2021511019 Jinho Song

1 Introduction to system of Linear Algebra

1.1 Linear Algebra

Definition : A Linear equation in the variables $X_1, \cdots, X_n$ is an equation that can be written as

$$a_1X_1 + \cdots + a_nX_n = b$$

Examples : which of the following equation are Linear

1. $4X_1 - 5X_2 + 2 = X_1$
2. $X_2 = 2(\sqrt{6} - X_1) + X_3$
3. $4X_1 - 6X_2 = X_1X_2$
4. $X_2 = 2\sqrt{X_1} - 7$

Linear System

Definition: A linear system is a collection of one or more linear equation involving the same set of variables $X_1, \cdots, X_n$. A solution of a linear system is a list $(s_1, \cdots, s_n)$ are substituted for $X_1, \cdots, X_n$ respectively.

Examples:

$$\begin{array}{ccc} (1) & (2) & (3) \\ X_1 + X_2 = 1 & X_1 - 2X_2 = -3 & 2X_1 + X_2 = 1 \\ -X_1 + X_2 = 0 & 2X_1 - 4X_2 = 8 & -4X_1 - 2X_2 = -2 \end{array}$$

(1) Only one solution (2) No solution (3) Infinitely many solution

Theorem : A Linear system has either

- no solution
- one unique solution
- infinitely many solution

Definition : A system is consistent if a solution exists

1.2 How to solve a linear system

Strategy : Replace a system with an equivalent system which is easier to solve.
Definition : Linear systems are equivalent if they have the same set of solutions.
Example :

$$\begin{aligned} X_1 + X_2 &= 1 \\ -X_1 + X_2 &= 0 \end{aligned} \qquad \xrightarrow{R_2 \leftarrow R_2 + R_1} \qquad \begin{aligned} X_1 + X_2 &= 1 \\ 2X_2 &= 1 \end{aligned}$$

Augmented matrix:
$$\begin{aligned} X_1 - 2X_2 &= -1 \\ -X_1 + 3X_2 &= 3 \end{aligned}$$

2 Solution of linear systems via row reduction

- Gaussian elimination : Row reduce to echelon form (**REF**)
- Gauss-Jordan elimination : Row reduce to reduced echelon form (**RREF**)

2.1 General solution in parametric form

$$\left[\begin{array}{ccccc|c} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -8 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array}\right]$$

$$\begin{aligned} X_1 - 6X_2 + 3X_4 &= 0 \\ X_3 + -8X_4 &= 5 \\ X_5 &= 7 \end{aligned}$$

General Solution in parametric form:

$$\begin{aligned} &X_1 = -6X_2 - 3X_4 \\ &X_2 = X_2 \quad \textit{(free variable)} \\ &X_3 = 8X_4 + 5 \\ &X_4 = X_4 \quad \textit{(free variable)} \\ &X_5 = 7 \end{aligned}$$

2.2 Questions of existence and awareness

Linear system is **consistent** if and only if an REF of the augmented matrix has no row of form

$$\left[\begin{array}{ccccc|c} 0 & 0 & 0 & \cdots & 0 & b \end{array}\right] \quad (b \neq 0)$$

If consistent,
One unique solution *(no free variable)*
Infinitely many solutions *(at least one free variables)*
Summary:

- Each linear systems corresponds to an augmented matrix
- From the Gaussian elimination , we can read off, whether the system has no, one, or infinitely many solutions
- We can further row reduce to reduced echelon form
  - General solution in parametric form
  - This form is unique
- The recipe to solve linear systems
  1. Augmented matrix
  2. Gauss-Jordan elimination
  3. Declare pivot & free variables, state the solutions in terms of free variables

  x‘

3 Linear system - Column picture

A vector equation
$$x_1\vec{a_1} + \cdots + x_m\vec{a_m} = \vec{b}$$
has the same solution set as the linear system with augmented matrix
$$\left[\begin{array}{ccc|c} \vec{a_1} & \cdots & \vec{a_m} & \vec{b} \end{array}\right]$$
In particular, $\vec{b}$ can be generated by a linear combination of $\vec{a_1}, \cdots, \vec{a_m}$ if and if only, linear system is consistent

3.1 The span of a set vectors

Definition : The span of vectors $\vec{v_1}, \cdots, \vec{v_m}$ is the set of all their linear combinations. We denote it by $span\{\vec{v_1}, \cdots, \vec{v_m}\}$

Example :
$$span \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} = a \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$$
if a exists, span is line. However, a does not exists, so span is plane.

3.2 Matrix operations

3.2.1 Matrix times matrix

$$AB = \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} \vec{b_1} & \cdots & \vec{b_m} \end{bmatrix} = \begin{bmatrix} A\vec{b_1} & \cdots & A\vec{b_m} \end{bmatrix}$$

Each column of AB is linear combination of the columns of A with weight given by the corresponding column of B

$$(AB)\vec{x} = A(B\vec{x})$$

- $A(BC) = (AB)C$
- $A(B + C) = AB + AC$
- $(A + B)C = AC + BC$

4 Transpose of a matrix

Definition : The transpose A^T of a matrix A is the matrix whose columns are formed from the corresponding row of A.

Example :

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}^T = \begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}$$

Theorem :

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$
- $(ABC)^T = C^T B^T A^T$

5 LU Decomposition

Definition : An elementary matrix is one that is obtained by performing a single elementary row operation on an identity matrix.

$$A = \begin{bmatrix} 2 & 1 \\ 4 & -6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 4 & -6 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & -8 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{[L]} \begin{bmatrix} 2 & 1 \\ 0 & -8 \end{bmatrix}^{[U]}$$

A = LU is known as the **LU** decomposition of A.
(**L**: lower triangular, **U**: upper triangular)

Remark

- Once we have $A = LU$, it is simple to solve $A\vec{x} = \vec{b}$
- Not always factor every matrix A as $A = LU$

Definition: A permutation matrix is one that is obtained by performing row exchanges on an identity matrix.

Theorem: For any matrix A, there is a permutation matrix P such that $PA = LU$

6 The inverse of matrix

Definition: An $n \times n$ matrix A is invertible if there is a matrix B such that $AB = BA = I$. In that case, B is inverse of A, and we write $A^{-1} = B$

Theorem: Let A be invertible, Then the system $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$

Remark

- If $A\vec{x} = \vec{b}$ does not have the unique solution, then A is not invertible.

6.1 Recipe for computing the inverse

1. Form the augmented matrix $[A|I]$
2. Compute the reduced echelon form(Gauss-Jordan elimination)
3. If A is invertible, the result is of form $[I|A^-1]$

Example:

Find the inverse of $A = \begin{bmatrix} 2 & 0 & 0 \\ -3 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 2 & 0 & 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3/2 & 1 & 0 \end{array}\right]$$

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

Theorem : Suppose A and B are invertible,

- A^{-1} is invertible, $(A^{-1})^{-1} = A$

- A^T is invertible, and $(A^T)^{-1} = (A^{-1})^T$
- $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$
- $(AB)^{-1} = B^{-1}A^{-1}$

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

- A is invertible.
- A is row equivalent to I
- A has n pivots
- $A\vec{x} = \vec{b}$ has a unique solution

* Matrice that are not invertible are often called singular.

6.2 LU decomposition vs. matrix inverse

- In many applications, we don't just solve $A\vec{x} = \vec{b}$ for a single $\vec{b}$, but for many different $\vec{b}$s.
- That is why LU decomposition saves us from repeating lots of computation in comparison with Gaussian elimination on $[A|\vec{b}]$
- Complexity of LU decomposition is $O(n)$. Complexity of inverse is $O(n^2)$.

7 Vector spaces and subspaces

7.1 Vector spaces

Definition: A vector space is a non-empty set V of vectors, which may be **added** and **scaled**.

- Axioms ($u, v, w \in V$, $c, d \in \mathbb{R}$).
  1. $u + v \in V$
  2. $u + v = v + u$
  3. $(u + v) + w = u + (v + w)$
  4. $\exists \mathbf{0} \in V$, s.t, $u + \mathbf{0} = u \; \forall u \in V$
  5. $\exists u \in V$, s.t, $u + (-u) = \mathbf{0}$
  6. $cu \in V$
  7. $c(u + v) = cu + cv$
  8. $(c + d)u = cu + du$
  9. $(cd)u = c(du)$
  10. $1 \cdot u = u$

7.2 Subspaces

Definition: A subset W of a vector space V is a subspace, if W is itself a vector space. Since the rules like associativity, commutativity, and distributivity stll hold, we only need to check the followings.

$$W \subseteq V \text{ is a subspace of } V \text{ if}$$

1. W contains the zero vector
2. W is closed under addition
3. W is closed under scaling

Theorem : If $A \in \mathbb{R}^{m \times n}$, then $Nul(A)$ is a subspace of $\mathbb{R}^n$

- $\vec{0} \in Nul(A)$ since, $A\vec{0} = \vec{0}$
- $\vec{x}, \vec{y} \in Nul(A) \to A\vec{x} = A\vec{y} = \vec{0}$
  Then $\vec{x} + \vec{y} \in Nul(A)$ since $A(\vec{x} + \vec{y}) = A\vec{x} = A\vec{y} = \vec{0}$
- $\vec{x} \in Nul(A) \to \gamma\vec{x} \in Nul(A)$
  since $\gamma A\vec{x} = \gamma\vec{0} = \vec{0}$

7.3 Linear independence of matrix columns

Each linear dependence relation of columns of A corresponds to a nontrival solution to Ax=0

Theorem : Let A be an $m \times n$ matrix. The columns of A are linearly independent

- $A\vec{x} = \vec{0}$ has only the solution $\vec{x} = \vec{0}$.
- $Nul(A) = \{\vec{0}\}$
- A has n pivots.

7.4 A basis of a vector space

Definition : A set of vectors $\{\vec{v_1}, \cdots, \vec{v_p}\}$ in V is a basis of V if

1. $V = span\{\vec{v_1}, \cdots, \vec{v_p}\}$
2. $\{\vec{v_1}, \cdots, \vec{v_p}\}$ are linearly independent

Theorem : If S is a basis of a vector space V, then every vector in V has exactly one representation as a linear combinations of elements of S.

Theorem : If V has a basis with n elements, then every set of m vectors which has more then n elements is linearly dependent. $(m > n)$

Theorem : If V has a basis with n elements, then every set of vectors with fewer than n elements does not span V

Theorem : If V has a basis with n elements, then all bases of V have the same number of elements.

Definition : The number of elements in a basis of a vectorspace is called dimension.

Theorem : Suppose that V has dimension d,

1. A set of d vectors in V are a basis if they span V.
2. A set of d vectors in V are a basis if they are linearly independent.

7.5 The Four fundamental subspaces

Definition :
The row space of A is $Col(A^T)$.
The left nullspace of A is $Nul(A^T)$

Theorem : Fundamental theorem of linear algebra(FTLA)
Let $A \in \mathbb{R}^{m\times n}$ of rank r.

- $dimCol(A) = r$
- $dimCol(A^T) = r$
- $dimNul(A) = n - r$
- $dimNul(A^T) = m - r$

8 Linear transforms

Definition : A map $T : V \mapsto W$ is a linear transform if
$T(c\vec{x} + d\vec{y}) = cT(\vec{x}) + dT(\vec{y})$, $\vec{x}, \vec{y} \in V$, $c, d \in \mathbb{R}$

8.1 Representing linear maps by matrices

Definition : Let $\vec{x_1}, \cdots, \vec{x_n}$ be a basis for V, and $\vec{y_1}, \cdots, \vec{y_m}$ be a basis for W. The matrix representing T with respect to these bases,

- has n column (one for each $\vec{x_j}$)
- the j-th column has m entries $a_{1j}, a_{2j}, \cdots, a_{mj}$ determined by

$$T(\vec{x_j}) = a_{1j}\vec{y_1} + a_{2j}\vec{y_2} + \cdots + a_{mj}\vec{y_m}$$

9 Orthogonality

Definition : The inner product of $\vec{v}, \vec{w} \in \mathbb{R}^n$

$$\vec{v} \cdot \vec{w} = \vec{v}^T\vec{w} = v_1w_1 + \cdots + v_nw_n$$

Definition : The norm of a vector $\vec{v} \in \mathbb{R}^n$ is

$$\|\vec{v}\| \triangleq \sqrt{\vec{v} \cdot \vec{v}}$$

The distance between $\vec{v}, \vec{w}$ is

$$dist(\vec{v}, \vec{w}) \triangleq \|\vec{v} - \vec{w}\|$$

Definition : The $\vec{v}, \vec{w} \in \mathbb{R}^n$ are orthogonal if

$$\vec{v} \cdot \vec{w} = 0$$

Theorem: Suppose that $\vec{v_1}, \cdots, \vec{v_n}$ are non zero, and pairwise orthogonal. Then, they are independent.

Definition : Let W be a subspace of $\mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$

- $\vec{v}$ is orthogonal to W, if $\vec{v} \cdot \vec{w} = 0$ for all $\vec{w} \in W$
- Another subspace V is orthogonal to W, if every vector in V is orthogonal to W.
- The orthogonal complement of $W^{\perp}$ of all vectors that are orthogonal to W

Theorem:
$Nul(A)$ is orthogonal to $Col(A^{\perp})$
$Nul(A^{\perp})$ is orthogonal to $Col(A)$

9.1 Orthogonal Bases

Definition : A basis $\vec{v_1}, \cdots, \vec{v_n}$ of vector space V is an orthgonal basis if the vectors are pairwise orthogonal.

9.2 Orthogonal Projection

Definition : The orthogonal projection of vector $\vec{x}$ onto vector $\vec{y}$ is

$$\hat{x} \triangleq (\frac{\vec{x} \cdot \vec{y}}{\vec{y} \cdot \vec{y}})\vec{y}$$

9.3 Orthogonal Projection on subspaces

Theorem : Let W be a subspace of $\mathbb{R}^n$. Then, each $\vec{x} \in \mathbb{R}^n$ can be uniquely written as

$$\vec{x} = \hat{x} + \hat{x}^{\perp}$$
$$\text{where } \hat{x} \in W \text{and } \hat{x}^{\perp} \in W^{\perp}$$

Definition : Let $\vec{v_1}, \cdots, \vec{v_n}$ be an orthogonal basis of $W \in \mathbb{R}^n$
Then the projection map $\pi_w : \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$\vec{x} \mapsto \hat{x} = \sum_{i=1}^{n} (\frac{\vec{x} \cdot \vec{v_i}}{\vec{v_i} \cdot \vec{v_i}})\vec{v_i}$$

and π_w is linear.
The matrix P representing π_w w, r, t the standard basis is the projection matrix.

9.4 Normal equation

Theorem : $\hat{x}$ is a least square solution of $A\vec{x} = \vec{b}$

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

$\hat{x}$ is a LS solution of $A\vec{x} = \vec{b}$

- $(A\vec{x} - \vec{b})$ is as small as possible
- $(A\vec{x} - \vec{b}) \perp Col(A)$
- $(A\vec{x} - \vec{b}) \in NUL(A^T)$
- $A^T(A\vec{x} - \vec{b}) = 0$
- $A^T A\hat{x} = A^T\vec{b}$

10 QR decomposition

Let A be an $m \times n$ matrix of rank n.
Then, the QR decomposition $A = QR$
where $Q \in \mathbb{R}^{m \times n}$ with orthonormal columns,
R is upper-triangular, $n \times n$, invertible.

$\hat{x}$ is a LS solution of $A\vec{x} = \vec{b} \Longleftrightarrow R\hat{x} = Q^T\vec{b}$

11 Determinant

Definition : The determinant is characterized by

- the normalization $det(I) = 1$
- and how it is affected by elementary row operations
  - (Replacement) Does not change the Det.
  - (Interchange) Reverse the sign of the Det.
  - (Scaling) Multiplies the Det with s.

12 Eigenvectors

Definition : An eigenvector of A is a nonzero $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$ for some scalar λ

Theorem : IF $\vec{x_1}, \cdots, \vec{x_n}$ are eigenvectors of A corresponding to different eigenvalues, then they are independent.