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.** → “linear operators”
- Matrices **transform** a vector into another vector. (Problem 1)
- Matrices transform a parallelogram into another one. (Problem 2-3)
- 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.

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

where

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

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**

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

- **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.

$$\left[\begin{array}{c} \mathbf{a}_1 \\ \hdashline \vdots \\ \hdashline \mathbf{a}_i \\ \hdashline \vdots \\ \hdashline \mathbf{a}_m \end{array}\right] \left[\begin{array}{c;{2pt/2pt}c;{2pt/2pt}c;{2pt/2pt}c;{2pt/2pt}c} \mathbf{b}_1 & \cdots & \mathbf{b}_j & \cdots & \mathbf{b}_r \end{array}\right] = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_j & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_r \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{a}_i \cdot \mathbf{b}_1 & \cdots & \mathbf{a}_i \cdot \mathbf{b}_j & \cdots & \mathbf{a}_i \cdot \mathbf{b}_r \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \cdots & \mathbf{a}_m \cdot \mathbf{b}_j & \cdots & \mathbf{a}_m \cdot \mathbf{b}_r \end{bmatrix}$$

- Example 3.7 → Application of matrix multiplication

Matrices and Linear Systems

- Example 3.8

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

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, $\mathbf{e}_i$ a $1 \times m$ standard unitvector, and $\mathbf{e}_j$ an $n \times 1$ standard unitvector. Then

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

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

$$[0 \;\cdots\; 1 \;\cdots\; 0] \left[\begin{array}{c} \mathbf{a}_1 \\ \hdashline \vdots \\ \hdashline \mathbf{a}_i \\ \hdashline \vdots \\ \hdashline \mathbf{a}_m \end{array}\right] = \mathbf{a}_i$$

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

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

- Notation: "$A \in \mathbb{R}^{m \times n}$" means "$A$ is an $m \times n$ matrix."
- **Outer product expansion** of AB:
  - $A \in \mathbb{R}^{m \times n}$ as a row vector of column vectors
  - $B \in \mathbb{R}^{n \times r}$ as a column vector of row vectors

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

$\rightarrow \mathbf{a}_k\mathbf{b}_k \in \mathbb{R}^{m \times r}$ ($\mathbf{a}_k\mathbf{b}_k$ is an $m \times r$ matrix.)

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

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

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

Block Multiplication

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

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

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

- Why is it possible?

Matrix Powers

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

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

For nonnegative integers r and s,

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

$\rightarrow$ Example 3.13 (Mathematical induction)

- For convenience, we *define* $A^0 := I_n = I$.

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 $\mathbf{u}$ and $\mathbf{v}$,

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

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. $(A + B) + C = A + (B + C)$ (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

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

- Example 3.16

"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

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

  is the trivial one: $c_1 = c_2 = \cdots = c_k = 0$.
- Example 3.18

Properties of Matrix Multiplication

- Example 3.19
  - Is matrix multiplication commutitative?
  - Is this statement true? "If $A^2 = O$, then $A = O$"

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

- $(A + B)^2 = A^2 + 2AB + B^2$? (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

- $(A_1 + A_2 + \cdots + A_k)^T =$?
- $(A_1 A_2 \cdots A_k)^T =$? $\to$ 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.

Outline

Solving an Equation

$$a + x = b \quad \Rightarrow -a + (a + x) = -a + (b) \quad \Rightarrow (-a + a) + x = b - a$$
$$\Rightarrow 0 + x = b - a \quad \Rightarrow x = b - a$$

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

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.
   $a' \star (a \star x) = a' \star b \Rightarrow (a' \star a) \star x = a' \star b \Rightarrow I \star x = a' \star b \Rightarrow x = a' \star b$
2. Now we have only x on the left-hand side therefore can solve the equation.
   $x = a' \star b$

- Is it always possible? Which properties should the operator $\star$ have?

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

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

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. $\rightarrow$ "pseudoinverse" (p.594)
- In fact, we only need to try either "$AA' = I$" or "$A'A = I$" to check if A' is the inverse of A. (Theorem 3.13)

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 inverse matrix?

Theorem 3.6

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

- Proving "uniqueness" $\rightarrow$ Show that there cannot be more than one
- "THE" inverse $\rightarrow$ A^{-1}
- Why "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 $A\mathbf{x} = \mathbf{b}$ <u>has</u> the <u>unique</u> solution $\mathbf{x} = A^{-1}\mathbf{b}$ for any $\mathbf{b} \in \mathbb{R}^n$.

- "Existence" and "uniqueness"
- Proving "existence" $\rightarrow$ Show that $\mathbf{x} = A^{-1}\mathbf{b}$ satisfies the equation
- Proving "uniqueness" $\rightarrow$ Show that any other solution, say $\mathbf{y}$, must be the same as $\mathbf{x}$

Inverse Matrix of a 2×2 Matrix

- How can we tell if a matrix is invertible?
- If a matrix is invertible, how can we find it?

Theorem 3.8

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

1. 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.

- $\det A = ad - bc \rightarrow$ *determinant* of A (Section 4.2)
- "A is invertible iff $\det A \neq 0$" $\rightarrow$ True for all square matrices.

Solving a Linear System

- Gauss-Jordan (or Gaussian) elimination vs. computing the inverse matrix
- Which is better? Why?
  (See the remark below Example 3.25 and try Excercise 13)
- Computing the inverse ...
  - is slower.
  - works only when the matrix is square & invertible.
  - does not handle well the case of infinitely many solutions.

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 $(A^n)^{-1} = (A^{-1})^n$

- $(A_1 A_2 \cdots A_n)^{-1} =$? $\rightarrow$ Products of invertible matrices are invertible
- $(A + B)^{-1} = A^{-1} + B^{-1}$? $\rightarrow$ Exercise 19
- $A^{-n} := (A^{-1})^n = (A^n)^{-1}$
  $\rightarrow$ "$A^r A^s = A^{r+s}$" and "$(A^r)^s = A^{rs}$" holds for *all integers* r and s, if A is invertible.

Elementary Matrices

- A difference perspective on the row reduction
- Row reduction as a series of matrix multiplications
- New and important insights into the nature of invertible matrices

Example:

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

$\rightarrow$ 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"?
- A set of equivalent characterizations of what it means for a matrix to be 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. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{R}^n$. (Theorem 3.7)

c. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. (Theorem 3.7)
   $\rightarrow$ Columns of A are linearly independent. (Theorem 2.6)

d. The reduced row echelon form of A is I_n.

e. A is a product of elementary matrices. (Example 3.29)

The Fundamental Theorem of Invertible Matrices (cont'd)

- How powerful is the "Fundamental Theorem"?

Theorem 3.13

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

- To check if $B = A^{-1}$, we only need to check if $AB = I$ OR $BA = I$, not both.

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}.

- Theorem 3.14 $\rightarrow$ We can compute A^{-1} via Gauss-Jordan elimination.

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

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

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

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

Outline

Matrix Factorization/Decomposition

- Integer/prime factorization

$$30 = 2 \cdot 3 \cdot 5$$

- 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 $\rightarrow$ Sec 3.4
  - QR factorization $\rightarrow$ Sec 5.3
  - SVD (Singular Value Decomposition) $\rightarrow$ Sec 7.4

Revisiting Gaussian Elimination

Example 3.33

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ -2 & 5 & 5 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftarrow R_3 + R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix}$$

$$\xrightarrow{R_3 \leftarrow R_3 + 2R_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} =: U$$

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

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

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

Revisiting Gaussian Elimination (cont'd)

Assuming no row interchange is required, let A be reduced to U (using Gaussian elimination) as $U = (E_m E_{m-1} \cdots E_1)A$.

- To reduce a matrix to row echelon form, we only need one type of elementary operation: $R_i \leftarrow R_i - kR_j$ where $i > j$. (Why?)
- The elementary matrix associated with the above operation is unit lower triangular (ULT) matrix. (Why?)
- Since
  - the inverse of a ULT matrix is also a ULT matrix, (Why? See Excercise 30) and
  - the product of ULT matrices is also a ULT matrix (Why? See Excercise 29)

  $E_1^{-1} E_2^{-1} \cdots E_m^{-1}$ is also a ULT matrix.
- Therefore,

$$
\begin{aligned}
& U = (E_m E_{m-1} \cdots E_1)A \\
\rightarrow & A = (E_m E_{m-1} \cdots E_1)^{-1} U = (E_1^{-1} E_2^{-1} \cdots E_m^{-1}) U = LU.
\end{aligned}
$$

LU Factorization

Example 3.33

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

$$A \qquad = \qquad L \qquad\qquad U$$

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.168/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 above Theorem 3.15

Solving a Linear System Using LU Factorization

For the linear system

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

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

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

- Example 3.34
- Why is this method good?

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 the remark)
  $\rightarrow$ from top to bottom, column by column from left to right
- Why does it work?
- Does this always work?

Is LU Factorization Unique for a Matrix?

Theorem 3.16

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

$P^T LU$ Factorization

- What if we need row exchange during Gauss elimination?

Example

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix} \xrightarrow[R_3+R_1]{R_2-3R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & 3 & 3 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix} = U = PE_1A$$

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

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

$$\begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 4 \\ 3 & 6 & 2 \end{bmatrix} \xrightarrow[R_3-3R_1]{R_2+R_1} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix} = U = E_2PA$$

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

P^TLU Factorization – With Row Interchange

Permutation matrix

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

With the **permutation matrix** P,

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

Theorem 3.17

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

- $A = P^{-1}LU = P^TLU$
- P is an *orthogonal matrix*. (Sec 5.1)

Definition: P^TLU Factorization

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

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

- Does $P^T LU$ factorization exist for any matrix?

Theorem 3.18

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

- Is it unique? $\rightarrow$ See the remark below Theorem 3.18
- How about the zero matrix?
- How can we solve the linear system $A\mathbf{x} = \mathbf{b}$ where $A = P^T LU$? (See Excercise 27, 28)
  1. $A\mathbf{x} = \mathbf{b} \rightarrow P^T LU\mathbf{x} = \mathbf{b} \rightarrow LU\mathbf{x} = P\mathbf{b}$
  2. Let $\mathbf{b}' := P\mathbf{b}$ then solve $UL\mathbf{x} = \mathbf{b}'$ via forward substitution followed by back substitution.
- How about rectangular matrices?
  $\rightarrow A, U \in \mathbb{R}^{m\times n}$ and $P, L \in \mathbb{R}^{m\times m}$

Outline

Geometry and Algebra

Geometry	Algebra
Lines & planes (through the origin)	*Subspaces*
Direction vectors for lines & planes	*Basis*
Dimension of lines & planes	How to define?

Let $\mathscr{P}$ be a plane through the origin in $\mathbb{R}^3$.

- What is the difference between $\mathbb{R}^2$ and $\mathscr{P}$?
- What is the difference between the vectors in $\mathbb{R}^2$ and $\mathscr{P}$?
- Operations on the vectors in $\mathscr{P}$?
- Are the vectors in $\mathscr{P}$ two-dimensional or three-dimensional?
- More in Chapter 6

Review on Lines and Planes Through The Origin

Let ℓ be a line through the origin with direction vector $\mathbf{d}$.

- The vector form of ℓ is "$\mathbf{x}(t) = t\mathbf{d}$."
- Any vector in ℓ is of the form $t\mathbf{d}$ for some t.
- Any vector in ℓ is a *linear combination* of $\mathbf{d}$
- $\ell = \operatorname{span}(\mathbf{d})$

Let $\mathscr{P}$ be a plane through the origin with direction vectors $\mathbf{u}$ and $\mathbf{v}$.

- The vector form of $\mathscr{P}$ is "$\mathbf{x}(s, t) = s\mathbf{u} + t\mathbf{v}$."
- Any vector in $\mathscr{P}$ is of the form $s\mathbf{u} + t\mathbf{v}$ for some s and t.
- Any vector in $\mathscr{P}$ is a *linear combination* of $\mathbf{u}$ and $\mathbf{v}$.
- $\mathscr{P} = \operatorname{span}(\mathbf{u}, \mathbf{v})$

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

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

- Example 3.37
  - Every line and plane through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$.
  - The dimension of vectors does not matter!
    $\rightarrow$ Can be generalized beyond $\mathbb{R}^3$

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 $\mathbb{R}^2$ or $\mathbb{R}^3$?
- $\{\mathbf{0}\}$?
- Example 3.39/3.40?

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

- $\mathbb{R}^2$ is the span 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 span of two linearly independent vectors.

Theorem 3.19

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

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

Subspaces Associated with Matrices: Row Spaces and Column Spaces

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

  $$A\mathbf{x}$$

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

  $$\mathbf{x}A$$

  with a row vector $\mathbf{x} \in \mathbb{R}^m$ 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.
   - A.k.a. *range* of A. (Why?)

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

Example 3.41

- $\mathbf{b} \in \mathrm{col}(A) \Leftrightarrow$ "There exist *some* $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$"
  $\Leftrightarrow$ "$A\mathbf{x} = \mathbf{b}$ is consistent"
- $\mathbf{w} \in \mathrm{row}(A) \Leftrightarrow$ "$\mathbf{w}$ can be represented as a l.c. of the rows of A"
  $\Leftrightarrow$ "$\mathbf{w}$ and the rows of A are linearly dependent."
  $\Leftrightarrow$ "$\begin{bmatrix} A \\ \mathbf{w} \end{bmatrix}$ can be reduced to $\begin{bmatrix} A' \\ \mathbf{0} \end{bmatrix}$ without moving $\mathbf{w}$" or
  "$A^T\mathbf{x} = \mathbf{w}^T$ is consistent"

  1. Elementary row operations create linear combination of rows.
  2. There is a linear combination of $\mathbf{w}$ and the rows of A which results in a zero vector $\mathbf{0}$. (Why?)
  3. $\mathbf{w}$ is a linear combination of the rows of A.

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

- 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.207/p.199.)

Subspaces Associated with Matrices: Null Spaces

- Is the set of solutions of a homogeneous 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 $A\mathbf{x} = \mathbf{0}$. Then N is a subspace of $\mathbb{R}^n$.

- What is 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 $A\mathbf{x} = \mathbf{0}$. It is denoted by $\mathrm{null}(A)$.

- A.k.a. *kernel*

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 $A\mathbf{x} = \mathbf{b}$, exactly one of the following is true:

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

- Can be proved using the fact that the null space of a matrix is a subspace.
- Except $\{\mathbf{0}\}$, there are infinitely many vectors in 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.

- A basis is a *maximal independent set* and a *minimal spanning set*. (Why?)
  - What happens if we *add* a vector to a basis?
  - What happens if we *remove* a vector from a basis?
- Example: $\mathbf{e}_1, \cdots, \mathbf{e}_n \in \mathbb{R}^n \rightarrow$ **standard basis**
- For a subspace, how many bases are there?

Finding a Basis for $\mathrm{row}(A)$

Let U be a row echelon form of A.

1. By Theorem 3.20, $\mathrm{row}(A) = \mathrm{row}(U)$.
2. Apparently, the nonzero rows of U span $\mathrm{row}(U)$ hence $\mathrm{row}(A)$.
3. In addition, the nonzero rows of U are linearly independent. (Why?)
4. Therefore, the set of the nonzero rows of U are a basis of $\mathrm{row}(U)$ hence $\mathrm{row}(A)$.

- How to know if the basis is correct?
  - Can all the rows of A be represented as a linear combination of the vectors in the basis?
  - Are they linearly independent?
- # of vectors in a basis of $\mathrm{row}(A)$ = # of leading variables of A
- Example 3.45

Finding a Basis for $\mathrm{col}(A)$

Let U be a row echelon form of A.

1. $A\mathbf{x} = \mathbf{0}$ and $U\mathbf{x} = \mathbf{0}$ have the same solution. I.e., $\mathrm{null}(A) = \mathrm{null}(U)$ (Why?)
2. If $U\mathbf{x} = \mathbf{0}$ has a nontrivial solution, any *non-pivot* column of U is a linear combination of the *pivot* columns U. (Why?)
   i The non-pivot columns correspond to *free variables*, therefore we can set any value for those variables.
   ii Assign 1 to one of the non-pivot columns and 0 to rest of them. Then that column can be represented by a linear combination of pivot columns. (Example 3.47)
3. Therefore, we do not need the non-pivot columns to span $\mathrm{col}(U)$.

Finding a Basis for $\mathrm{col}(A)$ (cont'd)

4. The pivot columns of $\mathrm{col}(U)$ are linearly independent. (Why?)
5. Therefore, the pivot columns of U are a basis of $\mathrm{col}(U)$.
   - All the columns of U can be represented by a l.c. of the pivot columns of U.
   - The pivot columns of U are linearly independent.
6. Since the columns of A have the same *dependence relation* as U, the set of the columns of A corresponding to the pivot columns of U is a basis of $\mathrm{col}(A)$.
   - A solution of $B\mathbf{x} = \mathbf{0}$ represents a dependence relation of the columns of B. (Why?)

- # of vectors in a basis of $\mathrm{col}(A)$ = # of leading variables of A
- Example 3.47

Finding a Basis for $\mathrm{null}(A)$

Let R be the **reduced** row echelon form of A.

1. $A\mathbf{x} = \mathbf{0}$ and $R\mathbf{x} = \mathbf{0}$ have the same solution.
2. From $R\mathbf{x} = \mathbf{0}$, any leading variable can be expressed as a linear combination of free variables.
3. Therefore, the solution can be expressed as a linear combination of (column) vectors where the coefficients are the free variables.
4. Since those vectors are linearly independent, (Why?) they form a basis of $\mathrm{null}(A)$.

- Why do we need a reduced row echelon form, not a row echelon form?
- # of vectors in a basis of $\mathrm{null}(A)$ = # of free variables of A
- Example 3.48

Finding a Basis for a Subspace (Summary)

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 $R\mathbf{x} = \mathbf{0}$ in terms of the free variables, set the free variables equal to parameters, substitute back into $\mathbf{x}$, 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.208/p.200)

Dimension

- How many direction vectors do we need to define a line/plane?
- 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 do we call the number?

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 do we call $\dim(\text{row}(A))$ or $\dim(\text{col}(A))$?

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.78/p.75?
- 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 $A\mathbf{x} = \mathbf{0}$
- Number of free variables in the solution of $A\mathbf{x} = \mathbf{0}$

All the above are the same. Why?

- See Theorem 2.2

  $\rightarrow$ 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. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$.

c. $A\mathbf{x} = \mathbf{0}$ 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

- A set of n vectors is a basis for $\mathbb{R}^n$ either one of the condition (p.204/p.196) of a basis is satisfied. (Why?)
- Example 3.52

Theorem 3.28

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

a. $\operatorname{rank}(A^TA) = \operatorname{rank}(A)$

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

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

Coordinates

- Given direction vectors, in how many ways can we represent a vector in a line/plane as a linear combination of them?
- 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 $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$ be a basis for S. For every vector $\mathbf{v}$ in S, there is exactly one way to write $\mathbf{v}$ as a linear combination of the basis vectors in $\mathcal{B}$:

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

- Once a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ for S is fixed, any vector $\mathbf{v}$ in S can be represented *uniquely* by the coefficients $c_1, \ldots, c_k$.

Coordinates (cont'd)

- What do we call the "way" (coefficients of unique linear combination for $\mathbf{v}$)?

Definition: Coordinates

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

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

is called the **coordinate vector of $\mathbf{v}$ 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: "An $m \times n$ matrix A **transforms** a column vector in $\mathbb{R}^n$ into another column vector in $\mathbb{R}^m$."

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

  - **transformation**, **mapping** or **function** T_A defined by the matrix A
  - **domain** (정의역定義域) of T_A: $\mathbb{R}^n$
  - **codomain** (공역共域) of T_A: $\mathbb{R}^m$
  - **image** of $\mathbf{x} \in \mathbb{R}^n$ under (the action of) T_A: $T_A(\mathbf{x}) = A\mathbf{x}$
  - **range** (치역値域) of T_A:
    $\{\mathbf{y} \in \mathbb{R}^m | \mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n\} = \mathrm{col}(A)$ (Excercise 54)

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

Remark

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

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2)$$

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

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

Linear Transformations (cont'd)

- Are all the matrix transformations 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

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

is a linear transformation.

- Examples: Example 3.56 (reflection), 3.57 (rotation)

Linear Transformations (cont'd)

- 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

$$A = \left[\begin{array}{c:c:c:c} T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{array} \right]$$

- A: "**standard matrix of the linear transformation** T"
- Examples: Example 3.58 (rotation), 3.59 (projection)

Linear Transformations (cont'd)

- Notation
  - T_A denotes the linear (matrix) transformation defined by the matrix A.
  - $[T]$ denotes the standard matrix of a linear transformation T.

  $\rightarrow [T_A] = A$ and $T_{[T]} = T$ (p.221)
- 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...?
- Non-linear transformations
  $\rightarrow$ Excercises 7-10 (p.222)

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 $S : \mathbb{R}^n \to \mathbb{R}^p$

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

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 $I_n(\mathbf{v}) = \mathbf{v}$ for every $\mathbf{v}$ 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."
  $\to$ "The (standard) matrix of the inverse (transformation) is the inverse (matrix) of the (standard) matrix (of the transformation)."

Proving 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 → More on "Computer Graphics" course!
- Markov chains
- Population growth
- Graphs and Digraphs
- Error-correcting codes

Markov Chain

- Represents an evolving process consisting of a finite number of states.
- At each step, the process may be in one of the states.
- At the next step, the process can remain in its present state or switch to one of the other states.
- The state changes based on the *transition probability* that depends *only* on the present state and not on the past history of the process.
- Every Markov chain has a unique steady state vector. (Chap 4)

Markov Chain (cont'd)

- "State vectors & transition matrix" / "Probability vectors & stochastic matrix"

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots$$

- Deterministic nature

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{for } k = 0, 1, 2, \ldots$$

- $(P^k)_{ij}$ is the probability of moving from state j to state i in k transitions.

Population Growth

- "Leslie model" by P.H.Leslie (1945)
- Describes the growth of the female portion of a population.
- Every female is assumed to have a maximum lifespan.
- The females are divided equally into age classes.
- *Leslie matrix*: Defined by birthrates and survival probabilities of each class.
- The proportion of the population in each class is approaching a steady state. (Chap 4)

Graphs and Digraphs

- A graph consists of a finite set of *vertices* and *edges*.
- A graph can be described by an *adjacency matrix*.
- *Path*, *length* of a path, k-*path*, *circuit* (closed path), *simple* path
- *Digraph*: a graph with directed edges