

# To-Do List Year 2 Fall Semester

## 1 Chapter 1

### 1.1 System of Linear Equation

A system of linear Equations has

1. no solution, or
2. exactly one solution, or
3. infinitely many solutions

Definition: If there is no solution exist, we say the linear system is inconsistent, otherwise, the linear system is consistent

### 1.2 Row Reduction and Echelon forms

1. Each matrix is row equivalent to one and only one reduced echelon matrix.
2. Columns contain leading 1(pivot position) are pivot columns.
3. From matrix to echelon form is forward phase while from echelon form to reduced echelon form called backward phase.
4. Variables in linear systems can be divided as basic and free variables(free to choose any value).

### 1.3 Vector Equations

Linear Combinations: In general, given vectors  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$  in  $\mathbb{R}^n$ , and given scalars  $C_1, C_2, \dots, C_p$ , the vector  $\vec{w} = C_1\vec{v}_1 + C_2\vec{v}_2 + \dots + C_p\vec{v}_p$  with weights coefficients  $C_1, C_2, \dots, C_p$ .

Definition: If  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$  are vectors in  $\mathbb{R}^n$ , then the set of all linear combinations of  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$  is denoted by  $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  and is called **the subset of  $\mathbb{R}^n$ (or generated) by  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$** .

That is,  $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  is the collection of all vectors that can be written in the form  $\vec{w} = C_1\vec{v}_1 + C_2\vec{v}_2 + \dots + C_p\vec{v}_p$ .

### 1.4 The matrix Equation $A\vec{x} = \vec{b}$

Let  $A$  is an  $m \times n$  matrix, the following statements are equivalent:

1. For each  $\vec{b}$  in  $\mathbb{R}^m$ , the equation  $A\vec{x} = \vec{b}$  has a solution.
2. Each  $\vec{b} \in \mathbb{R}^m$  is a linear combination of the column vectors of  $A$ .
3. The column vectors of  $A$  span  $\mathbb{R}^m$ .
4.  $A$  has a pivot position in every row.

## 1.5 Solution Sets of Linear Systems

Homogenous Linear Systems:

$$A\vec{x} = \vec{0}$$

It always has a trivial solution  $\vec{x} = \vec{0}$ . The system  $A\vec{x} = \vec{0}$  has nontrivial solution if and only if the system has at least one free variable.

Parametric Vector Form:

$$\vec{x} = s\vec{u} + t\vec{v} \text{ where } s, t \in \mathbb{R}$$

Theorem: Suppose  $A\vec{x} = \vec{b}$  is consistent for some given  $\vec{b}$  and let  $\vec{p}$  be a solution. Then the solution set of  $A\vec{x} = \vec{b}$  is the set of all vectors of the form  $\vec{w} = \vec{p} + r\vec{v}_n$ , where  $\vec{v}_n$  is any solution of  $A\vec{x} = \vec{0}$ .

Remarks: Describe the free variables as arbitrary.

A set of vectors  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  in  $\mathbb{R}^n$  is said to be **Linear Independent** if the vector equation  $C_1\vec{v}_1 + C_2\vec{v}_2 + \dots + C_p\vec{v}_p = \vec{0}$  has only the trivial solution.

## 1.6 Linear Independent

The column vectors of a matrix  $A$  are **Linearly Independent** if and only if  $A\vec{x} = \vec{0}$  has only the trivial solution.

Sets of One or Two Vectors:

1.  $\{\vec{v}\}$  is linearly independent if and only if  $\vec{v} \neq \vec{0}$
2.  $\{\vec{v}_1, \vec{v}_2\}$  is linearly independent if and only if neither of the vectors is a multiple of the other  $\iff \{\vec{v}_1, \vec{v}_2\}$  is linearly dependent if at least one of the vectors is a multiple of the other, where the constant can be 0.

Characterization of linearly dependent sets:  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  where  $p \geq 2$  is linearly dependent if and only if at least one of the vectors is a linear combination of the others.

Theorem: If  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  where  $p \geq 2$ ,  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$  are vectors in  $\mathbb{R}^n$ . If  $p > n$ , then  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$  must be **Linearly Dependent**.

## 1.7 Introduction to Linear Transformations

Definition: A transformation from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ ,  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  is a rule that assigns to each vector in  $\mathbb{R}^n$  to a vector  $T(\vec{x})$  in  $\mathbb{R}^m$ .

$\mathbb{R}^n$  : **Domain** of  $T$ .

$\mathbb{R}^m$  : **Codomain** of  $T$ .

For  $\vec{x} \in \mathbb{R}^n$ ,  $T(\vec{x})$  is called the image of  $\vec{x}$ . The set of all images  $T(\vec{x})$  is called the range of  $T$ .

Definition: A transformation (or mapping)  $T$  is **linear** if

1.  $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$  for all  $\vec{u}, \vec{v} \in D(T)$  of  $T$
2.  $T(c\vec{u}) = cT(\vec{u})$  for all scalars  $c$  and  $\vec{u} \in D(T)$ .
3.  $T(\vec{0}) = \vec{0}$

## 1.8 The Matrix of a Linear Transformation

Theorem: Let  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  be a linear transformation. Then there exists a unique matrix  $A$  such that  $T(\vec{x}) = A\vec{x}$  for all  $\vec{x}$  in  $\mathbb{R}^n$ .

Definition:

1. A map  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  is said to be **onto**  $\mathbb{R}^m$  if each  $\vec{b}$  in  $\mathbb{R}^m$  is the image of at least one  $\vec{x}$  in  $\mathbb{R}^n$ .
2. A map  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  is said to be **one-to-one** if  $T\vec{x} = T\vec{y}$ , then  $\vec{x} = \vec{y}$ , i.e. each  $\vec{b}$  in  $\mathbb{R}^m$  is the image of at most one  $\vec{x}$  in  $\mathbb{R}^n$ .

Theorem: Let  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  be a linear transformation. Then  $T$  is **one-to-one** if and only if the equation  $T\vec{x} = \vec{0}$  has only the trivial solution.

Theorem: Let  $T : \mathbb{R}^n \mapsto \mathbb{R}^m$  be a linear transformation, and let  $A$  be the standard matrix of  $T$ .

1.  $T$  maps  $\mathbb{R}^n$  onto  $\mathbb{R}^m$  iff the column vectors of  $A$  span  $\mathbb{R}^m$ .
2.  $T$  is one-to-one iff the column vectors of  $A$  are linear independent.

Remarks: Column of  $A$  does not span  $\mathbb{R}^3$  and  $T$  is not onto if there are no solution.

## 2 Chapter 2

### 2.1 Matrix Operation

$A \cdot (B\vec{x})$  is a linear transformation from  $\mathbb{R}^n$  to  $\mathbb{R}^m$

### 2.2 The Inverse of Matrix

Definition: An  $n \times n$  matrix  $A$  is said to be **invertible** if there is an  $n \times n$  matrix  $C$  such that

$$CA = I_n \text{ and } AC = I_n$$

In this case,  $C$  is an inverse of  $A$  and denoted by  $A^{-1}$  Remark: The inverse of  $A$  is unique.

Definition: If a matrix is not invertible, it is called singular matrix.

Theorem: Let  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . If  $ad - bc \neq 0$ , then  $A$  is invertible and  $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ . If  $ad - bc = 0$ , then  $A$  is not invertible and  $ad - bc$  is called the determinant of  $A$ , denoted by  $\det(A)$

Theorem:

1. If  $A$  is invertible matrix, then  $A^{-1}$  is invertible and  $(A^{-1})^{-1} = A$ .
2. If  $A$  and  $B$  are  $n \times n$  invertible matrices, then so is  $AB$ , and the inverse of  $A$  and  $B$  in the reverse order.  
That is,

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

3. 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,

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

Theorem: An  $n \times n$  matrix  $A$  is invertible iff  $A$  is row equivalent to  $I_n$ , and this case, any sequence of elementary row operations that reduces  $A$  to  $I_n$  also transform  $I_n$  into  $A^{-1}$ .

Find the third column of  $A^{-1}$  without computing the other columns.  $[A \ e_3] \sim [I \ b_3]$

## 2.3 The Characterization of Invertible Matrices

Theorem: Let  $A$  be an  $n \times n$  matrix. Then the following statements are equivalent.

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

Method to check whether  $A$  is invertible: use row operations to see the number of pivot position whether is equal to  $n$ .

## 3 Chapter 3

### 3.1 Introduction to Determinant

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

## 3.2 Properties of Determinant

Theorem:

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

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

**Corollary:**  $\det A = 0$  if the rows of  $A$  are linearly dependent

Theorem:

1.  $\det A^T = \det A$
2. Multiplicative Property:  $\det(AB) = \det A \cdot \det B$
3.  $\det(A^{-1}) = \frac{1}{\det(A)}$

Remark:  $\det(A + B) \neq \det A + \det B$

## 3.3 Cramer's Rule, Volume, and Linear Transformations

Theorem: Cramer's Rule

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

$$\vec{x}_i = \frac{\det A_i(b)}{\det A}, i = 1, 2, \dots, n$$

**Determinants as Area or Volume:**

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

**Linear Transformations:**

Theorem:

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

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

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

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

## 4 Chapter 4

### 4.1 Vector Spaces and Subspaces

Definition: A vector space is a nonempty set  $V$  of objects, called **vectors**, on which are defined two operations, called **addition** and **multiplication by scalars (real numbers)**, subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors  $\vec{v}, \vec{u}$  and  $\vec{w}$  in  $V$  and for all scalars  $c$  and  $d$ .

1. The sum of  $\vec{u}$  and  $\vec{v}$ , denoted by  $\vec{u} + \vec{v}$ , is in  $V$ .
2.  $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
3.  $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
4. There is a zero vector  $\vec{0}$  in  $V$  such that  $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
5. For each  $\vec{u}$  in  $V$ , there is a vector  $-\vec{u}$  in  $V$  such that  $\vec{u} + (-\vec{u}) = \vec{0}$
6. The scalar multiple of  $\vec{u}$  by  $c$ , denoted by  $c\vec{u}$ , is in  $V$ .
7.  $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
8.  $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
9.  $c(d\vec{u}) = (cd)\vec{u}$
10.  $1 \cdot \vec{u} = \vec{u}$

#### Subspaces:

Definition: A subspace of a vector space  $V$  is a subset  $H$  of  $V$  that has three properties:

1. The zero vector of  $V$  is in  $H$ . ( $\vec{0} \in H$ )
2.  $H$  is closed under vector addition. That is, for each  $\vec{u}$  and  $\vec{v}$  in  $H$ , the sum  $\vec{u} + \vec{v}$  is in  $H$
3.  $H$  is closed under multiplication by scalars. That is, for each  $\vec{u}$  in  $H$  and each scalar  $c$ , the vector  $c\vec{u}$  is in  $H$ .

An equivalent definition,  $H$  is a subspace of  $V$  if and only if for  $\vec{u}$  and  $\vec{v}$  in  $H$  and any scalar  $a$  and  $b$ ,  $a\vec{u} + b\vec{v}$  is in  $H$ .

For every vector spaces  $V$ , the subsets  $V$  and  $\{0\}$  are subspaces. Other subspaces are called **proper subspaces**.

### 4.2 Null Spaces, Column Spaces and Linear Transformation

#### The Null space of a matrix:

Definition: The null space of an  $m \times n$  matrix  $A$ , written as  $\text{Null}A$ , is the set of all solutions of the homogeneous equation  $A\vec{x} = \vec{0}$ . In set notation:

$$\text{Nul}A = \{\vec{x} : \vec{x} \text{ is in } \mathbb{R}^n \text{ and } A\vec{x} = \vec{0}\}$$

$\text{Nul}A$  is the set of all  $\vec{x} \in \mathbb{R}^n$  that are mapped into the zero vector of  $\mathbb{R}^m$  via the linear transformation  $\vec{x} \mapsto A\vec{x}$ .

Theorem: The null space of an  $m \times n$  matrix  $A$  is a subspace of  $\mathbb{R}^n$ .

Remark: When  $\text{Nul}A$  contains nonzero vectors, the number of vectors in the spanning set for  $\text{Nul}A$  equals the number of free variables in the equation  $A\vec{x} = \vec{0}$ .

**The Column space of a matrix:**

Definition: The column space of an  $m \times n$  matrix  $A$ , written as  $\text{Col}A$ , is the set of all linear combinations of the columns of  $A$ . If  $A = [\vec{a}_1, \dots, \vec{a}_n]$ , then

$$\text{Col}A = \text{Span}\{\vec{a}_1, \dots, \vec{a}_n\} = \{\vec{b} \mid \vec{b} = A\vec{x} \text{ for some } \vec{x} \in \mathbb{R}^n\}$$

Theorem: The column space of an  $m \times n$  matrix  $A$  is a subspace of  $\mathbb{R}^m$ .

The column space of an  $m \times n$  matrix  $A$  is all of  $\mathbb{R}^m$  if and only if the equation  $A\vec{x} = \vec{b}$  has a solution for each  $\vec{b}$  in  $\mathbb{R}^m$ .

Remark: to test a vector in  $\text{Col}A$  or not, we solve  $[A, \vec{u}]$  and see whether  $A\vec{x} = \vec{u}$  has a solution. Thus  $\vec{u}$  is in  $\text{Col}A$ .

**Contrast Between  $\text{Nul}A$  and  $\text{Col}A$  for an  $m \times n$  Matrix  $A$ :**

$\text{Nul}A$	$\text{Col}A$
$\text{Nul}A$ is a subspace of $\mathbb{R}^n$	$\text{Col}A$ is a subspace of $\mathbb{R}^m$
Solve $A\vec{x} = \vec{0}$ to get $\text{Nul}A$	$\text{Span}\{\vec{a}_1, \dots, \vec{a}_n\}$
$\vec{v} \in \text{Nul}A \iff A\vec{u} = \vec{0}$	$\vec{u} \in \text{Col}A \iff A\vec{x} = \vec{u} \text{ consistent}$

**Kernel and Range of a Linear Transformation:**

Definition: A linear transformation  $T$  from a vector space  $V$  into a vector space  $W$  is a rule that assigns to each vector  $\vec{x}$  in  $V$  a unique vector  $T(\vec{x})$  in  $W$ , such that  $T(a\vec{u} + b\vec{v}) = aT(\vec{u}) + bT(\vec{v})$  for all vectors  $\vec{u}$  and  $\vec{v}$  in  $V$  and all scalars  $a, b$ .

**Kernel (or Null space) of  $T$**   $= \{\vec{u} \in V \mid T(\vec{u}) = \vec{0}\}$

**Range of  $T$**   $= \{\vec{w} \in W \mid \vec{w} = T(\vec{u}) \text{ for some } \vec{u} \in V\}$

### 4.3 Linearly Independent Sets, Bases

Definition: An indexed set of vectors  $\{\vec{v}_1, \dots, \vec{v}_p\}$  in  $V$  is said to be **linearly independent** if the vector equation  $C_1\vec{v}_1 + C_2\vec{v}_2 + \dots + C_p\vec{v}_p = \vec{0}$  has only trivial solution  $C_1 = 0, \dots, C_p = 0$ .

Definition: Let  $H$  be a subspace of a vector space  $V$ . An indexed set of vectors  $B = \{\vec{b}_1, \dots, \vec{b}_p\}$  in  $V$  is a **basis** for  $H$  if

1.  $B$  is a linearly independent set, and
2. the subspace spanned by  $B$  coincides with  $H$ . That is,  $H = \text{Span}\{\vec{b}_1, \dots, \vec{b}_p\}$

Example:  $A$ :  $n \times n$  matrix,  $A = [\vec{a}_1, \dots, \vec{a}_n]$

If  $A$  is invertible ( $\det A \neq 0$ ), then  $\vec{a}_1, \dots, \vec{a}_n$  are linearly independent,  $\vec{b} = A\vec{x}$  always has a solution  $\vec{x} = A^{-1}\vec{b}$ .

Thus  $\mathbb{R}^n = \text{Span}\{\vec{a}_1, \dots, \vec{a}_n\}$  is a basis of  $\mathbb{R}^n$

Observation:

1. Each non-pivot column of  $B$  is a linear combination of the pivot columns.
2. If  $A$  is a matrix whose reduced echelon form is  $B$ , then  $A\vec{x} = \vec{0}$  iff  $B \cdot \vec{x} = \vec{0}$

Theorem: The pivot columns of a matrix  $A$  form a basis for  $\text{Col}A$ .

## 4.4 Coordinate Systems

Theorem: The Unique Representation Theorem

Let  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  be a basis for a vector space  $V$ . Then for each  $\vec{x}$  in  $V$ , there exists a unique set of scalars  $C_1, C_2, \dots, C_n$  such that

$$\vec{x} = C_1 \vec{b}_1 + \dots + C_n \vec{b}_n$$

Theorem: Let  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$  be a basis for a vector space  $V$ . Then the coordinate mapping  $\vec{x} \mapsto [\vec{x}]_B$  is a one-to-one linear transformation from  $V$  onto  $\mathbb{R}^n$ . And the transformation is called **isomorphism**.

## 4.5 The Dimension of a Vector Space

Theorem: If a vector space  $V$  has a basis  $B = \{\vec{b}_1, \dots, \vec{b}_n\}$ , then any set in  $V$  containing more than  $n$  vectors must be linearly dependent.

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

Definition: If  $V$  is spanned by a finite set, then  $V$  is said to be **finite-dimensional**, and the dimension of  $V$ , written as  $\dim V$ , is the number of vectors in a basis for  $V$ . The dimension of the zero vector space  $\{\vec{0}\}$  is defined to be zero. If  $V$  is not spanned by a finite set, then  $V$  is said to be **infinite-dimensional**.

**Subspace of a Finite-Dimensional Space:** Theorem: Let  $H$  be a subspace of a finite-dimensional vector space  $V$ . Any linearly independent set in  $H$  can be expanded, if necessary, to a basis for  $H$ . Also,  $H$  is finite-dimensional and

$$\dim H \leq \dim V.$$

**The Basis Theorem:** Let  $V$  be a  $p$ -dimensional vector space,  $p \geq 1$ , any linearly independent set of exactly  $p$  elements in  $V$  is automatically a basis for  $V$ . Any set of exactly  $p$  elements that spans  $V$  is automatically a basis for  $V$ .

**The Spanning Set Theorem:** Let  $S = \{\vec{v}_1, \dots, \vec{v}_p\}$  be a set in  $V$ , and let  $H = \text{Span}\{\vec{v}_1, \dots, \vec{v}_p\}$ .

1. If one of the vectors in  $S$ — say,  $\vec{v}_k$ — is a linear combination of the remaining vectors in  $S$ , then the set formed from  $S$  by removing  $\vec{v}_k$  still spans  $H$ .
2. If  $H \neq \{\vec{0}\}$ , some subset of  $S$  is a basis for  $H$ .

**The dimension of  $\text{Nul}A$  and  $\text{Col}A$ :**

1. The  $\dim(\text{Nul}A)$  = number of free variables in the equation  $A\vec{x} = \vec{0}$
2. The  $\dim(\text{Col}A)$  = number of pivot columns of  $A$ .

Exercise: Let  $V$  is a nonzero finite-dimensional vector space.



1. If  $\dim V = p$  and if  $S$  is a linearly dependent subset of  $v$ , then  $S$  contains more than  $p$  vectors. Answer: False
2. If  $S$  spans  $V$  and if  $T$  is a subset of  $V$  that contains more vectors than  $S$ , then  $T$  is linearly dependent. Answer: True

## 4.6 Rank

**The Row Space:** Theorem: The two matrices  $A$  and  $B$  are row equivalent, then their row spaces are the same. If  $B$  is in echelon form, the nonzero rows of  $B$  form a basis for the row space of  $A$  as well as for that of  $B$ .

Ex: Find basis for RowA, ColA, NulA

$$A = \begin{pmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{pmatrix}$$

Solution:  $A \sim B = \begin{pmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Basis for RowA:  $\{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$

Basis for ColA:  $\left\{ \begin{pmatrix} -2 \\ 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} -5 \\ 3 \\ 11 \\ 7 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 7 \\ 5 \end{pmatrix} \right\}$

NulA:  $\begin{cases} x_1 + x_3 + x_5 = 0 \\ x_2 - 2x_3 + 3x_5 = 0 \\ x_4 - 5x_5 = 0 \end{cases}$

$\Rightarrow \begin{cases} x_1 = -x_3 - x_5 \\ x_2 = 2x_3 - 3x_5 \\ x_4 = 5x_5 \end{cases}$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -x_3 - x_5 \\ 2x_3 - 3x_5 \\ x_3 \\ 5x_5 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 2x_3 \\ x_3 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -x_5 \\ -3x_5 \\ 0 \\ 5x_5 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{pmatrix}$$

Basis for NulA:  $\left\{ \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{pmatrix} \right\}$

**The Rank Theorem:** Definition: The rank of  $A$  is the dimension of the column space of  $A$ .

$$\text{Rank } A = \dim(\text{Col } A)$$

The rank Theorem: Let  $A : m \times n$

1.  $\text{Rank } A = \dim(\text{Col } A) = \dim(\text{Row } A) = \text{number of pivots in } A \leq m$
2.  $\text{Rank } A + \dim(\text{Nul } A) = n$

The Invertible Matrix Theorem:

Let  $A$  be an  $n \times n$  matrix. Then the following statements are equivalent.

1.  $A$  is an invertible matrix.
2.  $A$  is row equivalent to the  $n \times n$  identity matrix.
3.  $A$  has  $n$  pivot positions.
4. The equation  $A\vec{x} = \vec{0}$  has only the trivial solution.
5. The columns of  $A$  form a linearly independent set.
6. The linear transformation  $\vec{x} \mapsto A\vec{x}$  is one-to-one.
7. The equation  $A\vec{x} = \vec{b}$  has at least one solution for each  $\vec{b}$  in  $\mathbb{R}^n$ .
8. The columns span  $\mathbb{R}^n$ .
9. The linear transformation  $\vec{x} \mapsto A\vec{x}$  maps  $\mathbb{R}^n$  onto  $\mathbb{R}^n$ .
10. There is an  $n \times n$  matrix  $C$  such that  $CA = I$ .
11. There is an  $n \times n$  matrix  $D$  such that  $AD = I$ .
12.  $A^T$  is an invertible matrix.
13. The columns of  $A$  form a basis of  $\mathbb{R}^n$ .
14.  $\text{Col } A = \mathbb{R}^n$
15.  $\dim(\text{Col } A) = n$
16.  $\text{rank } A = n$
17.  $\text{Nul } A = \{\vec{0}\}$
18.  $\dim(\text{Nul } A) = 0$
19.  $\det(A) \neq 0$

Remarks:  $\text{Col } A^T = \text{Row } A$  and  $\text{Rank } A^T = \text{Rank } A$ .

## 5 Chapter 5

### 5.1 Eigenvectors and Eigenvalues

Definition: An **eigenvector** of an  $n \times n$  matrix  $A$  is a nonzero vector  $\vec{x}$  such that  $A\vec{x} = \lambda\vec{x}$  for some scalar  $\lambda$ .

1.  $\lambda$  is called an eigenvalue of  $A$ .
2.  $\vec{v}$  is called an eigenvector corresponding to  $\lambda$ .

Theorem: The eigenvalues of a triangular matrix are the entries on its main diagonal.

The scalar  $\lambda$  is an eigenvalue if and only if

1.  $(A - \lambda I)\vec{x} = \vec{0}$  has a nontrivial solution.
2.  $(A - \lambda I)\vec{x} = \vec{0}$  has a free variable
3. at least one of the entries on the diagonal of  $A - \lambda I$  is zero. (Number of pivots of  $A - \lambda I \leq n$ )
4.  $\lambda = a_{11}$  or  $\lambda = a_{22}$  or  $\dots$ ,  $\lambda = a_{nn}$

Example: If  $\vec{x}$  is an eigenvector of  $A$  corresponding to  $\lambda$ ,  $A^n \vec{x} = \lambda^n \vec{x}$

Theorem: If  $\vec{v}_1, \dots, \vec{v}_r$  are eigenvectors that correspond to distinct eigenvalues  $\lambda_1, \dots, \lambda_r$  of an  $n \times n$  matrix  $A$ , then the set  $\{\vec{v}_1, \dots, \vec{v}_r\}$  is linearly independent.  $\implies$  The eigenvectors are linearly independent.

### 5.2 The characteristic equation

Definition: The characteristic equation:  $A : n \times n$  matrix  $\det(A - \lambda I) = 0$  is called the characteristic equation of  $A$ .  $\det(A - \lambda I)$  is called the characteristic polynomial of  $A$ , a polynomial of degree  $n$  in the variable  $\lambda$ .

Theorem:  $\lambda$  is an eigenvalue of  $A$  if and only if  $\lambda$  satisfies the characteristic equation  $\det(A - \lambda I) = 0$ .

Multiplicity: The degree of the corresponding eigenvalue.

**Similarity** Definition: Let  $A, B$  be  $n \times n$  matrices.

$A$  is similar to  $B$  if there exists an invertible matrix  $P$  such that  $P^{-1}AP = B$  or equivalently,  $A = PBP^{-1}$

Changing  $A$  into  $P^{-1}AP$  is called a similarity transformation.

Theorem: If  $n \times n$  matrices  $A$  and  $B$  are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Warnings:

1. The matrices  $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$  and  $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$  are not similar even though they have the same eigenvalues.
2. Similarity is not the same as row equivalence. Row operations on a matrix usually change its eigenvalues.

## 5.3 Diagonalization

If  $A = PDP^{-1}$  where  $P$  is invertible and  $D$  is diagonal, then  $A^k = PD^kP^{-1}$ .

Definition: A square matrix  $A$  is said to be diagonalizable if  $A$  is similar to diagonal matrix. That is, if  $A = PDP^{-1}$  for some invertible matrix  $P$  and some diagonal matrix  $D$ .

The Diagonalization Theorem:

An  $n \times n$  matrix  $A$  is diagonalizable iff  $A$  has  $n$  linearly independent eigenvectors.

$A = PDP^{-1}$  with  $D$  is diagonal matrix iff the columns of  $P$  are  $n$  linearly independent eigenvectors of  $A$ . In this case, the diagonal entries of  $D$  are eigenvalues of  $A$  that correspond, respectively, to the eigenvectors in  $P$ .

### Diagonalizing Matrices

1. Find eigenvalues:  $\det(A - \lambda I) = 0$
2. Find linearly independent eigenvectors of  $A$ :  $(A - \lambda I)\vec{x} = \vec{0}$
3. Construct  $P$ :  $P = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$
4. Construct  $D$ .

Theorem: An  $n \times n$  matrix with  $n$  distinct eigenvalues is diagonalizable.

**Matrices Whose Eigenvalues Are Not Distinct:** Theorem: Let  $A$  be an  $n \times n$  matrix whose distinct eigenvalues are  $\lambda_1, \dots, \lambda_p$ .

1. For  $1 \leq k \leq p$ , the dimension of the eigenspace for  $\lambda_k$  is less than or equal to the multiplicity of the eigenvalue  $\lambda_k$ .
2. The matrix  $A$  is diagonalizable iff the sum of the dimensions of the eigenspace equals to  $nm$  and this happens if and only if
  - (a) The characteristic polynomial factors completely into linear factors
  - (b) The dimension of the eigenspace for each  $\lambda_k$  equals the multiplicity of  $\lambda_k$ .

Remarks:

1. If  $A = PDP^{-1}$ , then  $A^{-1} = PD^{-1}P^{-1}$ .
2. If  $A = PDP^{-1}$ , then  $A^T = P^T DP^{-1}$ .

## 5.4 Discrete Dynamical Systems

## 5.5 Applications to Differential Equations

# 6 Chapter 6

## 6.1 Inner Product, Length and Orthogonality

The Inner Product:

$$\vec{a}^T \vec{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = \vec{a} \cdot \vec{b}$$

Theorem: Let  $\vec{u}$ ,  $\vec{v}$  and  $\vec{w}$  be vectors in  $\mathbb{R}^n$ , and let  $c$  be a scalar. Then

1.  $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
2.  $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
3.  $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$
4.  $\vec{u} \cdot \vec{u} \geq 0$  and  $\vec{u} \cdot \vec{u} = 0$  iff  $\vec{u} = \vec{0}$

**The length of a vector:** Definition: The length (or norm) of  $\vec{u}$  is the non-negative scalar  $\|\vec{u}\|$  defined by

$$\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$$

Unit Vector:  $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$

**Distance in  $\mathbb{R}^n$ :** Definition: For  $\vec{u}$  and  $\vec{v}$  in  $\mathbb{R}^n$ , the distance between  $\vec{u}$  and  $\vec{v}$ , written as  $\text{dist}(\vec{u}, \vec{v})$ , is the length of the vector  $\vec{u} - \vec{v}$ . That is,

$$\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\|$$

**Orthogonal Vectors(Perpendicular):**

Definition: Two vectors  $\vec{u}$  and  $\vec{v}$  in  $\mathbb{R}^n$  are orthogonal (to each other) if  $\vec{u} \cdot \vec{v} = 0$ .

Remarks:  $\vec{0}$  is orthogonal to every vector in  $\mathbb{R}^n$  because  $\vec{0}^T \vec{0} = 0$

**Orthogonal Complements:**

Definition: Let  $\vec{z} \in \mathbb{R}^n$ ,  $W$  a subspace of  $\mathbb{R}^n$ . If  $\vec{z}$  is orthogonal to every vector in  $W$ , we say  $\vec{z}$  is orthogonal to  $W$ .

Definition: The set of all vectors  $\vec{z}$  that are orthogonal to  $W$  is called the orthogonal complement of  $W$ , and is denoted by  $W^\perp$ .

Fact:

1. A vector  $\vec{x}$  is in  $W^\perp$  iff  $\vec{x}$  is orthogonal to every vector in a set that spans  $W$ .
2.  $W^\perp$  is a subspace of  $\mathbb{R}^n$ .

Theorem: Let  $A$  be an  $m \times n$  matrix. The orthogonal complement of the row space of  $A$  is the null space of  $A$ , and the orthogonal complement of the column space of  $A$  is the null space of  $A^T$ .

$$(\text{Row } A)^\perp = \text{Nul } A$$

$$(\text{Col } A)^\perp = \text{Nul } A^T$$

Let  $W$  be a subspace of  $\mathbb{R}^n$ . Then  $W^\perp$  is also a subspace of  $\mathbb{R}^n$ .

$$\dim W + \dim W^\perp = n$$

## 6.2 Orthogonal Sets

Definition: A set of vectors  $\{\vec{u}_1, \dots, \vec{u}_p\}$  in  $\mathbb{R}^n$  is said to be an orthogonal set if each pair of distinct vectors from the set is orthogonal, that is, if  $\vec{u}_i \cdot \vec{u}_j = 0$  whenever  $i \neq j$ .

Theorem: If  $S = \{\vec{u}_1, \dots, \vec{u}_p\}$  is an orthogonal set of nonzero vectors in  $\mathbb{R}^n$ , then  $S$  is linearly independent and hence is a basis for the subspace spanned by  $S$ .

Definition: An orthogonal basis for a subspace  $S$  of  $\mathbb{R}^n$  is a basis for  $W$  that is also an orthogonal set.

Theorem: Let  $\{\vec{u}_1, \dots, \vec{u}_p\}$  be an orthogonal basis for a subspace  $W$  of  $\mathbb{R}^n$ . For each  $y$  in  $W$ , the weights in the linear combination

$$\vec{y} = C_1 \vec{u}_1 + C_2 \vec{u}_2 + \dots + C_p \vec{u}_p$$

are given by

$$C_j = \frac{\vec{y} \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}, j = 1, \dots, p$$

Remarks: If the basis were not orthogonal, it would be necessary to solve a system of linear equations in order to find the weights.

Orthogonal Projection:

$$\text{Proj}_{\vec{u}} \vec{y} = \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{\vec{u} \cdot \vec{y}}{\|\vec{u}\|^2} \vec{u}$$

$$\vec{w} = \vec{y} - \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \vec{u}$$

**Orthonormal Sets:**

A set contains unit vectors, i.e.  $\vec{u}_i \cdot \vec{u}_j = 0$  for  $i \neq j$  and  $\vec{u}_i \cdot \vec{u}_i = 1$ . If  $W$  is a subspace spanned by an orthonormal set, then the set is an orthonormal basis for  $W$ .

**Orthonormal columns:**

Theorem: An  $m \times n$  matrix  $U$  has orthonormal columns iff  $U^T U = I$ .

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

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

Remarks: Properties 1 and 3 say that the linear mapping  $\vec{x} \mapsto U\vec{x}$  preserves lengths and orthogonality.

## 6.3 Orthogonal Projection

In general, find a decomposition of  $\vec{u}$  in  $\mathbb{R}^n$  as  $\vec{u} = \vec{u}_1 + \vec{u}_2$ ,  $\vec{u}_1$  is in a subspace  $W$  and  $\vec{u}_2$  is in  $W^\perp$ .  $\vec{u}_1$  is called the orthogonal projection of  $\vec{u}$  onto the subspace  $W$ .

**The Orthogonal Decomposition Theorem:**

Let  $W$  be a subspace of  $\mathbb{R}^n$ . Then each  $\vec{y}$  in  $\mathbb{R}^n$  can be written uniquely in the form

$$\vec{y} = \hat{\vec{y}} + \vec{z}$$

where  $\hat{\vec{y}}$  is in  $W$  and  $\vec{z}$  is in  $W^\perp$ . In fact, if  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is any orthogonal basis of  $W$ , then

$$\hat{\vec{y}} = \frac{\vec{u}_1 \cdot \vec{y}}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \dots + \frac{\vec{u}_p \cdot \vec{y}}{\vec{u}_p \cdot \vec{u}_p} \vec{u}_p$$

and  $\vec{z} = \vec{y} - \hat{\vec{y}}$ .

### Properties of Orthogonal Projections:

Let  $W = \text{span}\{\vec{u}_1, \dots, \vec{u}_p\}$ , where  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is an orthogonal basis for  $W$ .

If  $\vec{y} \in W$ , then  $\text{Proj}_W \vec{y} = \vec{y}$

### 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  $\hat{\vec{y}}$  be the orthogonal projection of  $\vec{y}$  onto  $W$ . Then  $\hat{\vec{y}}$  is the closest point in  $W$  to  $\vec{y}$ , in the sense that

$$\|\vec{y} - \hat{\vec{y}}\| \leq \|\vec{y} - \vec{v}\|$$

for all  $\vec{v}$  in  $W$  distinct from  $\hat{\vec{y}}$ .

Definition: The distance from a point  $\vec{y}$  in  $\mathbb{R}^n$  to a subspace  $W$  is defined to be the distance from  $\vec{y}$  to the nearest point in  $W$ .

Theorem: If  $\{\vec{u}_1, \dots, \vec{u}_p\}$  is an orthonormal basis for a subspace  $W$  of  $\mathbb{R}^n$ , then

$$\hat{\vec{y}} = (\vec{y} \cdot \vec{u}_1) \vec{u}_1 + \dots + (\vec{y} \cdot \vec{u}_p) \vec{u}_p$$

If  $U = \{\vec{u}_1 \vec{u}_2 \dots \vec{u}_p\}$ , then

$$\hat{\vec{y}} = UU^T \vec{y}$$

for all  $\vec{y}$  in  $\mathbb{R}^n$ .

## 6.4 The Gram Schmidt Process

The Gram-Schmidt process is a simple algorithm for producing an orthogonal or orthogonal basis for any nonzero subspace of  $\mathbb{R}^n$ .

Theorem: Refer to the notes.

### The QR Factorization:

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

Example: Find a QR factorization of  $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ .

Solution: The columns of  $A$  are the vectors  $\vec{x}_1, \vec{x}_2, \vec{x}_3$  in Example 2. The orthonormal basis are

$$\vec{u}_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 0 \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$$

$$\text{Let } Q = [\vec{u}_1, \vec{u}_2, \vec{u}_3] = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

Note that  $Q^T Q = I$ , we have

$$R = Q^T Q R = Q^T A$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & \frac{3}{2} & 1 \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

## 6.5 Least-Square Problems

General least-squares problem: To find an  $\hat{\vec{x}}$  that makes  $\|\vec{b} - A\hat{\vec{x}}\|$  as small as possible.

Definition: If  $A$  is  $m \times n$  and  $\vec{b}$  is in  $\mathbb{R}^m$ , a least-squares solution of  $A\vec{x} = \vec{b}$  is an  $\hat{\vec{x}}$  in  $\mathbb{R}^n$  such that

$$\|\vec{b} - A\hat{\vec{x}}\| \leq \|\vec{b} - A\vec{x}\|$$

for all  $\vec{x}$  in  $\mathbb{R}^n$ .

Theorem: The set of least-squares solutions of  $A\vec{x} = \vec{b}$  coincides with the nonempty set of solutions of the normal equations

$$A^T A \vec{x} = A^T \vec{b}$$

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

1. The equation  $A\vec{x} = \vec{b}$  has a unique least-squares solution for each  $\vec{b}$  in  $\mathbb{R}^m$ .
2. The columns of  $A$  are linearly independent.
3. The matrix  $A^T A$  is invertible.

For given  $A$  and  $\vec{b}$ ,  $\hat{\vec{x}} = \begin{pmatrix} \frac{\vec{b} \cdot \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_1} \\ \frac{\vec{b} \cdot \vec{a}_2}{\vec{a}_2 \cdot \vec{a}_2} \end{pmatrix}$



Theorem: Given an  $m \times n$  matrix  $A$  with linearly independent columns, let  $A = QR$  be a  $QR$  factorization of  $A$ . Then, for each  $\vec{b}$  in  $\mathbb{R}^m$ , the equation  $A\vec{x} = \vec{b}$  has a unique least-squares solution, given by

$$\hat{\vec{x}} = R^{-1}Q^T\vec{b}$$

$$R\vec{x} = Q^T\vec{b}$$

## **6.6 Applications to Linear Models**

# **7 Chapter 7**

## **7.1 Diagonalization of Symmetric Matrices**

## **7.2 Quadratic Form**