A)$?

Theorem 3.26: The Rank Theorem

If A is an $m \times n$ matrix, then

$$\text{rank}(A) + \text{nullity}(A) = n$$

Fundamental Theorem of Invertible Matrices: Ver 2

Theorem 3.27

Let A be an $n \times n$ matrix. The following statements are equivalent:

a. A is invertible.
b. $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for every $\boldsymbol{b}$ in $\mathbb{R}^n$.
c. $A\boldsymbol{x} = \boldsymbol{0}$ has only the trivial solution.
d. The reduced row echelon form of A is I_n.
e. A is a product of elementary matrices.
f. $\text{rank}(A) = n$
g. $\text{nullity}(A) = 0$
h. The column vectors of A are linearly independent.
i. The column vectors of A span $\mathbb{R}^n$.
j. The column vectors of A form a basis for $\mathbb{R}^n$.
k. The row vectors of A are linearly independent.
l. The row vectors of A span $\mathbb{R}^n$.
m. The row vectors of A form a basis for $\mathbb{R}^n$.

Applications

Theorem 3.28

Let A be an $n \times m$ matrix. Then

a. $\text{rank}(A^T A) = \text{rank}(A)$

b. The $n \times n$ matrix $A^T A$ is invertible iff $\text{rank}(A) = n$.

$\rightarrow$ Prove them using the Rank Theorem and the Fundamental Theorem!

Coordinates

- What is the relation between vectors in a subspace and a basis for that subspace?

Theorem 3.29

Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k\}$ be a basis for S. For every vector $\boldsymbol{v}$ in S, there is exactly one way to write $\boldsymbol{v}$ as a linear combination of the basis vectors in $\mathcal{B}$:

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k$$

Coordinates (cont'd)

- What does the "way" (coefficients of unique linear combination for $\boldsymbol{v}$) called?

Definition: Coordinates

Let S be a subspace of $\mathbb{R}^n$ and let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k\}$ be a basis for S. Let $\boldsymbol{v}$ be a vector in S, and write $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k$. Then $c_1, c_2, \cdots, c_k$ are called the **coordinates of v with respect to $\mathcal{B}$**, and the column vector

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

is called the **coordinate vector of v with respect to $\mathcal{B}$**.

- What does the Cartesian coordinate of a vector mean?

Outline

Matrices as Functions

- "A function transforms a real number into another real number."

$$f : \mathbb{R} \to \mathbb{R}$$

- Matrices as functions acting on vectors: "A $m \times n$ matrix **transforms** a column vector in $\mathbb{R}^n$ into another column vector in $\mathbb{R}^m$."

$$A \in \mathbb{R}^{m \times n} : \mathbb{R}^n \to \mathbb{R}^m$$

  - **transformations**, **mapping** or **function**
  - **domain**: $\mathbb{R}^n$
  - **codomain**: $\mathbb{R}^m$
  - **image** of $\boldsymbol{x} \in \mathbb{R}^n$: $A\boldsymbol{x}$
  - **range** of A:
    $\{\boldsymbol{y} \in \mathbb{R}^m | \boldsymbol{y} = A\boldsymbol{x} \text{ for some } \boldsymbol{x} \in \mathbb{R}^n\} = \mathrm{col}(A)$

Linear Transformations

- What kind of transformations are they (transformations by matrices)?

Definition: Linear Transformation

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a **linear transformation** if

1. $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ for all $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$ and
2. $T(c\boldsymbol{v}) = cT(\boldsymbol{v})$ for all $\boldsymbol{v}$ in $\mathbb{R}^n$ and for all scalars c.

Remark

$T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if

$$T(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) = c_1T(\boldsymbol{v}_1) + c_2T(\boldsymbol{v}_2)$$

for all $\boldsymbol{v}_1, \boldsymbol{v}_2$ in $\mathbb{R}^n$ and scalars c_1, c_2.

- See Exercise 53.
- $T(c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k) =$?

Linear Transformations (cont'd)

- Are all the matrix transformatios linear transformations?

Theorem 3.30

Let A be an $m \times n$ matrix. Then the matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T_A(\boldsymbol{x}) = A\boldsymbol{x} \quad (\text{for } \boldsymbol{x} \text{ in } \mathbb{R}^n)$$

is a linear transformation.

- How about its converse? Are all the linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ matrix transformations?

Theorem 3.31

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is a matrix transformation. More specifically, $T = T_A$, where A is the $m \times n$ matrix

$$A = \left[\; T(\boldsymbol{e}_1) \;\vdots\; T(\boldsymbol{e}_2) \;\vdots\; \cdots \;\vdots\; T(\boldsymbol{e}_n) \;\right]$$

- A: "**standard matrix of the linear transformation** T"

Linear Transformations (cont'd)

- Notation
  $[T]$ means the standard matrix of a linear transformation T.
- What kinds ot linear transformations are there?
  - Reflection (Example 3.56)
  - Rotation (Example 3.57, 3.58)
  - Projection (Example 3.59)
  - ...And more – Scaling, Shearing, Squeezing
    See `http://en.wikipedia.org/wiki/Linear_transformation`.
  - Translation...?

Successive Linear Transformations

- **Composition** of two functions

$$(f \circ g)(x) = f(g(x))$$

- **Composition** of two linear transformations $T : \mathbb{R}^m \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^p$

$$(S \circ T)(\boldsymbol{x}) = S(T(\boldsymbol{x}))$$

Theorem 3.32

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ and $S : \mathbb{R}^n \to \mathbb{R}^p$ be linear transformations. Then $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$ is a linear transformation. Moreover, their standard matrices are related by

$$[S \circ T] = [S][T]$$

Inverse of Linear Transformations

- We can consider the **Identity transformation** defined as "$I_n : \mathbb{R}^n \to \mathbb{R}^n$ such that $I_n(\boldsymbol{v}) = \boldsymbol{v}$ for every $\boldsymbol{v}$ in $\mathbb{R}^n$."
- How can we define an **inverse transformation** of a linear transformation?

Definition

Let S and T be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then S and T are **inverse transformations** if $S \circ T = I_n$ and $T \circ S = I_n$.

- What is the standard matrix of the identity transformation?
- Does every linear transformation have its inverse?
  $\to$ **invertible** transformations
- Is it unique?

Inverse of Linear Transformations (cont'd)

Theorem 3.33

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then its standard matrix $[T]$ is an invertible matrix, and

$$[T^{-1}] = [T]^{-1}$$

- "The matrix of the inverse is the inverse of the matrix."
  $\rightarrow$ "The (standard) matrix of the inverse (transformation) is the inverse (matrix) of the (standard) matrix (of the transformation)."

Prooving the Associativity of Matrix Multiplication

- Associativity of matrix multiplication (Theorem 3.3(a) on p.156)

$$A(BC) = (AB)C$$

- Can be proved using the fact that

$$A(BC) = (AB)C \quad \text{iff} \quad R \circ (S \circ T) = (R \circ S) \circ T$$

where $R = T_A$, $S = T_B$ and $T = T_C$.

Outline

Applications

- Robotics
- Markov chains
- Population growth
- Graphs and Digraphs
- Error-correcting codes