Review Session 1 - Linear Algebra

Genghis Kehan

September 11, 2023

1 Linear System

Fundamental Question: How to solve a linear system?

The fundamental question in Linear Algebra is how to solve the linear system $A\vec{x}=\vec{b}$. In other words,

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array} \qquad (*)$$

Where

$$A_{m,n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

The most fundamental way to solve this linear system is by **elementary row operations**, which is also called **Gaussian Elimination**:

!! Please remember it is the only way for you to solve linear systems in the Linear Algebra course.

1. First, we find the **augmented matrix** of the linear system $(*)$:

$$A^*_{m,n+1} \triangleq (A|b) = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

2. Second, we do elementary row operations to reduce the augemented matrix $A^*_{m,n+1}$ into the **Reduced Echelon Form**. Then based on the reduced echelon form, we determine the free variables and find the solution(s) (or no solution). There are several cases:

I. If m≠n (in other words, A is not a sqaure matrix)

Thinking Point: If m≠n, the linear system $(*)$ can not have a unique solution. Either it has at least 1 free variable, or it is inconsistent.

For $A^r_{m,n+1}$(m≠n), either the leading entry 1 of every row is not in the perfect order to form an upper triangle matrix, or there is even no leading entry 1 for some rows (exist all-zero rows).

In such cases, the linear system turns out to be:

a. Situation of misaligned leading entries

$$A^r_{m,n+1} = \left[\begin{array}{ccccc|c} 1 & * & 0 & 0 & 0 & b_1' \\ 0 & 0 & 1 & 0 & 0 & b_2' \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_m' \end{array}\right]$$

Thinking Point: Here specifically, m=n-1, and x_2 is the free variable.

b. Situation of existing all-zero rows

$$A^r_{m,n+1} = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1' \\ 0 & 1 & \cdots & 0 & b_2' \\ \vdots & \vdots & \ddots 1 & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & b_m' \end{array}\right]$$

Thinking Point: Here if b_m' = 0, then x_n is the free variable; if $b_m' \neq$ 0, then $(*)$ **is inconsistent**.

c. Situation of the mixture

$$A^r_{m,n+1} = \left[\begin{array}{ccccc|c} 1 & * & 0 & 0 & 0 & b_1' \\ 0 & 0 & 1 & 0 & 0 & b_2' \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b_m' \end{array}\right]$$

II. If m=n (in other words, A is a square matrix)

In this situation, $A^r_{n,n+1}$ can still have the problem of misalignment, existing all-zero rows, as well as the mixture.

Thinking Point: For A is square, if one misalignment exists, then at least one existing all-zero row exists.

However, $A^r_{n,n+1}$ is possible to be in this perfect reduced echelon form:

$$A^r_{n,n+1} = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1' \\ 0 & 1 & \cdots & 0 & b_2' \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_n' \end{array}\right]$$

In this situation, the linear system $(*)$ is consistent. Moreover, it has a unique solution:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix}$$

2 Linear Combinations and Spanning

Solving a linear system $(*)$ has some geometric meanings.

First, let's rewrite $A_{m,n}$:

$$A_{m,n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \triangleq [\vec{A}_1, \quad \vec{A}_2, \quad \cdots \quad \vec{A}_n]$$

where

$$\vec{A}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

Then we can rewrite the linear system $(*)$ as:

$$span\{\vec{A}_1, \vec{A}_2, \cdots \vec{A}_n\} = \vec{A}_1 x_1 + \vec{A}_2 x_2 + \cdots + \vec{A}_n x_n = \vec{b}$$

Now we can say:

If $\vec{b}$ lies on the span of columns of the coefficient matrix A, then the linear system $(*)$ has a solution, in other words, $(*)$ is consistent.

3 Exercises

1. Solve the following linear systems:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 + 2x_2 + 3x_3 &= 2 \\ 6x_1 + 7x_2 + 8x_3 &= 3 \end{aligned} \quad (1)$$

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 7x_2 + 3x_3 &= 5 \\ 3x_1 + 5x_2 + 6x_3 &= 0 \end{aligned} \quad (2)$$

$$\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ 2x_1 + 4x_2 + 2x_3 + 5x_4 &= 2 \\ 3x_1 + 5x_2 + 6x_3 &= 3 \\ 2x_1 + 4x_2 + 2x_3 + x_4 &= 2 \end{aligned} \tag{3}$$

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

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 4x_2 + 3x_3 &= 5 \end{aligned} \tag{5}$$

2. Determine the values(s) of h such that the matrix is the augmented matrix of a consistent linear system.

(a) $\begin{bmatrix} 2 & 3 & h \\ 4 & 6 & 7 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & -4 & -3 \\ 6 & h & 9 \end{bmatrix}$

3. Suppose a system is a linear equation has a 3×5 augmented matrix whose fifth column is a pivot column. Is the system consistent? Why (or why not)?

4. What would you have to know about the pivot columns in an reduced echelon form in order to know that the linear system is consistent and has a unique solution?

5. A system of linear equations with fewer equation then unknowns is sometimes called an underdetermined system.

Suppose that such a system happens to be consistent. Explain, using pivots, why there must be an infinite number of solutions.

6. A system of linear equations with more equations than unknowns is sometimes called an overdetermined system. Can such a system be consistent? Illustrate your answer with a specific system of three equations in two variables where there is exactly one solution.

p.s. At the end, there is a Youtube video series I would love to recommend you! It visualizes the linear transformation by a matrix and provides you with a geometric sense on Linear Algebra. Feel free to check it! Essence of Linear Algebra

Midterm Review Session

Yuhan Liu

October 17, 2023

1 The Matrix Equation $A\mathbf{x} = \mathbf{b}$

1.1 Theorem 5.1——How to view the matrix equation?

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

$$A\mathbf{x} = \mathbf{b} \tag{4}$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{5}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{6}$$

Implication: Matrix equation = Vector equation (Matrix-vector product) = System of Linear equations

1.2 Theorem 5.3——How to determine the existence of solutions?

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, 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.

Note: **A is a coefficient matrix, not an augmented matrix.** If an augmented matrix $[A \quad \mathbf{b}]$ has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ may or may not be consistent.

Recipe: Row reduce A and use (d) to deduct the conclusions in (a), (b), (c).

1. Could a set of three vectors in $\mathbb{R}^4$ span all of $\mathbb{R}^4$? Justify. (Lecture 5)
2. Let $A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \\ -2 & 1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. Show that the equation $A\mathbf{x} = \mathbf{b}$ does not have a solution for all possible $\mathbf{b}$, and describe the set of all $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ does have a solution.

2 Linear Independence

Homogeneous equations always has a trivial solution ($\mathbf{x} = \mathbf{0}$).

2.1 Parametric vector form

Recipe:

WRITING A SOLUTION SET (OF A CONSISTENT SYSTEM) IN PARAMETRIC VECTOR FORM

1. Row reduce the augmented matrix to reduced echelon form.
2. Express each basic variable in terms of any free variables appearing in an equation.
3. Write a typical solution $\mathbf{x}$ as a vector whose entries depend on the free variables, if any.
4. Decompose $\mathbf{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

1. Write the solution set of the given homogeneous system in parametric vector form. (Lecture 6)

$$\begin{cases} -3x_1 + 5x_2 + 7x_3 = 0 \\ -6x_1 + 7x_2 + x_3 = 0 \end{cases} \tag{1}$$

2.2 Definitions of linear independence and linear dependence

An indexed set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{2}$$

Note: Equation (2) is called a **linear dependence relation** among $\mathbf{V_1}, \ldots, \mathbf{V_p}$ when the weights are not all zero. **An indexed set is linearly dependent if and only if it is not linearly independent.**

2.3 Proposition 7.2 ——How to determine?

The columns of a matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has *only* the trivial solution. (3)

Recipe:

1. Row reduce the augmented matrix to REF.
2. Identify if there is a free variable.
3. (Yes) At least a nonzero solution; (No) Only a trivial solution of $A\mathbf{x} = \mathbf{0}$.
4. Use Proposition 7.2.

2.4 Theorem 7.5

Characterization of Linearly Dependent Sets

An indexed set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $\mathbf{v}_1 \neq \mathbf{0}$, then some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

Note: It does not say that **every vector** in a linearly dependent set is a linear combination of the preceding vectors. A vector in a linearly dependent set may fail to be a linear combination of the other vectors.

2.5 Theorem 7.6——When columns > rows

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

1. Determine if the columns of the matrix form a linearly independent set. Justify. (Lecture 7 exercise 3)

$$\begin{bmatrix} 1 & 4 & -3 & 0 \\ -2 & -7 & 5 & 1 \\ -4 & -5 & 7 & 5 \end{bmatrix}$$

3 Linear Transformation

3.1 Transformation

A transformation (or function or mapping) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$.

- The set $\mathbb{R}^n$ is called the domain of T, and $\mathbb{R}^m$ is called the codomain of T.
- The vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the image of $\mathbf{x}$ (under the action of T).
- The set of all images $T(\mathbf{x})$ is called the range of T.

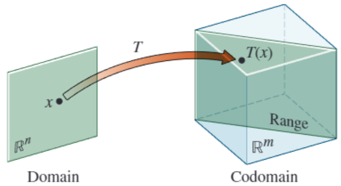


FIGURE 2 Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

3.2 Linear Transformation

Definition:

A transformation (or mapping) T is **linear** if

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;

(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

Properties: Proposition 8.1

If T is a linear transformation, then

$$T(\mathbf{0}) = \mathbf{0} \tag{3}$$

and

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{4}$$

for all vectors $\mathbf{u}$, $\mathbf{v}$ in the domain of T and all scalars c, d.

Note: **Always used to solve proof problems.**

1. Suppose vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$ span $\mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Suppose $T(\mathbf{v}_i) = \mathbf{0}$ for all $1 < i < p$. Show that T is the zero transformation. That is, show that if $\mathbf{x}$ is any vector in $\mathbb{R}^n$, then $T(\mathbf{x}) = \mathbf{0}$. (Lecture 8 Exercise 9)
2. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_3\}$ be a linearly dependent set in $\mathbb{R}^n$. Explain why the set $\{T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_3)\}$ is linearly dependent. (Lecture 8 Exercise 13)

3.3 The Matrix of a Linear Transformation

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that

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

In fact, A is the $m \times n$ matrix whose jth column is the vector $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth column of the identity matrix in $\mathbb{R}^n$:

$$A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \tag{3}$$

PROOF Write $\mathbf{x} = I_n\mathbf{x} = [\, \mathbf{e}_1 \ \cdots \ \mathbf{e}_n \,]\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, and use the linearity of T to compute

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) \\ &= \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x} \end{aligned}$$

The uniqueness of A is treated in Exercise 41. ■

Note: **A is the standard matrix of the Linear Transformation.**

1. Lecture 8 Exercise 15

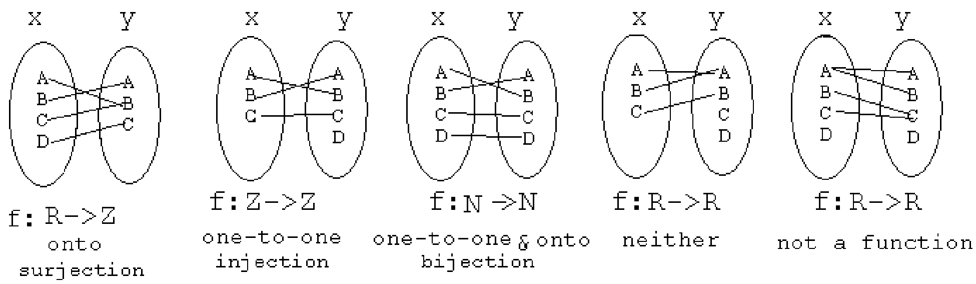
3.4 Onto VS One-to-one

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$.

Range = Codomain (The equation always has a solution.)

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most one* $\mathbf{x}$ in $\mathbb{R}^n$.

If T(u)=T(v), u = v (The equation has either a unique solution or none.)



Theorem 8.4:

> Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let A be the standard matrix for T. Then:
>
> a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;
>
> b. T is one-to-one if and only if the columns of A are linearly independent.

Implications:

1. (a) - pivots every row(Theorem 5.3); (b) - pivots every column(Proposition 7.2)

2. A is a square matrix and has n pivots to satisfy both onto and one-to-one requirements.

1. Lecture 8 Exercise 18
2. Lecture 8 Exercise 19

4 Matrix Operations

4.1 Matrix Multiplication

Definition: If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$. That is,

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

Row-column Rule:

ROW–COLUMN RULE FOR COMPUTING AB

If the product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

EXAMPLE 6 Find the entries in the second row of AB, where

$$A = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

SOLUTION By the row–column rule, the entries of the second row of AB come from row 2 of A (and the columns of B):

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}\begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

Note: The number of columns of A must match the number of rows in B in order for a linear combination such as Ab to be defined. **Also, the definition of AB shows that AB has the same number of rows as A and the same number of columns as B.**

Properties:

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B + C) = AB + AC$ (left distributive law)
c. $(B + C)A = BA + CA$ (right distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r
e. $I_m A = A = A I_n$ (identity for matrix multiplication)

4.2 Transpose of a Matrix: exchange (i,j) with (j,i), rows → columns

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. For any scalar r, $(rA)^T = rA^T$
d. $(AB)^T = B^T A^T$

1. Lecture 9 Exercise 13

5 The Inverse of a Matrix

5.1 Definition

An $n \times n$ matrix A is said to be invertible if there is an $n \times n$ matrix, denoted by A^{-1}, such that $A^{-1}A = I$ and $AA^{-1} = I$, where $I = I_n$ is the $n \times n$ identity matrix.
Theorem 10.1

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

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

Theorem 10.3

a. If A is an invertible matrix, then A^{-1} is invertible and

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

b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,

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

5.2 Finding the Inverse of a Matrix: Theorem 10.5

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 into A^{-1}.

Note: sometimes determine whether the matrix is invertible first.

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

2. Find the inverses of the matrices, if they exists. (Practice midterm question 18(e))
$\begin{bmatrix} a & a & a & a \\ a & a & a & a \\ a & a & a & a \end{bmatrix}$

5.3 Determine a Invertible Matrix: Theorem 10.7

The Invertible Matrix Theorem

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

	Singular matrix A	Non-singular matrix A
Determinant	det A $= 0$	det A $\neq 0$
Invertible Matrix	Not invertible	Invertible
$Ax = 0$	Many solutions (containing free variables)	Only trivial solution
$Ax = b$	Many solutions (containing free variables)	Unique solution
If A equivalent to I	No, may have an entire row/column equal to 0	Yes
Relationship between A's column vectors	Linearly dependent	Linear independent

1. Can a square matrix with 2 identical columns be invertible? Why or why not? (Lecture 10 Exercise 14)

2. Use determinants to decide if the set of vectors is linearly independent. (Practice midterm question 22)

6 Determinants

6.1 Definition

For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm a_{1j} \det A_{1j}$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \ldots, a_{1n}$ are from the first row of A. In symbols,

$$\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}$$
$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j}$$

Cofactor expansion across the first row of A:
Given $A = [a_{ij}]$, the (i, j)-cofactor of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

Then,

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}$$

6.2 Properties

Theorem 11.2: (Triangular: all entries below the main diagonal is 0.)

If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

Theorem 11.3: **"row" can be replaced by "column"**

Row Operations

Let A be a square matrix.

a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.
b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.
c. If one row of A is multiplied by k to produce B, then $\det B = k \det A$.

Proposition 11.4: $\det A = \begin{cases} (-1)^r \begin{pmatrix} \text{product of} \\ \text{pivots in } U \end{pmatrix} & \text{when } A \text{ is invertible} \\ 0 & \text{when } A \text{ is not invertible} \end{cases}$

Proposition 11.5: A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 11.6: If A is an $n \times n$ matrix, then $\det A^T = \det A$.

Theorem 11.7: **Multiplicative Property** If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

EXAMPLE 2 Compute $\det A$, where $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

SOLUTION To simplify the arithmetic, we want a 1 in the upper-left corner. We could interchange rows 1 and 4. Instead, we factor out 2 from the top row, and then proceed with row replacements in the first column:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Next, we could factor out another 2 from row 3 or use the 3 in the second column as a pivot. We choose the latter operation, adding 4 times row 2 to row 3:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Finally, adding $-1/2$ times row 3 to row 4, and computing the "triangular" determinant, we find that

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2\,(1)(3)(-6)(1) = -36$$

■

6.3 Two ways of computing determinant

1) Using cofactor expansion
2) Row reduce to echelon form and use theorem 11.2

6.4 Cramer's Rule

Cramer's Rule

Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, \ldots, n \tag{1}$$

6.5 Determinants as Area/Volume

Theorem 12.4

If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|\det A|$. If A is a 3×3 matrix, the volume of the parallelepiped determined by the columns of A is $|\det A|$.

1. Lecture 12 Exercise 16

Theorem 12.5/6

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation determined by a 2×2 matrix A. If S is a parallelogram in $\mathbb{R}^2$, then

$$\{\text{area of } T(S)\} = |\det A| \cdot \{\text{area of } S\} \tag{5}$$

If T is determined by a 3×3 matrix A, and if S is a parallelepiped in $\mathbb{R}^3$, then

$$\{\text{volume of } T(S)\} = |\det A| \cdot \{\text{volume of } S\} \tag{6}$$

The conclusions of Theorem 10 hold whenever S is a region in $\mathbb{R}^2$ with finite area or a region in $\mathbb{R}^3$ with finite volume.

References

Linear Algebra Final Review (Lec 13-27)

Genghis Luo, Tina Liu

December 8, 2023

1 Vector Space and Subspace

1.1 Vector Space

Definition: A vector space is a **nonempty** set V of **vectors**, on which are defined two operations, called **addition** and **multiplication** by scalars, subject to ten axioms.

Example: Polynomial vector space (The set P_n of polynomials of degree at most n):

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

1.2 Subspace

Difference between subset and subspace: A subspace not only needs to be a subset of the original vector space but also must satisfy the ten axioms of a vector space.

A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:

a. The zero vector is in H.

b. For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.

c. For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

Note:

1. Subset is the premise: R^2 is not a subspace of R^3 because R^2 is not even a subset of R^3.
2. Every subspace is a vector space.

Describe a subspace: Spanning space and set

Theorem 13.2:

If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in a vector space V, then Span$\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is a subspace of V.

1.3 Exercises

1. Show that set W is a subspace of the vector space V below.

$$W = \{\mathbf{p} \in P : \frac{d^2\mathbf{p}(t)}{dt^2} = \mathbf{0}\}, \quad V = P$$

2. Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

, is $\mathbf{w}$ in the subspace spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

2 Null Space, Column Space, Row Space

For an $m \times n$ matrix A,

The null space is the set of all solutions of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Column/Row space is the span of its columns/rows.

Note:

1. If matrix A represents a function, Col$A$ is the range of the function.
2. Determine whether a linear system has a solution: whether $\mathbf{b}$ is in Col$A$.
3. ColA^T = Row$A$.

2.1 Relations between Concepts

Contrast between Nul$A$ and Col$A$:

Nul A	**Col A**
1. Nul A is a subspace of $\mathbb{R}^n$.	**1.** Col A is a subspace of $\mathbb{R}^m$.
2. Nul A is implicitly defined; that is, you are given only a condition ($A\mathbf{x} = \mathbf{0}$) that vectors in Nul A must satisfy.	**2.** Col A is explicitly defined; that is, you are told how to build vectors in Col A.
3. It takes time to find vectors in Nul A. Row operations on $[\ A \quad \mathbf{0}\ ]$ are required.	**3.** It is easy to find vectors in Col A. The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A.	**4.** There is an obvious relation between Col A and the entries in A, since each column of A is in Col A.
5. A typical vector $\mathbf{v}$ in Nul A has the property that $A\mathbf{v} = \mathbf{0}$.	**5.** A typical vector $\mathbf{v}$ in Col A has the property that the equation $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector $\mathbf{v}$, it is easy to tell if $\mathbf{v}$ is in Nul A. Just compute $A\mathbf{v}$.	**6.** Given a specific vector $\mathbf{v}$, it may take time to tell if $\mathbf{v}$ is in Col A. Row operations on $[\ A \quad \mathbf{v}\ ]$ are required.
7. Nul $A = \{\mathbf{0}\}$ if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.	**7.** Col $A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.	**8.** Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$.

2.2 Generalization: Kernel/null space and range of a linear transformation

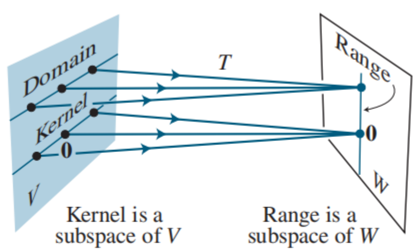


FIGURE 2 Subspaces associated with a linear transformation.

2.3 Exercises

1. Find an explicit description of Col A and Nul A.

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

2. $T : M_{2\times 2} \to \mathbb{R}^2, \quad T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & d \\ 0 & 0 \end{bmatrix}$

   (a) Show that T is a linear transformation.

   (b) Find a 2×2 matrix A in $M_{2\times 2}$ that spans the kernel of T, and describe the range of T.

3 Basis and Dimension

Definition: Let H be a subspace of a vector space V. A set of vectors $\mathcal{B}$ in V is a basis for H if:

1. The set is linearly independent.
2. $H = \text{Span}\mathcal{B}$

(Simplified: Let V be a nonzero subspace of $\mathbb{R}^n$, a set of linearly independent vectors that span V is called a basis.)

3.1 Features of Basis

1. A basis is the smallest spanning set.
2. A basis is the largest independent set in the subspace.

Theorem 18.4 The basis theorem:

> **The Basis Theorem**
>
> Let H be a p-dimensional subspace of $\mathbb{R}^n$. Any linearly independent set of exactly p elements in H is automatically a basis for H. Also, any set of p elements of H that spans H is automatically a basis for H.

3. **Theorem 18.2**

> If a vector space V has a basis of n vectors, then every basis of V must consist of exactly n vectors.

4. The number of vectors in the basis of subspace V is called the **dimension** of V.

3.2 Basis for Nul$A$, Col$A$ and Row$A$

Recipes (lecture 16)

Summary: For an $m \times n$ matrix A,

	Dimension	Basis
Col$A$	Rank$A$	The pivot columns of$A$
Nul$A$	Nullity A = n - Rank A	The vectors in the parametric form of the solution of $A\mathbf{x} = \mathbf{0}$
Row$A$	Rank$A$	The nonzero rows of the RREF of$A$

Rank = dim Col A = number of pivot columns = number of pivot rows = dim Row A
Nullity = dim Nul A = number of free variables
Rank A + Nullity A = dim V (Theorem 18.5)

Original matrix A VS its RREF U:

For any REF U of a matrix A,

Row U = Row A (because elementary row operations are linear combinations of rows), Col $U \neq$ Col A (e.g. elimination results in a row of zeros at the bottom of the matrix, causing the column space to decrease), and rank remains the same.

3.3 The invertible matrix theorem (continued)

The Invertible Matrix Theorem (continued)

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

m. The columns of A form a basis of $\mathbb{R}^n$.

n. Col $A = \mathbb{R}^n$

o. rank $A = n$

p. dim Nul $A = 0$

q. Nul $A = \{\mathbf{0}\}$

3.4 Exercises

1. Find the basis for the null space of the matrix.

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

2. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$, with $k < n$. Use a theorem from lecture 5 to explain why S cannot be a basis for $\mathbb{R}^n$.

3. Consider the polynomials $\mathbf{p}_1(t) = 4t - t^2$, $\mathbf{p}_2(t) = 4 - t^2$, and $\mathbf{p}_3(t) = 4 - t$. Is $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ a linearly independent set in P_3? Why or why not?

4. Let H be an n-dimensional subspace of an n-dimensional vector space V. Show that $H = V$.

4 Coordinate Systems

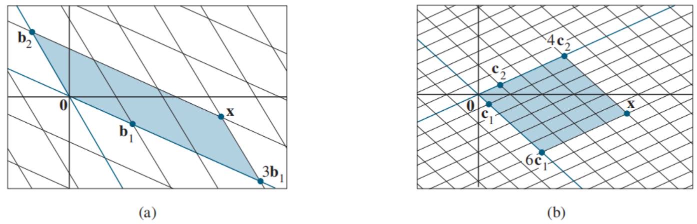


FIGURE 1 Two coordinate systems for the same vector space.

4.1 Represent Vectors

For a vector space, vectors can be uniquely expressed using the standard basis. (Theorem 17.1 The unique representation theorem)

B-coordinates of a vector $\mathbf{x}$ ($[\mathbf{x}]_{\mathcal{B}}$): given a vector space V with basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, it represent the unique set of coefficients such that:

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n$$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Coordinate mapping:

$$[\cdot]_{\mathcal{B}} : V \to \mathbb{R}^n, \quad \mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$$

Questions:

1. Find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ and the given basis $\mathcal{B}$.

Hint: This vector can be represented as a linear combination of the basis vectors, with the coefficients being its coordinates.

$x = [b_1\, b_2\, \ldots][\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$ (matrix-vector product)

2. Find the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of $\mathbf{x}$ relative to the given basis $\mathcal{B}$.

$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}\mathbf{x}$.

4.2 Generalization: Change of Basis

$P_{\mathcal{B}}$ ($P_{\mathcal{E}\leftarrow\mathcal{B}}$): the change-of-coordinates matrix from $\mathcal{B}$ to the standard matrix in $\mathbb{R}^n$.
$P_{\mathcal{C}\leftarrow\mathcal{B}}$: transforms the coordinates from the basis $\mathcal{B}$ to $\mathcal{C}$.

Theorem 17.3:

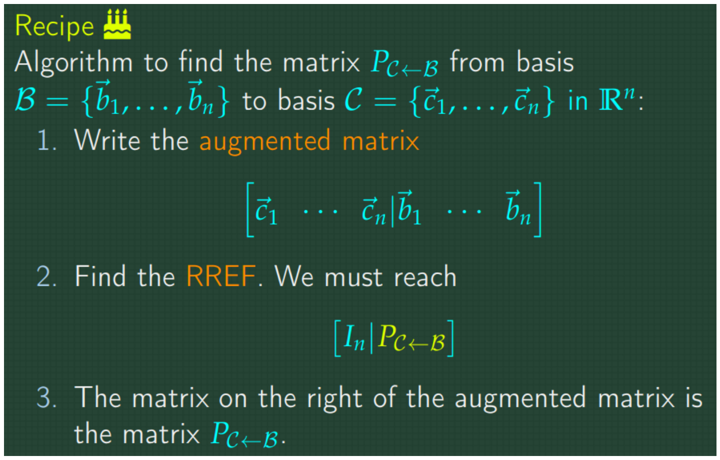
Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be bases of a vector space V. Then there is a unique $n \times n$ matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ such that

$$[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}} \qquad (4)$$

The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ are the $\mathcal{C}$-coordinate vectors of the vectors in the basis $\mathcal{B}$. That is,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \qquad (5)$$

Recipe: Find the change-of-coordinates matrix



4.3 Exercises

1. Consider the following set of polynomial functions

$$B = \{4 + t + 2t^2, 2t - t^2, 1 + 2t^3, t^3\}$$

   (a) Show that these polynomials form a basis of P_3.
   (b) Let $\mathbf{p}(t) = 1 + t^2$. Find the coordinate vector of $\mathbf{p}$ relative to B.

5 Eigenvalues, Eigenvectors and Eigenspaces

5.1 Computation

Step 1: Find the eigenvalues λ

$$\det(\text{A-}\lambda\text{I}) = 0$$

Step 2: Find the eigenvectors $\vec{v}$ corresponding to each eigenvalue λ_i

$$\text{A}\vec{v} = \lambda_i \vec{v}$$

Step 3: Find an eigenvector basis $\{\vec{v_1}, \cdots, \vec{v_k}\}$ (k≥1) for each eigenspace E_{λ_i}

$$E_{\lambda_i} \triangleq \text{Nul(A-}\lambda_i\text{I)} = \text{span}\{\vec{v_1}, \cdots, \vec{v_k}\}$$

p.s. The eigenspace E_{λ_i} of each eigenvalue λ_i is the set of all eigenvectors corresponding to λ_i, denoted as Nul(A-λ_iI)

5.2 Key notes [Theorem 20.3]

1. The geometric multiplicity of the eigenvalue λ_i is the dimension of E_{λ_i}, noted as dim Nul(A-λ_iI)

2. The algebraic multiplicity of the eigenvalue λ_i is the power of $(\lambda - \lambda_i)$ in the linear factorization of the characteristic polynomial

3. The geometric multiplicity of each λ_i is always smaller or equal to the algebraic multiplicity of each λ_i

4. For a n×n matrix A, A is diagonalizable if and only if the geometric multiplicity equals the algebraic multiplicity for every λ_i, (equivalent to "the sum of geometric multiplicity equals n").

5. $B_i \triangleq \{\vec{v_1}, \cdots, \vec{v_k}\}$ is a (eigenvector) basis of the eigenspace E_{λ_i}. Say $\lambda_1, \cdots, \lambda_p$ are eigenvalues of A. A is diagonalizable if and only if the total collection of the (eigenvector) basis for each eigenspace $\bigcup_{i=1}^{p} B_i$ forms an eigenvector basis for $\mathbb{R}^n$

5.3 Exercise

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

1. What are the eigenvalues and corresponding eigenvectors?
2. What are the eigenspaces corresponding to each eigenvalue?
3. What are the algebraic and geometric multiplicity of each eigenvalue?
4. Is this matrix diagonalizable?

6 Diagonalization

6.1 Diagonalizability

Theorem: For an n×n matrix A, A is diagonalizable if and only if it has n linearly independent eigenvectors.

p.s. When A has n linearly independent eigenvectors, considering the key note 5 for Theorem 20.3, these eigenvectors form an eigenvector basis for $\mathbb{R}^n$; considering the key note 4 for Theorem 20.3, the geometric multiplicity equals to the algebraic multiplicity.

Corollary: Every n×n matrix with n distinct eigenvalues is diagonalizable.

p.s. Every eigenvalue has algebraic multiplicity 1. Since the geometric multiplicity must greater or equal to 1, and smaller or equal to the algebraic multiplicity, then it is also 1. Then the algebraic multiplicity of every eigenvalue equals to its geometric multiplicity, equals to 1.

6.2 Computation

If A is diagonalizable, then $A = PDP^{-1}$, where $P = [v_1, \cdots, v_n], D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, and

$v_1, \cdots, v_n$ should be linearly independent eigenvectors corresponding to $\lambda_1, \cdots, \lambda_n$

p.s. Be sure to match each eigenvector column in P with each eigenvalue column in D. It is possible that same eigenvalue appears several times in D, as long as its geometric multiplicity coincides with its algebraic multiplicity which is given by the diagonalizability.

6.3 Generalization

Question 1: What if we have complex eigenvalues instead of real eigenvalues? In other words, what if the characteristic polynomial can not be linearly factorized in $\mathbb{R}$?

Example: The characteristic polynomial of $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is $\lambda^2 + 1 = 0$, which leads to two complex eigenvalues: i and $-i$. It is not diagonalizable on $\mathbb{R}$, but it is diagonalizable on $\mathbb{C}$. $A = PDP^{-1}$ where $P = \begin{bmatrix} -i & i \\ 1 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$

Generalization 1 (Lec 22): From real eigenvalues to complex eigenvalues

Remark 1: Not every matrix is diagonalizable. For instance, $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is not diagonalizable both on $\mathbb{R}$ and $\mathbb{C}$.

Remark 2: For 2×2 real matrix with the complex eigenvalue $a - bi$ and its corresponding complex eigenvector $\vec{v}$ (the other eigenvalue is its conjugation, $a + bi$), there exists a real decomposition akin to but not the same as diagonalization: $A = PCP^{-1}$, where P=[Re$\vec{v}$, Im$\vec{v}$], and C= $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

....

Question 2: Can we learn diagonalizability of linear transformation not only from $\mathbb{R}^n$ to $\mathbb{R}^n$ (matrices), but from general vector space to vector space (operators)?

Example: The linear transformation T: $\mathbb{P}_2 \to \mathbb{P}_2$ given by $T(\vec{p}) = \vec{p}(t) + (t+1)\vec{p}'(t)$ is diagonalizable since it has an eigenvector basis $\{1, 1+t, (1+t)^2\}$ (Recipe: Lec 21, page 33)

Generalization 2 (Lecture 21): From $\mathbb{R}^n$ to general vector space

Remark: $[T]_{B \to C}$ is called the representation matrix of the linear transformation T from $\mathbb{V}$ with the basis $B = \{b_1, \cdots, b_n\}$ to $\mathbb{W}$ with the basis $C = \{c_1, \cdots, c_m\}$. $[T]_{B \to C} = [[T(b_1)]_C, \cdots, [T(b_n)]_C]$, which should be an $m \times n$ matrix. $[T]_B$ is the abbreviation of the representation matrix of the linear transformation T from $\mathbb{V}$ to $\mathbb{V}$ both with the same basis $B = \{b_1, \cdots, b_n\}$, which should be an $n \times n$ square matrix.

$$[T(x)]_C = [T]_{B \to C}[x]_B$$
$$\text{In particular, } [T(x)]_B = [T]_B[x]_B$$

Thinking point: What are the similarities and differences between representation matrix and the change-of-coordinates matrix?

6.4 Exercises

1. Diagonalize $\begin{bmatrix} 3 & 0 & 0 \\ -3 & 4 & 9 \\ 0 & 0 & 3 \end{bmatrix}$

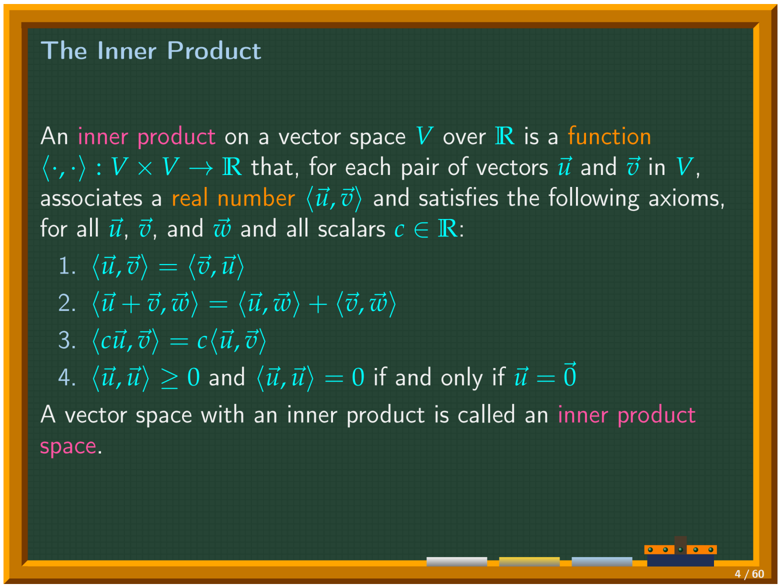
2. Find an invertible matrix P and a matrix C of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ such that the matrix A=$\begin{bmatrix} 2 & 2 \\ -1 & 0 \end{bmatrix}$ has the form $A = PCP^{-1}$.

3. Diagonalize the linear transformation T: $M_{2\times 2} \to M_{2\times 2}$, $T(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = \begin{bmatrix} b & b \\ c & c \end{bmatrix}$

4. Suppose that A is a diagonalizable matrix where all the eigenvalues are real. Prove that the rank of A is the number of nonzero eigenvalues of A, including repetition.

7 Inner Product, Norm and Orthogonality

7.1 Inner Product



p.s. Inner product needs to be understood as an "operator" satisfying the four axioms. Dot product is one of the inner product on $\mathbb{R}^n$.

7.2 Examples for Inner Product (Lec 24)

I. When V=$\mathbb{R}^n$

i. Dot Product: $< \vec{u}, \vec{v} > = \vec{u} \cdot \vec{v} = \vec{u}^\top \vec{v}$

ii. $< \vec{u}, \vec{v} > = au_1v_1 + bu_2v_2$, for $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

II. When V=$\mathbb{P}_n$

i. $< \vec{p}, \vec{q} > = \sum_{i=0}^{n} \vec{p}(t_i)\vec{q}(t_i)$ for fixed $t_0, \cdots, t_n \in \mathbb{R}$

III. When V=$\mathbb{M}_{m\times n}$

i. $< A, B > = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}b_{ij}$

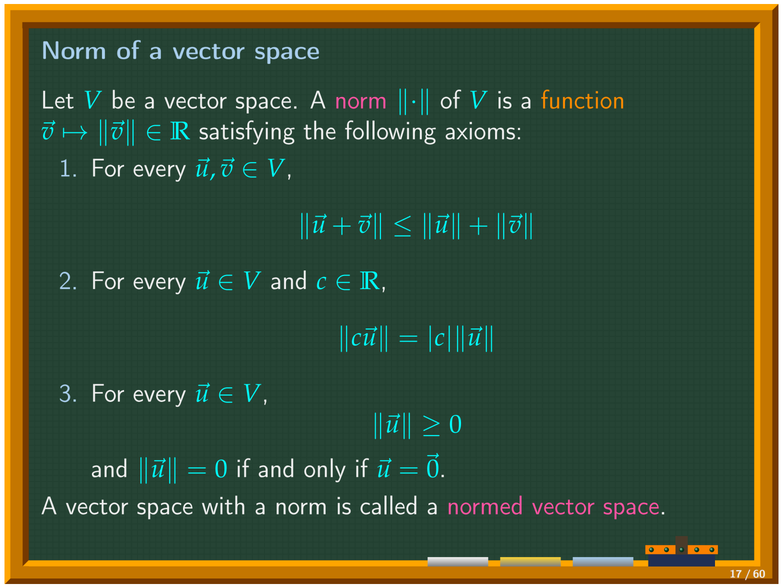
IV. When V $= L^2 = \{f \in \mathbf{F} : \int_a^b f(x)^2 dx < \infty\}$, $\mathbf{F}$ is the set of all functions f: $\mathbb{R} \to \mathbb{R}$

i. $< f, g > = \int_a^b f(x)g(x)dx$

V. When V $= \ell^2 = \{(a_n)_{n\geq 1} : \sum_{n=1}^{\infty} a_n^2 < \infty\}$, $(a_n)_{n\geq 1} = (a_1, a_2, a_3. \cdots)$ is an $\mathbb{R}$ sequence

i. $< (a_n)_{n\geq 1}, (b_n)_{n\geq 1} > = \sum_{n=1}^{\infty} a_n b_n$

7.3 Norm



p.s. Norm needs to be understood as a "function" satisfying the three axioms. The Euclidean magnitude/length of the vector on $\mathbb{R}^n$ defined by the square root of dot product, is one of the norm on $\mathbb{R}^n$.

7.4 Examples for Norm (Lec 24)

$I.$ When V = $\mathbb{R}$

$i.$ $||x|| = |x|$

$I.$ When V = $\mathbb{R}^n$

$i.$ Euclidean Length (ℓ^2 norm): $||\vec{u}|| = \sqrt{\sum_{i=1}^{n} u_i^2} = \sqrt{\vec{u}^\top \vec{u}} = \sqrt{\vec{u} \cdot \vec{u}}$

$ii.$ (ℓ^1 norm): $||\vec{u}|| = \sum_{i=1}^{n} |u_i|$

$ii.$ (ℓ^∞ norm): $||\vec{u}|| = \max_{1 \leq i \leq n} |u_i| = \max\{|u_1|, \cdots, |u_n|\}$

7.5 Relations Between Inner Product Space and Normed Vector Space

Finite Dimensional Vector Space $\subset$ Inner Product Space $\subset$ Normed Vector Space

Theorem 24.1: Finite Dimensional Vector Space $\subset$ Inner Product Space

Every finite dimensional vector space is an inner product space. Moreover, if B is a basis for a vector space V with dimension n, then an inner product for V is

$$< \vec{u}, \vec{v} > = [\vec{u}]_B \cdot [\vec{v}]_B$$

Theorem 24.4: Inner Product Space $\subset$ Normed Vector Space

Every inner product space is a normed vector space. Moreover, if V is an inner product space, the following function

$$||\vec{u}|| = \sqrt{< \vec{u}, \vec{u} >}$$

is a norm in V.

Theorem 24.5: Normed Vector Space + Parallelogram Law $\rightarrow$ Inner Product Space

Let (V, $||\cdot||$) be a normed vector space satisfying the following identity: For every $\vec{u}, \vec{v} \in$ V,

$$||\vec{u}+\vec{v}||^2 + ||\vec{u}-\vec{v}||^2 = 2(||\vec{u}||^2 + ||\vec{v}||^2),$$

then there exists an inner product on V such that $||\vec{v}|| = \sqrt{< \vec{v}, \vec{v} >}$. Moreover, the function

$$< \vec{u}, \vec{v} >= \frac{||\vec{u}+\vec{v}||^2 - ||\vec{u}-\vec{v}||^2}{4}$$

forms an inner product and satisfies the desired condition.

Two Properties on the Inner Product Space

1. Pythagorean Theorem (Thm 24.2)
Let V be an inner product space and $||\vec{u}|| = \sqrt{< \vec{u}, \vec{u} >}$ for all $\vec{u} \in$ V. If $\vec{u}, \vec{v} \in$ V is such that $< \vec{u}, \vec{v} >= 0$, then

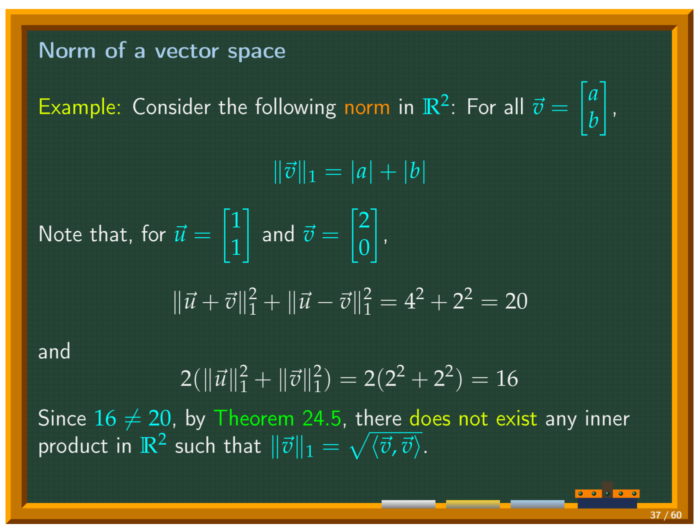
$$||\vec{u} + \vec{v}||^2 = ||\vec{u}||^2 + ||\vec{v}||^2$$

2. Cauchy-Schwarz Inequality (Thm 24.3)
Let V be an inner product norm and $||\vec{u}|| = \sqrt{< \vec{u}, \vec{u} >}$ for all $\vec{u} \in$ V. If $\vec{u}, \vec{v} \in$ V, then

$$| < \vec{u}, \vec{v} > | \leq ||\vec{u}|| \, ||\vec{v}||$$

Equality holds if and only if $\vec{u}$ and $\vec{v}$ are linearly dependent.

7.6 Example of Normed Vector Space but Not Inner Product Space



p.s. We equip $\mathbb{R}^2$ with ℓ^1 norm that does not satisfy Parallelogram Law, then $(\mathbb{R}^2, || \cdot ||_1)$ is a finite dimensional normed vector space which is not an inner product space. If we use ℓ^2 norm, then $(\mathbb{R}^2, || \cdot ||_2)$ is an inner product space.

Thinking point: examples of infinite dimensional inner product space; examples of unnormed vector space

7.7 Orthogonality and Orthonormality

- Orthogonality:
  We say a set of vectors B $= \{\vec{v_1}, \cdots, \vec{v_n}\}$ is **orthogonal** in the inner product space V if $< \vec{v_i}, \vec{v_j} >= 0$ for every $i \neq j$
- Orthonormality:
  We say a set of vectors B $= \{\vec{v_1}, \cdots, \vec{v_n}\}$ is **orthonormal** in the inner product space V if B is orthogonal and $||v_i|| = \sqrt{< \vec{v_i}, \vec{v_i} >} = 1$ for every $i \in \{1, \cdots, n\}$ (An inner product space is a normed vector space)
- Orthogonal/Orthonormal Basis:
  A basis that is an orthogonal/orthonormal set
- Orthogonal Complement

Let W be a subspace of V. The orthogonal complement of W, denoted by $\underline{W^{\perp} = \{\vec{z} \in V : <\vec{z}, \vec{w}> = 0 \text{ for all } \vec{w} \in W\}}$, is the set of all vectors $\vec{z}$ in V that are orthogonal to every vector $\vec{w}$ in W.

• Theorem 23.4: Connection Between Row Space, Column Space and Null Space of the Matrix ($\mathbb{R}^n$ with dot product)

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

$$(RowA)^{\perp} = NulA$$
$$(ColA)^{\perp} = NulA^T$$

p.s. The geometric interpretation of row space is the largest subspace that makes A, as a linear transformation, injective.

• Theorem 26.1: An m $\times$ n matrix U has orthonormal columns (the column vectors form an orthonormal set) if and only if $U^T U = I_n$.

• Theorem 26.2: Let U be an m $\times$ n matrix with orthonormal columns, and let $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$. Then

1. $||U\vec{x}|| = ||\vec{x}||$
2. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$
3. $(U\vec{x}) \cdot (U\vec{y}) = 0$ if and only if $\vec{x} \cdot \vec{y} = 0$

p.s. U is called orthogonal matrix if it is a square matrix, in other words, m=n.

7.8 Exercises

1. Verify all the examples for inner product and examples for norm
2. Prove the Bessel's inequality: Let V be a finite dimensional inner product space and suppose $\{\vec{b}_1, \cdots, \vec{b}_n\}$ is an orthonormal basis in V. Then, for every $\vec{x} \in$ V, we have

$$\sum_{k=1}^{n} | < \vec{x}, \vec{b}_k > |^2 \leq ||\vec{x}||^2$$

3. Prove Parseval's identity: Let $\{\vec{u}_1, \cdots, \vec{u}_n\}$ be an orthonormal basis of an inner product space V. Show that, for every $\vec{v}, \vec{w} \in$ V:

$$< \vec{v}, \vec{w} > = \sum_{i=1}^{n} < \vec{v}, \vec{u}_i >< \vec{w}, \vec{u}_i >$$

8 Orthogonal Projection and the Gram-Schmidt Process

We only consider $\mathbb{R}^n$ with dot product in orthogonal projection as well as the Gram-Schmidt Process in this course. **In fact, all of them can be generalized to inner product space.**

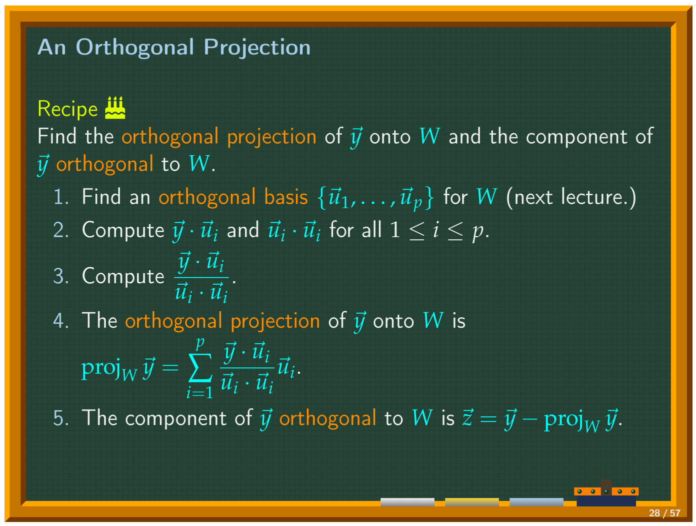
8.1 The Orthogonal Projection

The orthogonal projection of $\vec{y}$ onto W, a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{\vec{u}_1, \cdots, \vec{u}_p\}$, is denoted as

$$\text{proj}_W\vec{y} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \cdots + \frac{\vec{y}\cdot\vec{u}_p}{\vec{u}_p\cdot\vec{u}_p}\vec{u}_p$$

The vector $\vec{z} = \vec{y} - \text{proj}_W\vec{y}$ is called the component of $\vec{y}$ orthogonal to W.

Theorem 25.3 (The Orthogonal Decomposition Theorem) guarantees that the $\text{proj}_W\vec{y} \in$ W and $\vec{z} \in W^\top$ are unique.

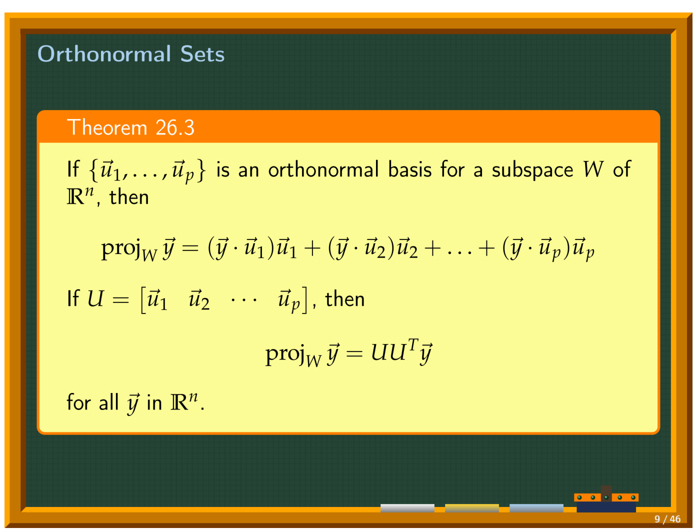


• Theorem 25.4: The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\vec{y}$ be any vector in $\mathbb{R}^n$, and let $\text{proj}_W\vec{y}$ be the orthogonal projection of $\vec{y}$ onto W. Then $\text{proj}_W\vec{y}$ is the closest point in W to $\vec{y}$, in the sense that

$$||\vec{y} - \text{proj}_W\vec{y}|| < ||\vec{y} - \vec{v}||$$

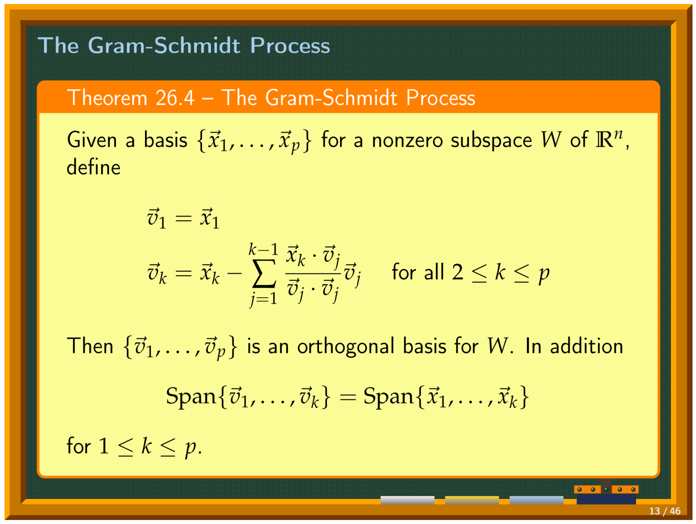
for all $\vec{v}$ in W distinct from $\text{proj}_W\vec{y}$.
p.s. $\text{proj}_W\vec{y}$ is the best approximation.
• Theorem 26.3: An Alternative Way to Calculate Orthogonal Projection



p.s. Normalize the orthogonal basis into orthonormal basis

8.2 The Gram-Schmidt Process

The goal of the Gram-Schmidt Process is to find an orthogonal (and orthonormal) basis for any nonzero subspace of $\mathbb{R}^n$. It is generated by a given basis.

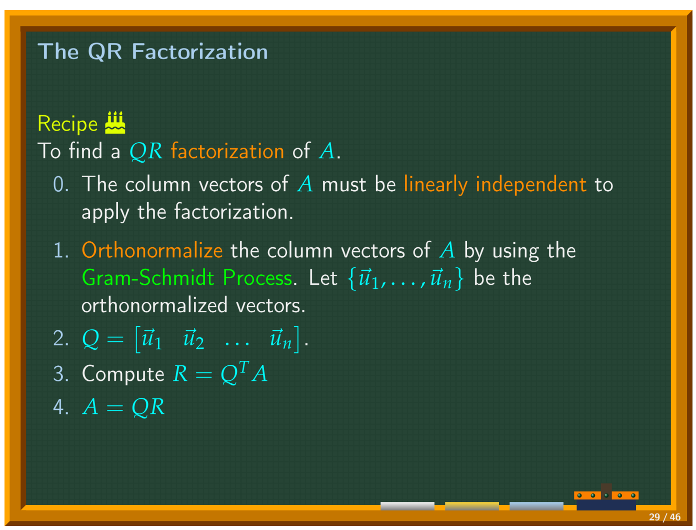


8.3 QR Factorization/Decomposition

It is a natural deduction from the Gram-Schmidt Process.

Theorem 26.6

If A is an m × n matrix with linearly independent columns, then A can be factored as A = QR, where Q is an m × n matrix whose columns form an orthonormal basis for Col A and R is an n × n upper triangular matrix with positive entries on its diagonal.



Example

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix} = QR$$

8.4 Exercises

1. Let $\vec{y} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. Write $\vec{y}$ as the sum of two orthogonal vectors, one in Span$\{\vec{u}\}$ and the other one orthogonal to $\vec{u}$.

2. Find the closest point and the best approximation to $\vec{y}$ in the subspace W spanned by $\vec{v}_1$ and $\vec{v}_2$.

$$\vec{y} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}, \vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix}$$

3. Let $\vec{u}$ and $\vec{v}$ be vectors in $\mathbb{R}^n$. Show that $\text{proj}_{\text{span}\{\vec{v}\}}\, \vec{u} = \vec{0}$ if and only if $\vec{u}$ and $\vec{v}$ are orthogonal.

4. The given set is a basis for a subspace W. Use the Gram-Schmidt Process to produce an orthogonal basis for W.

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

5. Verify the QR-Factorization of

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