# Lecrture 1

Zecheng Zhang

January 18, 2023

## 1 Matrix and vector

## 2 Row operations

Elementary row operations (EROs)

1 (Replacement) Replace one row by the sum of itself and a multiple of another row

2 (Interchange) Interchange two rows

3 (Scaling) Multiply all entries in a row by a nonzero constant

We have some remarks.

1 Two matrices are called row equivalent if there is a sequence of elementary row operations that transforms one matrix into the other.

2 If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

### 2.1 Gaussian elimination (self review)

Apply EROs to reduce the matrix to row echelon form (REF) or reduced row echelon form (RREF). We will discuss the LU decomposition, which is the realization of the Gaussian elimination by matrix multiplication.

# 3 Linearly independent

## 3.1 Linear combination

Linear combination of vectors $v_1, v_2, ..., v_n$ is still a vector of the form:

$$v = c_1 v_1 + c_2 v_2 + ... + c_n v_n, \tag{1}$$

where $c_1, c_2, ...c_n$ are scalars. $span\{v_1, v_2, ..., v_n\}$ is a set which contains all possible linear combinations of $v_1, v_2, ..., v_n$.

## 3.2 Linearly independent

An indexed set of vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ in $\mathbb{R}^n$ is said to be linearly independent if the vector equation

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

has only a trivial solution. Equivalently, the linear system

$$Ax = 0,$$

has zero (trivial) solution only, where $A = [v_1, v_2, ...v_p] \in \mathbb{R}^{n\times p}$ and $x = [x_1, ..., x_p]^t$. The set $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_p$ is said to be linearly dependent if there exist weights $c_1, c_2, \cdots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

# 4 Linear transformation

## 4.1 Transformation

A transformation $T$ 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 $\mathbb{R}^m$. Denote by $T : \mathbb{R}^n \to \mathbb{R}^m$.

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

## 4.2 Linear transformation

A transformation $T$ is linear if:

1 $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}))$ for all $\mathbf{u}, \mathbf{v}$ in the domain of $T$,

2 $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars $c$ and all $\mathbf{u}$ in the domain of $T$.

If $T$ is a linear transformation, then

1 $T(\mathbf{0}) = \mathbf{0}$

2 $T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$

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

**Example 4.1.** Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly dependent set in $\mathbb{R}^n$. Explain why the set $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ is linearly dependent. What about if $v_1, v_2, v_3$ are linearly independent?

**Theorem 4.2.** 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 $j$th column is the vector $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the $j$th column of the identity matix in $\mathbb{R}^n$:

$$A = [T(\mathbf{e}_1) \quad \cdots \quad T(\mathbf{e}_n)] \tag{2}$$

The matrix $A$ in (2) is called the standard matrix for the linear transformation $T$.

**Remark 1.** For some other topics, such as one-to-one and onto, please refer to sec 1.9 notes.

# 5 Matrix multiplications

## 5.1 Dot products

## 5.2 Row column rule

## 5.3 One equivalence

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

$$AB = A[\mathbf{b}_1, \cdots, \mathbf{b}_p] = [A\mathbf{b}_1, \cdots, A\mathbf{b}_p]$$

**Remark 2.** Each column of $AB$ is a linear combination of the columns of $A$ using weights from the corresponding column of $B$.

**Example 5.1.** Find the third column of $D$ given $CD$ and $C$.

**Remark 3.** Each row of $AB$ is a linear combination of the rows of $B$ using weights from the corresponding row of $A$.

### 5.4 Another equivalence

Multiply columns 1 to $n$ of $A$ times rows 1 to $n$ of $B$. Add those matrices.

### 5.5 LU decomposition

### 5.6 Row operations and the matrix multiplication

Perform the row operation on matrix $A \in \mathbb{R}^{m \times n}$ is equivalent to multiplying $A$ by the corresponding elementary matrix.

## 5.7 LU decomposition

The idea of transforming A into an upper triangular matrix $U$ is equivalent to multiplying a sequence of lower triangular matrices $L_k$ on the left:

$$\underbrace{L_{m-1}...L_1}_{L^{-1}} A = U.$$

Note that we restrict the process to the replacement and scaling row operations (lower triangular assumption). Setting $L = L_1^{-1}...L_{m-1}^{-1}$ yields $A = LU$, where $L$ is lower-triangular and $U$ is upper-triangular.

Consider solving $Ax = b$. Suppose one can establish a $LU$ decomposition for $A$, i.e., $A = LU$, it follows that,

$$LUx = b.$$

Now denote $y = Ux$, then it is equivalent to solving two triangular systems:

$$Ly = b,$$
$$Ux = y.$$

**Example 5.2.** Find LU decomposition of the following matrix:

# 6 Subspace

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

1. The zero vector is in $H$.
2. For each $\mathbf{u}$ and $\mathbf{v}$ in $H$, the sum $\mathbf{u} + \mathbf{v}$ is in $H$.
3. For each $\mathbf{u}$ in $H$ and each scalar $c$, the vector $c\mathbf{u}$ is in $H$.

**Remark 4.** A subspace is closed under addition and scalar multiplication.

## 6.1 Column space

The column space of a matrix $A$ is the set Col $(A)$ of all linear combinations of the columns of $A$. If $A = [a_1, a_2, ..., a_n]$, then $col(A) = span\{a_1, a_2, ..., a_n\}$.

## 6.2 Null space

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

## 6.3 Row space

The row space of a matrix A is the set $row(A)$ of all linear combinations of the rows of $A$.

**Theorem 6.1.** If two matrices $A$ and $B$ are row equivalent, their row spaces are the same.

*Proof.* Row operations are indeed the linear combinations of rows. If $B$ is obtained from $A$ by the EROs, the rows of $B$ are the linear combinations of rows of $A$. As a result, the row space of $B$ is in the row space of $A$. The other way is the same. □

What about the column space?

**Remark 5.**
- The column space of an $m \times n$ matrix is a subspace of $\mathbb{R}^m$.
- The null space of an $m \times n$ matrix $A$ is a subspace of $\mathbb{R}^n$.

### 6.4 Basis

A basis for a subspace $H$ of $\mathbb{R}^n$ is a linearly independent set in $H$ that spans $H$.

**Theorem 6.2.** The pivot columns of a matrix $A$ form a basis for the column space of $A$.

**Theorem 6.3.** The nonzero rows of the $ref(A)$ form a basis for $row(A)$.

### 6.5 Dimension

The dimension of a nonzero subspace $H$, denoted by $\dim(H)$, is the number of vectors on any basis for $H$. The dimension of the zero subspace $\{\mathbf{0}\}$ is defined as zero.

### 6.6 Rank

The rank of a matrix $A$, denoted by $\operatorname{rank}(A)$, is the dimension of the column space of $A$. In addition, the dimension of the null space is called nullity. In addition, the row space dimension is called the matrix's row rank.

**Remark 6.** The EROs do not change the dimension of the column space; hence the EROs do not change the rank of the matrix.

**Theorem 6.4** (Rank theorem)**.** If a matrix $A$ has $n$ columns, then $rank(A)+dim(null(A)) = n$.

**Example 6.5.** Find the rank, column space, and null space of the matrix.

**Theorem 6.6.** Some facts about the rank.

1. $rank(AB) \leq \min(rank(A), rank(B))$.
2. $rank(A+B) \leq rank(A) + rank(B)$.

3. $rank(AA^T) = rank(A^TA) = rank(A) = rank(A^T)$.

*Proof.* I will only show the first statement and leave the other two as the homework questions. Since the columns of $AB$ are the linear combinations of columns of $A$ by $B$, this implies that $dim(col(AB) \leq dim(col(A))$. It follows that $rank(AB) \leq rank(A)$.

Suppose $x \in null(B)$, this implies that $Bx = 0$, consequently, $ABx = 0$, or, $x \in null(AB)$. This indeed shows that $null(B) \subset null(AB)$, or $dim(null(B)) \leq dim(null(AB))$. Since $B$ and $AB$ have the same number of columns, it follows from the Rank theorem that $rank(AB) \leq rank(B)$.

□

### 6.6.1 Rank decomposition

Every rank $r$ matrix $A \in \mathbb{R}^{m \times n}$ matrix has a rank decomposition $A = CR$, where $C \in \mathbb{R}^{m \times r}$, $R \in \mathbb{R}^{r \times n}$ and columns of $C$ form a basis for $col(A)$. One can construct $C$ by taking all linearly independent columns of $A$. Because each column of $A$ is the linear combination of columns of $C$ by weights from the corresponding columns of $A$, the $A$ matrix can be constructed easily. One way is to remove all zero rows from $ref(A)$.

**Theorem 6.7.** For $A \in \mathbb{R}^{m \times n}$, we have $rank(A) = rank(A^t)$.

*Proof.* Suppose $rank(A) = r$ and admits the rank-decomposition $A = CR$. We have $A^t = R^t C^t$. Since the columns of $A^t$ is the linear combination of columns of $R^t$, this implies that $col(A^t) \subset col(R^t)$, or $rank(A^t) \leq rank(R^t) \leq r = rank(A)$. Consider $A = (A^t)^t$ and complete the proof by yourself. $\square$

**Theorem 6.8.** The row rank is equal to the column rank of a matrix.